

# Complexity of Robust Orbit Problems for Torus Actions and the *abc*-conjecture

Peter Bürgisser ✉

Institute of Mathematics, Technische Universität Berlin, Germany.

Mahmut Levent Doğan ✉

Institute of Mathematics, Technische Universität Berlin, Germany.

Visu Makam ✉

Radix Trading, Amsterdam.

Michael Walter ✉

Faculty of Computer Science, Ruhr-Universität Bochum, Germany.

Avi Wigderson ✉

School of Mathematics, Institute for Advanced Study, Princeton.

---

## Abstract

When a group acts on a set, it naturally partitions it into orbits, giving rise to *orbit problems*. These are natural algorithmic problems, as symmetries are central in numerous questions and structures in physics, mathematics, computer science, optimization, and more. Accordingly, it is of high interest to understand their computational complexity. Recently, [16] gave the first polynomial-time algorithms for orbit problems of *torus actions*, that is, actions of commutative continuous groups on Euclidean space. In this work, motivated by theoretical and practical applications, we study the computational complexity of *robust* generalizations of these orbit problems, which amount to *approximating* the distance of orbits in  $\mathbb{C}^n$  up to a factor  $\gamma \geq 1$ . In particular, this allows deciding whether two inputs are approximately in the same orbit or far from being so. On the one hand, we prove the NP-hardness of this problem for  $\gamma = n^{\Omega(1/\log \log n)}$  by reducing the closest vector problem for lattices to it. On the other hand, we describe algorithms for solving this problem for an approximation factor  $\gamma = \exp(\text{poly}(n))$ . Our algorithms combine tools from invariant theory and algorithmic lattice theory, and they also provide group elements witnessing the proximity of the given orbits (in contrast to the algebraic algorithms of prior work). We prove that they run in polynomial time if and only if a version of the famous number-theoretic *abc*-conjecture holds – establishing a new and surprising connection between computational complexity and number theory.

**2012 ACM Subject Classification** Computing methodologies → Algebraic algorithms; Computing methodologies → Combinatorial algorithms; Theory of computation → Algebraic complexity theory

**Keywords and phrases** computational invariant theory, geometric complexity theory, orbit problems, *abc*-conjecture, closest vector problem

**Funding** *Peter Bürgisser*: Supported by the ERC under the European Union’s Horizon 2020 research and innovation programme (grant agreement no. 787840).

*Mahmut Levent Doğan*: Supported by the ERC under the European Union’s Horizon 2020 research and innovation programme (grant agreement no. 787840).

*Visu Makam*: Partially supported by NSF Grant CCF-1900460 and the University of Melbourne.

*Michael Walter*: Supported by the European Union (ERC, SYMOPTIC, 101040907), by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy - EXC 2092 CASA - 390781972, by the BMBF (QuBRA, 13N16135), and by the Dutch Research Council (NWO grant OCENW.KLEIN.267).

*Avi Wigderson*: Supported by the NSF Grant CCF-1900460.

**Acknowledgements** The authors thank Matías Bender, Alperen A. Ergür, Jonathan Leake, and Philipp Reichenbach for productive discussions.

## 1 Introduction

Computational invariant theory was a central topic of 19th century algebra, see the historical account in [57]. In the second half of the 20th century, structural progress was made through Mumford’s invention of geometric invariant theory [73] and computational progress came through the theory of Gröbner bases, see [83]. More recently, it was realized that algorithmic questions in invariant and representation theory are deeply connected with the core complexity questions of P vs. NP and VP vs. VNP.

On the one hand, Mulmuley and Sohoni’s Geometric Complexity Theory (GCT) [71] highlights the inherent symmetries of complete problems of these complexity classes. This was the starting point of several specific invariant theoretic and representation theoretic attacks on the VP vs. VNP questions, which has led to many new questions, techniques, and much faster algorithms: for example, see [70, 31, 15, 63].

The other connection is through the work of Impagliazzo and Kabanets [52], which uses Valiant’s completeness theory for VP and VNP to construct efficient deterministic algorithms for the basic PIT (Polynomial Identity Testing) problem. This problem, again thanks to Valiant’s completeness result, has natural symmetries which resemble basic problems of invariant theory. Major progress was recently achieved in this direction in [40, 33, 50, 51, 25]. This work was followed by [26, 43, 49, 34, 19, 18, 20, 17]; we refer to [17] for a description of the state-of-art.

In addition to these fundamental connections to computational complexity, orbit problems appear in many other areas, for example in physics, since the symmetries are often given by the actions of Lie groups, see [6, 65, 39]. Recently, a connection of algebraic statistics to orbit problems was discovered [4, 5]: the problem of finding maximum likelihood estimates in certain statistical models and the ‘flip flop’ algorithm from statistics is precisely related to the invariant theory and the scaling algorithms of [33, 19]; see also [32, 27, 28]. In particular, the setting of log-linear models in [5] relates precisely to the setting of this work.

### 1.1 Background and high-level summary of results

In this paper, we study actions of the commutative group  $T := (\mathbb{C}^\times)^d$  and its subgroup  $K := (S^1)^d$  on vector spaces  $V := \mathbb{C}^n$ .<sup>1</sup> The group  $T$  is called an *algebraic torus* of rank  $d$ , and the subgroup  $K$  is known as a *compact torus*. The action of either group partitions the vector space  $V$  into *orbits*. That is, for a vector  $v \in V$ , we consider the set of all vectors reachable from  $v$  by applying group elements:

$$\mathcal{O}_v := \{t \cdot v \mid t \in T\}, \quad \mathcal{C}_v := \{k \cdot v \mid k \in K\}.$$

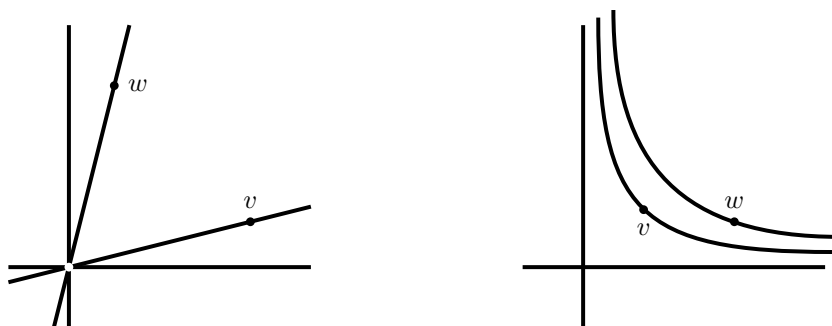
The sets  $\mathcal{O}_v$  and  $\mathcal{C}_v$  are called the *T-orbit* and *K-orbit* of  $v$ , respectively. As  $K$  is compact, so is  $\mathcal{C}_v$ . However, the *T-orbit*  $\mathcal{O}_v$  is in general not closed and hence it is natural to also consider its *orbit closure*  $\overline{\mathcal{O}_v}$ .<sup>2</sup> See Figure 1 for two illustrative examples. While torus actions are very well-understood from a structural perspective, they give rise to interesting and challenging computational problems:

► **Problem 1.1.** *Given a torus action on a vector space  $V$  and  $v, w \in V$ :*

1. **Orbit equality:** *Decide if  $\mathcal{O}_v = \mathcal{O}_w$ .*

<sup>1</sup> Here,  $\mathbb{C}^\times$  denotes the multiplicative group of the nonzero complex numbers and  $S^1 := \{z \in \mathbb{C} \mid |z| = 1\}$ .

<sup>2</sup> We take the closure with respect to either the Euclidean or the Zariski topology, as they coincide here [72, I §10].



■ **Figure 1** (a) Action of  $T = \mathbb{C}^\times$  on  $V = \mathbb{C}^2$  given by  $t \cdot (x, y) = (tx, ty)$ : All orbits are lines through the origin, with the origin excluded. Thus, all orbit closures intersect in the origin. (b) Action of  $T = \mathbb{C}^\times$  on  $V = \mathbb{C}^2$  given by  $t \cdot (x, y) = (tx, t^{-1}y)$ : All orbits of vectors with non-zero coordinates  $x, y \neq 0$  are closed (they are hyperbolas).

2. **Orbit closure intersection:** Decide if  $\overline{\mathcal{O}_v} \cap \overline{\mathcal{O}_w} \neq \emptyset$ .
3. **Orbit closure containment:** Decide if  $\mathcal{O}_w \subseteq \overline{\mathcal{O}_v}$ .
4. **Compact orbit equality:** Decide if  $\mathcal{C}_v = \mathcal{C}_w$ .

These problems capture and relate to a broad class of “isomorphism”, “classification”, or “transformation” problems. Their computational complexity was determined only very recently: [16] found that all four problems can be decided in polynomial time. This is perhaps surprising, because for actions of noncommutative groups these problems are believed to differ in their computational complexity and, e.g., Problem 1.1 (3) is known to be hard for some actions [10].

In many applications, e.g., the ones in statistics or physics mentioned above, the vectors  $v, w$  are not given exactly, but only up to a certain error. If this is the case, the orbit equality problem has to be replaced by a robust generalization, which meaningfully tolerates small perturbations in the input.

The goal of this paper is to define such robust generalizations of the orbit equality problem and to investigate their computational complexity. More precisely, we aim to *approximate the distance between two orbits* given by vectors  $v, w \in \mathbb{C}^n$  up to an approximation factor  $\gamma > 1$ . In particular, this allows deciding whether two points are approximately in the same orbit or far from being so.

At first glance, this appears to simplify the situation, since we no longer need to distinguish between an orbit and its closure. However, surprisingly, the picture arising is far more intricate than for the problems that were dealt with in [16]. On the one hand, we prove NP-hardness when  $\gamma = n^{\Omega(1/\log \log n)}$ . On the other hand, using a combination of tools from invariant theory and algorithmic lattice theory, we give algorithms that achieve an approximation factor  $\gamma = \exp(\text{poly}(n))$  and run in polynomial time if and only if a version of the well-known number-theoretic *abc*-conjecture holds! Our algorithms also return a witness  $t \in T$  such that the distance between  $t \cdot v$  and  $w$  is at most  $\gamma$  times the distance between the orbits. We give precise technical statements below.

We note that our results are not the first examples of an open problem in number theory having an impact on the complexity of algorithms. As is widely known, the *generalized (or extended) Riemann hypothesis* has been used, e.g., for testing primality in polynomial time [68] (an unconditional deterministic polynomial time algorithm was later given in [1]), the containment of knottedness in NP [58], or the containment of Hilbert’s Nullstellensatz

in the polynomial hierarchy [56]. To our best knowledge, however, our work, together with [30], constitutes one of the first examples that the *abc*-conjecture has some bearing on the performance of numerical algorithms.

In the remainder of this introduction we first recap the structure of torus actions and define the input model and complexity parameters used in all our computational problems, algorithms, and results (Section 1.2). We then state the computational problems (Section 1.3) and discuss our hardness and algorithmic results (Sections 1.4 and 1.5). Next, we explain how the “separation hypotheses” that ensure that our algorithms run in polynomial time are intimately related to well-known variants of the *abc*-conjecture in number theory (Section 1.6). Finally, we sketch the proof idea of our lattice lifting result (Section 1.7), and we conclude with open problems for future research (Section 1.9).

## 1.2 Torus actions: structure, input model, parameters

Every rational action of  $T = (\mathbb{C}^\times)^d$  or  $K = (S^1)^d$  on a finite-dimensional complex vector space  $V = \mathbb{C}^n$  can be parameterized by a *weight matrix*  $M \in \mathbb{Z}^{d \times n}$ , as follows: For  $t = (t_1, t_2, \dots, t_d) \in T$  and  $v = (v_1, v_2, \dots, v_n) \in V$ ,

$$t \cdot v := \left( v_1 \prod_{i=1}^d t_i^{M_{i1}}, v_2 \prod_{i=1}^d t_i^{M_{i2}}, \dots, v_n \prod_{i=1}^d t_i^{M_{in}} \right). \quad (1.1)$$

See Figure 1 for two examples with weight matrix  $M = \begin{pmatrix} 1 & 1 \end{pmatrix}$  and  $M = \begin{pmatrix} 1 & -1 \end{pmatrix}$ , respectively.

Our computational problems and algorithms take as their input torus actions as well as vectors  $v, w \in \mathbb{C}^n$ . The former are encoded in terms of the weight matrix  $M$ , with each entry represented in binary. The latter are assumed to be vectors  $v, w \in \mathbb{Q}(i)^n$ , i.e., their components are Gaussian rationals (complex numbers of the form  $q + ir$  with  $q, r \in \mathbb{Q}$ ) and given by encoding the numerators and denominators of the real and imaginary parts in binary.

Throughout the paper, we denote by  $B$  the maximum bit-length of the entries of  $M$ , and by  $b$  the maximum bit-length of the components of  $v$  and  $w$ , respectively. The parameters  $B$  and  $b$  will have different effects in the bounds.

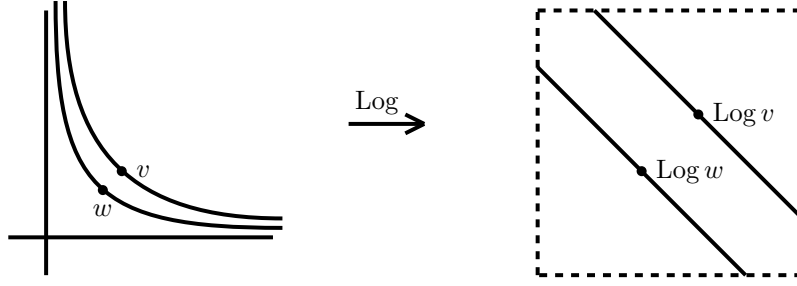
## 1.3 Robust orbit problems: definitions

We will now define precisely the computational problems that we address in this paper. Given a torus action on a vector space  $V = \mathbb{C}^n$  and two vectors  $v, w \in V$ , we wish to approximate the distance of the orbit under either the action of the algebraic torus  $T$  or the compact torus  $K$ .

We start with the latter. Because  $K = (S^1)^d$  is compact, so are its orbits. Thus, if two  $K$ -orbits are distinct then they must have a finite distance with respect to any norm on  $V$ . We find it natural to use the Euclidean norm  $\|v\| := \sqrt{\sum_{j=1}^n |v_j|^2}$ . Given two vectors  $v, w \in \mathbb{C}^n$ , we denote their Euclidean distance by  $\text{dist}(v, w) := \|v - w\|$  and we extend this notation to arbitrary subsets  $A, B \subseteq \mathbb{C}^n$  by setting  $\text{dist}(A, B) := \inf_{a \in A, b \in B} \|a - b\|$ . Then we are interested in the following approximation problem, where the approximation factor  $\gamma$  is a parameter that may depend on  $d, n, b$ , or  $B$ .

► **Problem 1.2.** *The compact orbit distance approximation problem with approximation factor  $\gamma \geq 1$  is defined as follows: Given  $v, w \in (S^1)^n$ , compute a number  $D$  such that*

$$\text{dist}(\mathcal{C}_v, \mathcal{C}_w) \leq D \leq \gamma \text{dist}(\mathcal{C}_v, \mathcal{C}_w).$$



■ **Figure 2** The logarithm maps the “hyperbolic” orbits (left) into parallel “lines” (right), which have a finite Euclidean distance. We define the logarithmic distance of the former as the Euclidean distance of the latter.

We now consider the algebraic torus  $T = (\mathbb{C}^\times)^d$ . Since it is not compact, in general neither are its orbits and hence two distinct orbits need not have any finite distance – in other words, we can have  $\mathcal{O}_v \neq \mathcal{O}_w$  but  $\text{dist}(\mathcal{O}_v, \mathcal{O}_w) = 0$ . This can be due to two phenomena: (a) when orbits are not closed, they can still intersect in their closure; (b) even if orbits are closed, they may still become arbitrarily close at infinity. See Figure 1 for an illustration. The first phenomenon is natural: if  $\overline{\mathcal{O}_v} \cap \overline{\mathcal{O}_w} \neq \emptyset$  then it can be natural to define the distance to be zero. The second phenomenon however is undesirable – clearly we would like to be able to tell apart, e.g., the two hyperbolas in Figure 1 (b)! To overcome this problem, we will instead consider a *logarithmic distance*, which always assigns a positive distance between distinct orbits, even when their closures intersect.

For simplicity, we assume that  $v, w$  have non-zero coordinates, that is, they are elements of  $(\mathbb{C}^\times)^n \subseteq V$ . The key idea is to “linearize” or “flatten” the group action by using the (coordinatewise) complex logarithm map:

$$\text{Log}: (\mathbb{C}^\times)^n \rightarrow \mathbb{C}^n / 2\pi i \mathbb{Z}^n, \quad v \mapsto (\log v_1, \log v_2, \dots, \log v_n), \quad (1.2)$$

which is a group isomorphism. This is natural because it converts the multiplicative action of  $T$  into an additive one. Indeed, one finds from (1.1) that  $\text{Log}$  maps any  $T$ -orbit  $\mathcal{O}_v$  to the image of the affine subspace  $\text{im}(M^T) + \text{Log}(v)$  under the quotient map  $\mathbb{C}^n \rightarrow \mathbb{C}^n / 2\pi i \mathbb{Z}^n$ . For distinct orbits these will have a finite Euclidean distance because the corresponding subspaces are parallel (see Figure 2). This motivates using the following distance: Given  $v, w \in (\mathbb{C}^\times)^n$ , we define their *logarithmic distance* by

$$\delta_{\log}(v, w) := \text{dist}(\text{Log}(v) + 2\pi i \mathbb{Z}^n, \text{Log}(w) + 2\pi i \mathbb{Z}^n) := \min_{\alpha \in 2\pi i \mathbb{Z}^n} \|\text{Log}(v) - \text{Log}(w) - \alpha\|, \quad (1.3)$$

using the natural Euclidean distance on the quotient  $\mathbb{C}^n / 2\pi i \mathbb{Z}^n$ , and we extend this to arbitrary subsets  $A, B \subseteq (\mathbb{C}^\times)^n$  as above. In particular, we have  $\delta_{\log}(\mathcal{O}_v, \mathcal{O}_w) = 0$  if and only if  $\mathcal{O}_v = \mathcal{O}_w$ , which is exactly the property that we wanted. Thus we arrive at the following approximation problem, where the approximation factor  $\gamma$  may depend on  $d, n, b$ , or  $B$  as above.<sup>3</sup>

<sup>3</sup> One can also use the metric  $\delta_{\log}$  in definition of Problem 1.2 in place of  $\text{dist}$ . However, this is essentially equivalent, since for  $v, w \in (S^1)^n$  the two metrics are related by  $\frac{2}{\pi} \delta_{\log}(v, w) \leq \text{dist}(v, w) \leq \delta_{\log}(v, w)$ , see Section 3.

► **Problem 1.3.** *The algebraic orbit distance approximation problem with approximation factor  $\gamma \geq 1$  is defined as follows: Given  $v, w \in (\mathbb{C}^\times)^n$ , compute a number  $D$  such that*

$$\delta_{\log}(\mathcal{O}_v, \mathcal{O}_w) \leq D \leq \gamma \delta_{\log}(\mathcal{O}_v, \mathcal{O}_w).$$

#### 1.4 Robust orbit problems: hardness

The appearance of a lattice in the above suggests that there might be some hardness lurking behind this problem. In fact, we find that *both* orbit distance approximation problems are NP-hard for a sufficiently small approximation factor. We will prove the following in Section 6 by showing that there is a polynomial time reduction from the *closest vector problem* (CVP) to Problems 1.2 and 1.3.

► **Theorem 1.4.** *There is a constant  $c > 0$  such that Problems 1.2 and 1.3 for  $\gamma = n^{c/\log \log n}$  are NP-hard.*

We note that a solution to Problem 1.2 allows distinguishing between the cases  $\text{dist}(\mathcal{C}_v, \mathcal{C}_w) \leq \varepsilon$  and  $\text{dist}(\mathcal{C}_v, \mathcal{C}_w) \geq \gamma\varepsilon$  on input  $\varepsilon$ , and similarly for Problem 1.3. The CVP problem has recently attracted significant interest due to its relevance to *lattice-based cryptosystems*, which are conjectured to be secure against quantum computers [37, 47, 48, 35].

Interestingly, our reduction from CVP is not straightforward, but rather uses a quantitative lattice lifting result in the spirit of [23] that might be of independent interest. We sketch the main idea for the  $K$ -action and for  $\delta_{\log}$  instead of the Euclidean distance (which by Footnote 3 makes no difference). Let  $\theta := \frac{1}{2\pi i} \text{Log}(v)$  and  $\phi := \frac{1}{2\pi i} \text{Log}(w)$ . Then,

$$\frac{1}{2\pi} \delta_{\log}(\mathcal{C}_v, \mathcal{C}_w) = \text{dist}(P(\theta - \phi), P(\mathbb{Z}^n)), \quad (1.4)$$

where  $P$  denotes the orthogonal projection from  $\mathbb{R}^n$  onto the orthogonal complement of the row space of  $M$ . Note that the right-hand side quantity amounts to a CVP for the lattice  $P(\mathbb{Z}^n)$ . Our lattice lifting result states that, up to scaling, *every* full-dimensional lattice  $\mathcal{L}$  can be obtained as the orthogonal projection of a cubic lattice  $\mathbb{Z}^n$ , where  $n$  is not much larger than  $\dim(\mathcal{L})$ . Moreover, this projection can be computed in polynomial time. This yields the desired reduction from CVP.

► **Theorem 1.5 (Lattice lifting).** *Suppose  $\mathcal{L} \subset \mathbb{R}^m$  is a lattice of rank  $m$  given by a generator matrix  $G \in \mathbb{Z}^{m \times m}$ , i.e.,  $\mathcal{L} = G(\mathbb{Z}^m)$ . Then we can, in polynomial time, compute  $n \geq m$ , a scaling factor  $s \in \mathbb{Z}_{>0}$ , and an orthonormal basis  $v_1, v_2, \dots, v_n \in \mathbb{Q}^n$  such that*

$$\mathcal{L} = sP(\mathcal{L}'), \quad \text{where } \mathcal{L}' := \mathbb{Z}v_1 + \mathbb{Z}v_2 + \dots + \mathbb{Z}v_n$$

and  $P: \mathbb{R}^n \rightarrow \mathbb{R}^m$  denotes the orthogonal projection onto the first  $m$  coordinates. Moreover, we may assume that  $n = O(m \log m + m \log \langle G \rangle)$ , where  $\langle G \rangle$  denotes the bit-length of  $G$ .

To the best of our knowledge, Theorem 1.5 was not previously observed in the literature. We sketch its proof in Section 1.7 below and give the full proof in Section 6.2.

#### 1.5 Robust orbit problems: algorithms

The picture changes markedly if we allow for larger approximation ratios  $\gamma$ , where we recall that  $B$  denotes the maximum bit-length of the entries of  $M$ . Roughly speaking, our intuitive result is the following:

*If the distance between any two distinct orbits is no smaller than  $\exp(-\text{poly}(d, n, B, b))$ , then these distances can be approximated to  $\gamma = \exp(\text{poly}(d, n, B))$  in polynomial time.*

To state this precisely, let us define the *compact* and the *algebraic separation parameters* as

$$\text{sep}_K(d, n, B, b) := \min_{M, v, w} \text{dist}(\mathcal{C}_v, \mathcal{C}_w) \quad \text{and} \quad \text{sep}_T(d, n, B, b) := \min_{M, v, w} \delta_{\log}(\mathcal{O}_v, \mathcal{O}_w), \quad (1.5)$$

respectively, where the minima are taken over all weight matrices  $M \in \mathbb{Z}^{d \times n}$  with entries of maximum bit-length at most  $B$  and over all vectors  $v, w \in (\mathbb{Q}(i)^\times)^n$  with components of bit-length at most  $b$  such that  $v$  and  $w$  are in distinct orbits.

► **Separation Hypothesis 1.6** (for the compact torus).  $\text{sep}_K(d, n, B, b) \geq \exp(-\text{poly}(d, n, b, B))$ .

► **Separation Hypothesis 1.7** (for the algebraic torus).  $\text{sep}_T(d, n, B, b) \geq \exp(-\text{poly}(d, n, b, B))$ .

The latter hypothesis is readily seen to imply the former (this follows essentially from Footnote 3, see Corollary 3.9). While both hypotheses appear geometric in nature, it turns out that they are intimately related to well-known quantitative versions of the *abc*-conjecture in number theory. We will explain this connection in Section 1.6. Assuming these hypotheses, we can solve the orbit distance approximation problems with an exponential approximation factor:

► **Theorem 1.8.** *If Separation Hypothesis 1.6 holds, then there is a polynomial-time algorithm that solves Problem 1.2 for an approximation factor  $\gamma = \exp(\text{poly}(d, n, B))$ . Similarly, if Separation Hypothesis 1.7 holds, then there is a polynomial-time algorithm that solves Problem 1.3 for  $\gamma = \exp(\text{poly}(d, n, B))$ .*

In fact, we show in Section 5 without any hypotheses that for  $\gamma = \exp(\text{poly}(d, n, B))$ , there is an algorithm solving Problem 1.3 in time  $O(\text{poly}(d, b, n, B) \log(\text{sep}_T^{-1}(d, n, B, b)))$ , and similarly for Problem 1.2. This algorithm runs in polynomial time if and only if the separation hypothesis is true, see Remark 5.11. Our algorithm also solves the corresponding *search problem* (see Theorem 5.4): When  $\mathcal{O}_v \neq \mathcal{O}_w$ , it finds a group element  $t \in T$  such that

$$\delta_{\log}(t \cdot v, w) \leq \gamma \delta_{\log}(\mathcal{O}_v, \mathcal{O}_w). \quad (1.6)$$

In particular,  $t$  can serve as a witness of the approximate distance of the orbits. In general, the coordinates of any  $t$  satisfying Equation (1.6) may require a superpolynomial bit-length. Accordingly, our algorithm will output instead a vector  $x \in \mathbb{Q}(i)^d$  such that  $t = \text{Exp}(x)$ . When  $\mathcal{O}_v = \mathcal{O}_w$ , then it will in general not be possible to output such a group element, since  $t \cdot v = w$  will in general not have a rational solution (neither in  $t$  nor in  $x$ ).

To explain the main idea behind proving Theorem 1.8 and motivate the separation hypothesis, we recall the algorithm for solving the orbit equality problem in [16]. For simplicity, we consider vectors  $v, w \in (\mathbb{C}^\times)^n$  with closed  $T$ -orbits. The algorithm in [16] first computes a basis  $\mathcal{B}$  of the lattice  $\mathcal{K} := \{\alpha \in \mathbb{Z}^n \mid M\alpha = 0\}$  in polynomial time. Each basis vector  $\alpha \in \mathcal{B}$  determines an invariant function (Laurent monomials) of the form

$$f_\alpha(x) := x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}.$$

Then,  $\mathcal{O}_v = \mathcal{O}_w$  if and only if  $v$  and  $w$  take the same value on all these functions [16, Corollary 5.2]. Even though the  $\alpha$  have polynomial bit-length, testing whether  $f_\alpha(v) = f_\alpha(w)$  is not obvious, as the bit-length of the evaluations can be exponentially large in the input size. Nevertheless, it is possible to determine equality of these numbers in polynomial time without actually computing them, see [16, Proposition 5.5].



When dealing with the problem of approximating orbit distances, one may use similar ideas. One finds that one now needs to test whether  $f_\alpha(v)$  and  $f_\alpha(w)$  are *close* (rather than equal), for a suitable choice of a metric. This turns out to be substantially more difficult than the equality testing,<sup>4</sup> and indeed it must be by the hardness results in Section 1.4. To solve it with an exponential approximation factor we can proceed as follows: Let  $H \in \mathbb{Z}^{k \times n}$  denote the matrix with rows the vectors in  $\mathcal{B}$ . We prove in Section 3 that the logarithmic distance between the orbits  $\mathcal{O}_v$  and  $\mathcal{O}_w$  can be approximated by  $\|H \operatorname{Log}(v) - H \operatorname{Log}(w)\|$ , up to a factor  $\sigma_{\max}(H)/\sigma_{\min}(H) = \exp(\operatorname{poly}(d, n, B))$ . Here,  $\sigma_{\max}(H), \sigma_{\min}(H)$  denote the largest and smallest singular values of  $H$ , respectively. It remains to approximate  $\|H \operatorname{Log}(v) - H \operatorname{Log}(w)\|$  with relative error. The key challenge is in computing rational approximations of the logarithms. This can be done by standard algorithms, but we have to make sure that polynomially many bits of accuracy are sufficient. This is guaranteed by the separation hypothesis! The idea of relating the distance between two orbits with the difference between the values that the invariants take on these orbits, and approximating the distance between the orbits in this way is inspired by the invariant theory and it will be important later when we explain the close relationship between the robust orbit problems and the *abc*-conjecture. On the other hand, we will also develop an alternative algorithm to the robust orbit problems in Section 6.1 by providing a polynomial time reduction to CVP under the assumption that Separation Hypothesis 1.7 is true by using Equation (1.4).

Finally, we study the complexity of the orbit distance approximation problems when  $n$  is fixed. In this situation, Separation Hypotheses 1.6 and 1.7 are true, as we will discuss in Section 1.6, and hence there are polynomial-time algorithms for Problems 1.2 and 1.3 with  $\gamma = \exp(\operatorname{poly}(d, B))$ . However, we can in fact choose  $\gamma$  to be much smaller:

► **Theorem 1.9.** *For fixed  $n$ , Problem 1.2 can be solved in polynomial time for the approximation factor  $\gamma = 2$ .*

► **Theorem 1.10.** *For fixed  $n$ , and any fixed  $\gamma > 1$ , Problem 1.3 can be solved in polynomial time. More specifically, there is a polynomial-time algorithm that on input  $M \in \mathbb{Z}^{d \times n}$ ,  $v, w \in \mathbb{Q}(i)^n$ , and  $\varepsilon > 0$  outputs a number  $D$  such that  $\delta_{\log}(\mathcal{O}_v, \mathcal{O}_w) \leq D \leq (1 + \varepsilon) \delta_{\log}(\mathcal{O}_v, \mathcal{O}_w)$ .*

We prove these results at the end of Section 6.1.

## 1.6 Separation hypotheses and the *abc*-conjecture

The famous *abc*-conjecture by Oesterlé and Masser [75] states that for every  $\varepsilon > 0$  there exists  $\kappa_\varepsilon > 0$  such that, if  $a$  and  $b$  are coprime positive integers and  $c = a + b$ , then

$$c < \kappa_\varepsilon \operatorname{rad}(abc)^{1+\varepsilon},$$

where  $\operatorname{rad}(abc)$  is the *radical* of  $abc$ , i.e., the product of all primes dividing  $abc$ . Motivated by an earlier conjecture of Szpiro, the *abc*-conjecture is one of the most important open problems in number theory, due to its many consequences [38, 87].

Baker [8] observed that the *abc*-conjecture is intimately related to lower bounds for linear forms in logarithms. Let  $v_1, v_2, \dots, v_n$  be positive integers and  $e_1, e_2, \dots, e_n$  be integers such that the product  $v_1^{e_1} v_2^{e_2} \cdots v_n^{e_n} \neq 1$  is different from one. How close to 1 can such a product be? Equivalently, how close can the linear combination of logarithms,

$$\Lambda(v, e) := e_1 \log v_1 + \dots + e_n \log v_n, \tag{1.7}$$

<sup>4</sup> Just testing an inequality  $f_\alpha(v) \geq f_\alpha(w)$  in polynomial time is only known assuming a certain strengthening of the *abc*-conjecture; see [30].



be to zero? We note that if  $|v_1^{e_1} \dots v_n^{e_n} - 1| \leq \frac{1}{2}$ , then

$$\frac{1}{2} |\Lambda(v, e)| \leq |v_1^{e_1} \dots v_n^{e_n} - 1| \leq 2 |\Lambda(v, e)|,$$

so lower bounds for  $|\Lambda(v, e)|$  and for  $|v_1^{e_1} \dots v_n^{e_n} - 1|$  are equivalent; see Equation (4.4). Improving work by Baker and Wüstholz [88], Matveev [66] proved that if  $\Lambda(v, e) \neq 0$ , then

$$|\Lambda(v, e)| \geq e^{-B \cdot b^{O(n)}}. \quad (1.8)$$

Here and below  $B$  and  $b$  denote upper bounds on the bit-length of the  $e_j$  and  $v_j$ , respectively. Note that for fixed  $n$ , this bound is polynomial in  $B$  and  $b$ . Furthermore, every lower bound for  $|\Lambda(v, e)|$  implies a version of the *abc*-conjecture: Stewart, Yu and Tijdemann [81, 80] used the Baker-Wüstholz-Matveev bound (1.8) and its  $p$ -adic version by Van der Poorten [84] to prove the inequality

$$\log c < \kappa_\varepsilon \operatorname{rad}(abc)^{\frac{1}{3} + \varepsilon}.$$

To our knowledge, this bound is the strongest bound to date given towards a solution to the *abc*-conjecture and demonstrates the strong connection between the *abc*-conjecture and lower bounds for linear forms in logarithms. Baker [8] proved that a stronger lower bound than (1.8) (together with its  $p$ -adic version) is equivalent to a certain strengthening of the *abc*-conjecture. We will say more about this connection in Section 4.

It is widely conjectured that (1.8) is *not* optimal. In particular, famous conjectures in number theory, such as Waldschmidt's conjectures [85, Conjecture 14.25, p. 547], [86, Conjecture 4.14], or the Lang-Waldschmidt conjecture for Gaussian rationals [59, Introduction to Chapters X and XI] imply that the following hypothesis holds (in fact, they make even stronger predictions!):

► **Number-Theoretic Hypothesis 1.11.** *For any  $n \in \mathbb{Z}_{\geq 0}$ ,  $v_1, v_2, \dots, v_n \in \mathbb{Q}(i)^\times$ , and  $e_1, e_2, \dots, e_n \in \mathbb{Z}$  such that, with  $\operatorname{Log}$  the principal branch of the complex logarithm,*

$$\Lambda(v, e) := e_1 \operatorname{Log} v_1 + e_2 \operatorname{Log} v_2 + \dots + e_n \operatorname{Log} v_n \neq 0,$$

*we have*

$$|\Lambda(v, e)| \geq \exp(-\operatorname{poly}(n, b, B)),$$

*where  $B$  and  $b$  are the maximum bit-lengths of the  $e_j$  and  $v_j$ , respectively.*

Here we prove a novel connection between this number-theoretic conjecture and the separation hypothesis of group orbits, which can be seen as further evidence for the latter:

► **Theorem 1.12.** *Separation Hypothesis 1.7 and Number-Theoretic Hypothesis 1.11 are equivalent. Moreover, Separation Hypothesis 1.6 is equivalent to Number-Theoretic Hypothesis 1.11 when the latter is restricted to  $|v_1| = \dots = |v_n| = 1$ .*

We prove this in Section 4 by using similar ideas as sketched in Section 1.5. There, we also show that Matveev's bound (1.8) implies the lower bound  $\operatorname{sep}_T(d, n, B, b) \geq e^{-B^{O(1)} \cdot b^{O(n)}}$ , which is only exponentially small for constant  $n$ .

## 1.7 Proof sketch of the lattice lifting theorem

We now give a summary of the idea behind the proof of Theorem 1.5. We call a collection  $u_1, u_2, \dots, u_n$  of vectors in  $\mathbb{R}^m$ ,  $n \geq m$ , an *eutactic star of scale  $s$*  if each  $u_j$  can be extended to a vector  $v_j \in \mathbb{R}^n$  such that the  $v_j$  are pairwise orthogonal and of the same norm  $s$ . Note that  $u_i = P(v_i)$  where  $P$  denotes the projection onto the first  $m$  coordinates. Let  $X \in \mathbb{R}^{m \times n}$  denote the matrix whose  $i$ -th column is  $u_i$ . It is easy to see that the collection  $u_1, u_2, \dots, u_n$  forms an eutactic star of scale  $s$  iff  $XX^T = s^2 I_m$ .

Assume that the lattice  $\mathcal{L} \subset \mathbb{R}^m$  is given as the  $\mathbb{Z}$ -span of the columns of the matrix  $G \in \mathbb{Z}^{m \times m}$ . Our aim is to understand when  $\mathcal{L}$  is generated (as a lattice) by an eutactic star. We first note that the collection  $u_1, u_2, \dots, u_n$  is contained in  $\mathcal{L}$  iff there exists an integer matrix  $L \in \mathbb{Z}^{m \times n}$  such that  $X = GL$ . To enforce further that the vectors  $u_i$  generate  $\mathcal{L}$  as a lattice, we require  $L$  to be *right invertible*, meaning there exists  $R \in \mathbb{Z}^{n \times m}$  such that  $LR = I_m$ . In Proposition 6.8 we prove:

$$\begin{aligned} \mathcal{L} \text{ is generated by an eutactic star of scale } s \in \mathbb{Z}_{>0} &\iff \\ \exists \text{ right invertible } L \in \mathbb{Z}^{m \times n} \text{ s.t. } (GL)(GL)^T &= s^2 I_m. \end{aligned}$$

Therefore, writing  $\mathcal{L}$  as in Theorem 1.5 amounts to finding such  $L$  and  $s$ . In Section 6.3 we show that for given  $G \in \mathbb{Z}^{m \times m}$ , such  $L$  and  $s$  can be computed in polynomial time. For this, we choose the integer  $s$  so large that the matrix

$$A := s^2(G^{-1}G^{-T}) - I_m.$$

is positive definite. In Theorem 6.10 we prove that a decomposition of the form  $A = (L')(L')^T$  can be computed in polynomial time (efficient Waring decomposition). Note that this amounts to writing the given positive definite quadratic form  $x^T A x$  as sum of squares of rational linear forms. Such decompositions are known to exist if  $n = m + 3$  [69], and randomized polynomial time algorithms computing them are available [77]. However, to the best of our knowledge, it is open whether such small decompositions can be computed in deterministic polynomial time. Instead, we will use the simple Lemma 6.12, which writes an integer  $D$  as a sum of  $O(\log \log D)$  many squares, but can be easily shown to run in deterministic polynomial time. Finally, one checks that  $L := [I_m \quad L']$  satisfies  $(GL)(GL)^T = s^2 I_m$ , which is what we wanted to get.

## 1.8 The method of Kempf-Ness

A major challenge is to extend the results in this paper beyond torus actions! The Kempf-Ness theorem [55] indicates a strategy to do so. This is a general theorem on rational actions of reductive algebraic groups, that in principle allows to reduce orbit closure intersection problems to compact orbit equality problems. It states that the vectors of minimal norm in the orbit closure of  $v \in V$  form a single  $K$ -orbit that we denote by  $\mathcal{C}_{v^*}$ . Here  $K$  denotes a maximal compact subgroup of a reductive algebraic group  $G$  that is assumed to act isometrically on  $V = \mathbb{C}^n$ . Moreover, the Kempf-Ness theorem states that for  $v, w \in \mathbb{C}^n$ ,

$$\overline{\mathcal{O}_v} \cap \overline{\mathcal{O}_w} \neq \emptyset \iff \mathcal{C}_{v^*} = \mathcal{C}_{w^*}.$$

In an ideal world, this reduces the problem of testing  $\overline{\mathcal{O}_v} \cap \overline{\mathcal{O}_w} \neq \emptyset$  to the problem of testing  $\mathcal{C}_{v^*} = \mathcal{C}_{w^*}$ , provided that we can compute  $v^*$  on input  $v$ . Unfortunately, even when the coordinates of  $v$  are rational,  $v^*$  may not have rational entries. Instead one may work with a numerical approximation of  $v^*$ . This was successfully carried out in [3] for the left-right

action on tuples of complex matrices. As a result, polynomial time algorithms for the orbit closure problems for the left-right action were obtained along this way.

For torus actions, approximating  $v^*$  becomes a convex optimization problem. More precisely, we consider the *Kempf-Ness function* corresponding to  $v$ ,

$$f(x) := \log \|e^{x/2} \cdot v\|^2 = \log \left( \sum_{i=1}^n q_i e^{\omega_i^T x} \right),$$

where  $\omega_1, \omega_2, \dots, \omega_n \in \mathbb{Z}^d$  denote the columns of  $M$  and  $q_i := |v_i|^2$ . We have  $\min_x f(x) = 2 \log \|v^*\|$  by the definition of  $v^*$ . This function finds applications in a surprising number of different areas: In machine learning [21, 22, 46, 64],  $f(x)$  is known as the *log-sum-exp* function and it is used as a smooth approximation to the piecewise affine function  $L(x) := \max_i (\langle x, \omega_i \rangle + \log q_i)$ : we have  $L(x) \leq f(x) \leq L(x) + \log n$ . The optimization of the Kempf-Ness function also has uses in the area of statistical physics: The convex program  $\inf_x f(x)$  is the Lagrange dual of an entropy maximization problem with mean constraints. More precisely, one has

$$\inf_x f(x) = \sup \left\{ -D_{KL}(p||q) \mid \sum_{i=1}^n p_i \omega_i = 0, \sum_{i=1}^n p_i = 1, \forall i, p_i \geq 0 \right\}$$

where  $D_{KL}(p||q) := \sum_{i=1}^n p_i \log(p_i/q_i)$  denotes the *Kullback-Leibler divergence* between the probability distribution  $p$  and (possibly non-normalized) distribution  $q$ . For more on this connection, we refer to [79, 82, 60, 20].

In Section 7, we consider the problem of approximating the optimal solution of  $f(x)$ . We conjecture that a point  $x \in \mathbb{R}^d$  that is  $\varepsilon$ -close to  $\arg \min f(x)$  in the Euclidean distance can be computed in polynomial time, see Conjecture 7.3. We show that if the conjecture is true then the logarithmic distance  $\delta_{\log}(\mathcal{C}_{v^*}, \mathcal{C}_{w^*})$  between two Kempf-Ness orbits can be efficiently approximated. This leads to new polynomial time algorithms for the orbit equality problems, provided the separation hypothesis holds. We interpret our finding as some positive evidence towards the feasibility of the Kempf-Ness approach in more general settings, but also interpret it as a warning about the difficulties to be encountered.

## 1.9 Conclusion and outlook

Summarizing, we defined the problem of approximating orbit distances for torus actions and showed that these problems are NP-hard for an almost-polynomial factor, but can be solved in polynomial time within an at most single exponentially large factor under the assumption that a variant of the famous *abc*-conjecture is true. Our results point to an exciting connection between orbit problems and deep theorems and conjectures in the areas of number theory and algorithmic lattice theory.

Let us point out that our approximation algorithm for the robust orbit problem, jointly with our reduction result from CVP to this problem, yield a new polynomial time approximation algorithm for CVP, which is not based on the LLL algorithm. Unfortunately, the resulting approximation factor  $\gamma$  not only depends on the rank  $m$  of the lattice, but also on the bit-length of the generators of the lattice, which cannot compete with the well-known  $2^{O(m)}$ -approximation algorithms for CVP (see, for example, [67, Lemma 2.12]). It is an interesting question whether competitive algorithms for CVP can be developed along this way and whether there are other uses of our lattice lifting result (Theorem 1.5). A further interesting problem is if one can remove the separation hypothesis by allowing for significantly larger approximation factors  $\gamma$ .

## 1.10 Organization of the paper

The preliminary Section 2 collects known facts from different areas that we need to establish our main results. Section 3 is devoted to an analysis of the structure of the orbits in the logarithmic space  $\mathbb{C}^n/2\pi i\mathbb{Z}^n$ , and the logarithmic distance metric  $\delta_{\log}$  defined in Equation (1.3). We show that  $\delta_{\log}(\mathcal{O}_v, \mathcal{O}_w)$  is approximated by a linear form in  $\text{Log } v, \text{Log } w$ . The goal of Section 4 is to explain the connection between the *abc*-conjecture and Number-Theoretic Hypothesis 1.11 in more detail and to formally prove Theorem 1.12. Section 5 describes our algorithms to solve Problem 1.3 and Problem 1.2 in polynomial time, proving Theorem 1.8. The goal of Section 6 is to prove the hardness results by reducing the closest vector problem to the orbit distance problems (Problem 1.2 and Problem 1.3). There, we also provide the proof of Theorem 1.5 on efficient lattice lifting and we prove Theorem 1.9 and Theorem 1.10. Finally, Section 7 is devoted to the analysis of the Kempf-Ness approach for general torus actions. We conjecture a result on the complexity of approximating the optimal solution of the Kempf-Ness function and based on that, we devise a numerical algorithm for deciding the equality of two orbits, which runs in polynomial time if Separation Hypothesis 1.7 is true.

## 2 Preliminaries

The goal of this section is to collect known facts from different areas: invariants for torus actions, singular values of matrices, the complexity of numerically computing elementary functions. We end with general observations on quotient metrics.

### 2.1 Notation

Throughout the paper,  $T = (\mathbb{C}^\times)^d$  denotes the algebraic torus of rank  $d$  and  $V = \mathbb{C}^n$ . We always denote by  $M \in \mathbb{Z}^{d \times n}$  a weight matrix that determines the action of  $T$  on  $V$ . The Euclidean norm of  $v \in \mathbb{C}^n$  is denoted  $\|v\|$ . We write  $\text{dist}(v, w) := \|v - w\|$  for the Euclidean distance of  $v, w \in \mathbb{C}^n$  and we extend this notation to nonempty subsets  $A, B \subseteq \mathbb{C}^n$  by setting  $\text{dist}(A, B) := \inf_{a \in A, b \in B} \|a - b\|$ .

We denote by  $\Re(z)$  and  $\Im(z)$  the real and imaginary part of a complex number  $z \in \mathbb{C}$ , respectively. Moreover,  $\log(z) \in \mathbb{C}/2\pi i\mathbb{Z}$  denotes complex logarithm. The componentwise defined logarithm and exponentiation functions are denoted by  $\text{Log}: (\mathbb{C}^\times)^n \rightarrow \mathbb{C}^n/2\pi i\mathbb{Z}^n$  and  $\text{Exp}: \mathbb{C}^n/2\pi i\mathbb{Z}^n \rightarrow (\mathbb{C}^\times)^n$ .

The letters  $v, w$  denote vectors in  $V$ , while the Greek letters  $\eta, \zeta$  refer to vectors that are exponentiated, as in  $v = \text{Exp}(\eta)$ . We always denote by  $b$  a bound for the bit-lengths of the components of the input vectors  $v, w \in \mathbb{Q}(i)^n$  (or  $\eta, \zeta$ ), and  $B$  always denotes a bound for the bit-length of the entries of the weight matrix  $M \in \mathbb{Z}^{d \times n}$ .

### 2.2 Invariant theory of torus actions

In this section, we present a brief summary of the setting and the results of [16]. We will discuss the proof strategy sketched in Section 1.5 in more detail. As always,  $T = (\mathbb{C}^\times)^d$  is the  $d$ -dimensional torus and  $V = \mathbb{C}^n$ .

Any rational action of  $T$  on  $V$ , up to some base change, can be simultaneously diagonalized and brought to the form Equation (1.1). The matrix  $M = [M_{ij}] \in \mathbb{Z}^{d \times n}$  occurring in Equation (1.1) is called the *weight matrix* and its columns  $\omega_1, \omega_2, \dots, \omega_n \in \mathbb{Z}^d$  are called the *weights* of the action. There is an induced action of  $T$  on the polynomial ring  $\mathbb{C}[x_1, x_2, \dots, x_n]$ ,

given by

$$(t \cdot f)(v) := f(t^{-1} \cdot v), \quad f \in \mathbb{C}[x_1, x_2, \dots, x_n], \quad t \in T, \quad v \in V.$$

A *polynomial invariant* is a fixed point of this action, i.e.,  $t \cdot f = f$  for all  $t \in T$ . Note that a polynomial is an invariant iff it is constant along orbits. We denote the ring of polynomial invariants by  $\mathbb{C}[V]^T$ . For a monomial  $x^\alpha := x^{\alpha_1} x^{\alpha_2} \dots x^{\alpha_n} \in \mathbb{C}[x_1, x_2, \dots, x_n]$ , we have

$$x^\alpha(t \cdot v) = t^{M\alpha} x^\alpha(v), \quad (2.1)$$

Equation (2.1) implies that the line  $\mathbb{C}x^\alpha$  is preserved by the action of  $T$ . Moreover,  $x^\alpha$  is an invariant iff  $M\alpha = 0$ . Hence a polynomial  $f$  is an invariant if and only if each monomial appearing in  $f$  is invariant. In particular, the space of invariant polynomials is linearly spanned by the invariant monomials.

The action of  $T$  leaves  $X := (\mathbb{C}^\times)^n$  invariant. We have an induced action of  $T$  on the algebra

$$\mathbb{C}[X] = \mathbb{C}[x_1, x_1^{-1}, x_2, x_2^{-1}, \dots, x_n, x_n^{-1}],$$

of Laurent polynomials, the ring of regular functions  $\mathbb{C}[X]$  of  $X$ , which is easier to study. The above observations extend: Equation (2.1) also holds for *Laurent monomials*  $x^\alpha$  with exponent vector  $\alpha \in \mathbb{Z}^n$ , which is thus invariant iff  $M\alpha = 0$ . The space of invariant Laurent polynomials is linearly spanned by the invariant Laurent monomials. We call

$$\mathcal{K} := \{\alpha \in \mathbb{Z}^n \mid M\alpha = 0\}, \quad (2.2)$$

the *lattice of rational invariants* defined by  $M$ . Note that the rank of  $\mathcal{K}$  is given by  $k = n - \text{rk } M$ .

Thus  $\alpha_1, \alpha_2, \dots, \alpha_k$  is a lattice basis of  $\mathcal{K}$ , then the Laurent monomials  $x^{\alpha_1}, x^{\alpha_2}, \dots, x^{\alpha_k}$  generate  $\mathbb{C}[X]^T$  as an algebra. Then, for  $v, w \in (\mathbb{C}^\times)^n$ , we have (see [16, Proposition 4.1])

$$\mathcal{O}_v = \mathcal{O}_w \iff \forall i \in [k] \quad x^{\alpha_i}(v) = x^{\alpha_i}(w). \quad (2.3)$$

There is an analogous result for the orbits of the compact torus  $K = (S^1)^d$  (see [16, Proposition 8.1]): For two vectors  $v, w \in (\mathbb{C}^\times)^n$ , we have

$$\begin{aligned} \mathcal{C}_v = \mathcal{C}_w &\iff \forall \alpha \in \mathcal{K}, \quad x^\alpha(v) = x^\alpha(w) \text{ and } \forall i \in [n], |v_i| = |w_i| \\ &\iff \mathcal{O}_v = \mathcal{O}_w \text{ and } \forall i \in [n], |v_i| = |w_i|. \end{aligned} \quad (2.4)$$

The next theorem (see [16, Corollary 4.4 and Proposition 5.5]) shows that one can decide  $\mathcal{O}_v = \mathcal{O}_w$  in polynomial time.

- **Theorem 2.1. (a)** *We can compute in  $\text{poly}(d, n, B)$ -time a basis for the lattice  $\mathcal{K}$  of  $T$ -invariant Laurent monomials. In particular, basis elements of  $\mathcal{K}$  have bit-lengths bounded by  $\text{poly}(d, n, B)$ .*
- (b) *Suppose  $v, w \in (\mathbb{Q}(i)^\times)^n$  and  $\alpha \in \mathbb{Z}^n$ . Then in  $\text{poly}(n, b, \langle \alpha \rangle)$ -time one can decide whether  $x^\alpha(v) = x^\alpha(w)$ .*

Let  $H \in \mathbb{Z}^{k \times n}$  be a matrix whose rows  $\alpha_1^T, \alpha_2^T, \dots, \alpha_k^T$  form a basis of the lattice  $\mathcal{K}$  defined by  $M$ , see Equation (2.2). We call  $H$  a *matrix of rational invariants*.

- **Proposition 2.2.** *We have  $\ker H = \text{im } M^T$  and  $\text{rk } H = k$ . Moreover,  $H(\mathbb{Z}^n) = \mathbb{Z}^k$ , i.e., for every  $\beta \in \mathbb{Z}^k$ , there exists  $\alpha \in \mathbb{Z}^n$  such that  $H\alpha = \beta$ . If the rows of  $H$  are produced from  $M$  by a polynomial time algorithm (as in Theorem 2.1), then the bit-length of  $H$  is at most  $\text{poly}(d, n, B)$ .*

**Proof.** The first claim is obvious from the construction of  $\mathcal{K}$  and  $H$ . The second claim is obvious. For the third claim, we are going to use the integral analogue of Farkas' lemma (see [78, Corollary 4.1a]), which states that for  $H \in \mathbb{Z}^{k \times n}, \beta \in \mathbb{Z}^k$ ,

$$H\alpha = \beta \text{ has a solution } \alpha \in \mathbb{Z}^n \iff \forall \gamma \in \mathbb{Q}^k \text{ with } H^T \gamma \in \mathbb{Z}^n, \text{ it holds that } \beta^T \gamma \in \mathbb{Z}.$$

To reach a contradiction, suppose  $H(\mathbb{Z}^n) \neq \mathbb{Z}^k$ , i.e., there exists a vector  $\beta \in \mathbb{Z}^k \setminus H(\mathbb{Z}^n)$ . Then there exists a rational vector  $\gamma \in \mathbb{Q}^k$  such that  $H^T \gamma \in \mathbb{Z}^n$  but  $\beta^T \gamma \notin \mathbb{Z}$ .

We note that if  $H^T \gamma \in \mathbb{Z}^n$  for some  $\gamma \in \mathbb{Q}^k$ , then  $H^T \gamma \in \mathcal{K}$  since  $MH^T = 0$ . Moreover, since the rows of  $H$  generate  $\mathcal{K}$ , we further have  $H^T \gamma = H^T \tilde{\gamma}$  for some integral vector  $\tilde{\gamma} \in \mathbb{Z}^k$ . However,  $\text{rk } H = k$  so we must have  $\gamma = \tilde{\gamma}$ , so  $\gamma$  is integral. This contradicts  $\beta^T \gamma \notin \mathbb{Z}$ . ◀

► **Remark 2.3.** The property that  $H(\mathbb{Z}^n) = \mathbb{Z}^k$  will be used frequently throughout the paper. We note that this is equivalent to the diagonal entries of the Smith normal form of  $H$  being all one. Moreover, the proof of Proposition 2.2 shows that  $H(\mathbb{Z}^n) = \mathbb{Z}^k$  holds whenever the rows of  $H$  form a basis of a *saturated lattice* in  $\mathbb{Z}^n$  with rank  $k$ . We recall that this means for  $s \in \mathbb{Z}_{>0}$  and  $\alpha \in \mathbb{Z}^n$ ,  $s\alpha \in \mathcal{K}$  implies  $\alpha \in \mathcal{K}$ .

► **Remark 2.4.** We could ignore the dependence of the complexity parameters on  $d$  since  $d \leq n$  can be assumed without loss of generality. For seeing this, assume  $d > n$  and let  $A \in \mathbb{Z}^{d \times n}$  denote the (row reduced) Hermite normal form of the weight matrix  $M \in \mathbb{Z}^{d \times n}$ . Then the orbits with respect to the actions defined by  $M$  and by  $A$  are the same. However, since the last  $d - n$  rows of  $A$  are zero, the last  $d - n$  coordinates of the torus act trivially on  $V$  and can be ignored.

## 2.3 Singular values

Suppose  $k \leq n$ . For  $H \in \mathbb{C}^{k \times n}$ , there exist unitary matrices  $X \in \text{U}(k), Y \in \text{U}(n)$  such that  $XHY$  is a diagonal matrix with non-negative real diagonal entries  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_k \geq 0$ . The values  $\sigma_i$  are called the *singular values* of the matrix  $H$ . We denote by  $\sigma_{\max}(H) := \sigma_1$  and  $\sigma_{\min}(H) := \sigma_k$ . We have  $\sigma_{\min}(H) \neq 0$  if and only if  $\text{rk } H = k$ . Recall that  $\|x\|$  stands for the Euclidean norm of  $x \in \mathbb{C}^n$  and  $\text{dist}$  denotes the corresponding distance for subsets of  $\mathbb{C}^n$ .

The maximum and the minimum singular values  $\sigma_{\max}(H), \sigma_{\min}(H)$  are characterized by the following properties:

$$\sigma_{\max}(H) = \max_{\|x\|=1} \|Hx\| \quad \text{and} \quad \forall x \quad \|Hx\| \geq \sigma_{\min}(H) \cdot \text{dist}(x, \ker H). \quad (2.5)$$

► **Lemma 2.5.** Suppose  $k \leq n$  and  $H \in \mathbb{C}^{k \times n}$ . For nonempty subsets  $A, B \subset \mathbb{C}^n$  we have

$$\sigma_{\min}(H) \text{dist}(A + \ker H, B + \ker H) \leq \text{dist}(HA, HB) \leq \sigma_{\max}(H) \text{dist}(A + \ker H, B + \ker H).$$

**Proof.** We first note that  $\text{dist}(A + \ker H, B + \ker H)$  is the infimum of  $\|a - b + u\|$  over  $a \in A, b \in B, u \in \ker H$ , and  $\text{dist}(HA, HB)$  is the infimum of  $\|H(a - b)\|$  over  $a \in A, b \in B$ .

For any  $a \in A, b \in B$  and  $u \in \ker H$  we have, by the equation in (2.5),

$$\text{dist}(HA, HB) \leq \|H(a - b)\| = \|H(a - b + u)\| \leq \sigma_{\max}(H) \|a - b + u\|.$$

Taking the infimum over  $a, b, u$  we have  $\text{dist}(HA, HB) \leq \sigma_{\max}(H) \text{dist}(A + \ker H, B + \ker H)$ .

For the other inequality we use the inequality in (2.5): For any  $a \in A, b \in B$

$$\|H(a - b)\| \geq \sigma_{\min}(H) \text{dist}(a - b, \ker H) \geq \sigma_{\min}(H) \text{dist}(A + \ker H, B + \ker H).$$

Taking the infimum over  $a, b$  we get  $\text{dist}(HA, HB) \geq \sigma_{\min}(H) \text{dist}(A + \ker H, B + \ker H)$ . ◀

The distance  $\text{dist}(A + \ker H, B + \ker H)$  depends only on  $\ker H$  but not on  $H$  itself. Consequently, Lemma 2.5 gives different approximations for different matrices with the same kernels. The optimal choice is the orthogonal projection onto  $\ker(H)^\perp$ . The singular values of the orthogonal projection are all 1, so in this case the inequality from Lemma 2.5 becomes an equality and we get the following.

► **Corollary 2.6.** *For nonempty subsets  $A, B \subset \mathbb{C}^n$  and an orthogonal projection  $P : \mathbb{C}^n \rightarrow \mathbb{C}^n$ , we have*

$$\text{dist}(A + \ker P, B) = \text{dist}(A + \ker P, B + \ker P) = \text{dist}(P(A), P(B)).$$

We will use Lemma 2.5 mostly in the case where  $H$  is the matrix of rational invariants defined in the previous section. In this case  $H$  is integral and the singular values can be bounded in terms of the bit-length of  $H$ .

► **Lemma 2.7.** *Suppose  $k \leq n$  and  $H \in \mathbb{Z}^{k \times n}$  is an integer matrix of rank  $\text{rk } H = k$ . Then*

$$\sigma_{\max}(H) \leq n \|H\|_{\max}, \quad \sigma_{\min}(H) \geq n^{-(n-1)} \|H\|_{\max}^{-(n-1)},$$

where  $\|H\|_{\max} := \max_{i,j} |H_{ij}|$  is the max norm of  $H$ . Consequently, if  $B := \langle H \rangle$  is the bit-length of  $H$ , then

$$\kappa(H) := \sigma_{\max}(H) / \sigma_{\min}(H) \leq n^n 2^{Bn}.$$

Given an  $H$  as above, one can compute in polynomial time a number  $D \in \mathbb{Q}$  such that  $\sigma_{\min}(H) \leq D \leq 2\sigma_{\min}(H)$ .

**Proof.** The upper bound on  $\sigma_{\max}(H)$  follows from  $\|Hx\| \leq \sqrt{nk} \|H\|_{\max} \|x\|$ . For the lower bound, note that the product of the singular values of  $H$  equals  $\prod_{i=1}^k \sigma_i(H) = \sqrt{\det(HH^T)}$ . Since  $H$  has rank  $k$ , the matrix  $HH^T$  is invertible and positive definite. Hence,  $\det(HH^T) \geq 1$  and we deduce with part one that

$$\sigma_{\min}(H) \geq \sigma_{\max}(H)^{1-k} \sqrt{\det(HH^T)} \geq n^{-(k-1)} \|H\|_{\max}^{-(k-1)} \geq n^{-(n-1)} \|H\|_{\max}^{-(n-1)},$$

which shows the second inequality.

We refer to [76] for the algorithmic claim on computing  $\sigma_{\min}(H)$ . ◀

## 2.4 Complexity of elementary functions

Two important functions used in this paper are the complex logarithm and exponentiation,  $\log$  and  $\exp$ . Both functions are transcendental and rarely assume rational values on rational inputs. Fortunately, they can be efficiently approximated by rational functions with the arithmetic mean-geometric mean iteration of Gauss, Lagrange and Legendre. We refer to the paper [13] and the book [12] for a detailed study of the AM-GM iteration and for the following results.

► **Lemma 2.8.** *Suppose  $\log$  denotes the standard branch of the complex logarithm whose imaginary part satisfies  $\Im(\log(z)) \in (-\pi, \pi]$ , and  $\text{Log}$  is the componentwise logarithm. Assume that  $v \in (\mathbb{Q}(i)^\times)^n$  is a vector with Gaussian rational entries. Given  $\varepsilon \in \mathbb{Q}_{>0}$ , in  $\text{poly}(n, \langle v \rangle, \log \varepsilon^{-1})$ -time we can compute an approximation to  $\text{Log } v$  with absolute error  $\varepsilon$ . That is, we can compute a vector  $\eta \in \mathbb{Q}^n + 2\pi i \mathbb{Q}^n$  such that*

$$\|\text{Log } v - \eta\| < \varepsilon.$$



► **Lemma 2.9.** *Suppose  $\exp$  denotes the complex exponentiation,  $\exp(\rho + i\theta) = e^\rho(\cos(\theta) + i\sin(\theta))$  and  $\text{Exp}$  is the componentwise exp. Assume that  $\eta \in \mathbb{Q}^n + 2\pi i\mathbb{Q}^n$ . Given  $\varepsilon \in \mathbb{Q}_{>0}$ , in  $\text{poly}(n, \langle \eta \rangle, \log \varepsilon^{-1})$ -time we can compute an approximation to  $\text{Exp}(\eta)$  with relative error  $\varepsilon$ . That is, we can compute a vector  $v \in (\mathbb{Q}(i)^\times)^n$  such that*

$$\frac{\|\text{Exp}(\eta) - v\|}{\|\text{Exp}(\eta)\|} < \varepsilon.$$

## 2.5 Quotient topology and quotient metric

Throughout the paper, we consider various group actions on various metric spaces. The orbit space, i.e., the set of orbits of the action, has a natural quotient topology and it is an important question for us to decide when it is possible to carry over the metric space structure to the orbit space. This is possible in all cases we consider in this section.

► **Definition 2.10.** *Suppose  $X$  is a topological space and  $\sim$  is an equivalence relation on  $X$ . We denote by  $X/\sim := \{[x] \mid x \in X\}$  the set of equivalence classes of  $X$ , where  $[x]$  is the equivalence class of  $x$ . Let  $\pi : X \rightarrow X/\sim$  be the canonical projection map. The quotient topology on  $X/\sim$  is defined by calling a subset  $U$  of  $X/\sim$  open iff  $\pi^{-1}(U)$  is open in  $X$ .*

Assume that  $(X, \delta)$  is a metric space with distance function  $\delta$ . We define the induced distance function on  $X/\sim$  by

$$\delta([x], [y]) := \inf\{\delta(x', y') \mid x \sim x', y \sim y'\}.$$

Clearly, this is a *pseudo-metric* on  $X/\sim$ , which means that  $\delta(x, x) = 0$ ,  $\delta(x, y) = \delta(y, x)$  and  $\delta(x, z) \leq \delta(x, y) + \delta(y, z)$  for all  $x, y, z \in X$ . Moreover, one can show that  $\delta$  is *compatible* with the quotient topology on  $X/\sim$ , which means that the balls  $\{[x] \mid \delta([x], [y]) < r\}$  form a basis of the quotient topology on  $X/\sim$ , see [45, Theorem 4]. A pseudo-metric is a metric iff  $\delta(x, y) = 0$  implies  $x = y$ .

We note that the induced distance is not always a metric. Indeed, the quotient topology on  $X/\sim$  does not need to be Hausdorff. (The so-called line with two origins is an example for this.)

However, in the following two cases, the pseudo-metric  $\delta([x], [y])$  is indeed a metric.

► **Proposition 2.11.** *Suppose  $(X, \delta)$  is a metric space and  $K$  is a compact topological group acting continuously on  $X$ . For  $x \in X$ , we denote by  $\mathcal{C}_x$  the orbit of  $x$ . Then  $\delta(\mathcal{C}_x, \mathcal{C}_y)$  defines a metric on the space of orbits  $X/K := \{\mathcal{C}_x \mid x \in X\}$ , which is compatible with the quotient topology.*

**Proof.** The orbit  $\mathcal{C}_x$  is compact since it is the image of  $K$  under the continuous map  $k \mapsto k \cdot x$ . By the above observation,  $\text{dist}(\mathcal{C}_x, \mathcal{C}_y)$  defines a compatible pseudo-metric on  $X/K$ . Different orbits  $\mathcal{C}_x, \mathcal{C}_y$  are disjoint. Since they are compact, they have a positive distance:  $\text{dist}(\mathcal{C}_x, \mathcal{C}_y) > 0$ . This shows that the induced distance is a metric. ◀

For the second example, let  $V$  be a finite dimensional real vector space and  $\mathcal{L} \subseteq V$  be a lattice. This means that  $\mathcal{L}$  is the set of  $\mathbb{Z}$ -linear combinations of a collection of linearly independent vectors in  $V$ . Alternatively, a lattice  $\mathcal{L}$  can be defined as an additive subgroup of  $V$  that is *discrete*, i.e., there exists a positive constant  $\varepsilon > 0$  such that for any distinct lattice points  $x, y \in \mathcal{L}$  we have  $\text{dist}(x, y) \geq \varepsilon$ . Thus, any convergent sequence of lattice points must be eventually constant.

► **Proposition 2.12.** *Let  $V$  be a finite dimensional real vector space,  $\mathcal{L}$  a lattice in  $V$ , and  $\delta$  a translation invariant metric on  $V$  that is compatible with the standard topology on  $V$ . Then  $\delta(x + \mathcal{L}, y + \mathcal{L})$  defines a metric on the quotient space  $V/\mathcal{L}$  which is compatible with the quotient topology.*

**Proof.** We argue as in the proof of Proposition 2.11. Suppose that  $\delta(x + \mathcal{L}, y + \mathcal{L}) = 0$ . Then there exist sequences  $u_i$  and  $v_i$  in  $\mathcal{L}$  such that  $\delta(x + u_i, y + v_i) = \delta(x - y, v_i - u_i)$  converges to zero. Since the sequence  $v_i - u_i$  of lattice points converges, it must be eventually constant. Hence  $x - y \in \mathcal{L}$ , which completes the proof. ◀

Sometimes in the paper the lattice  $\mathcal{L}$  will be given as an orthogonal projection of another lattice. Note that the orthogonal projection of a lattice is not always a lattice. (For instance, the orthogonal projection of  $\mathbb{Z}^2$  onto  $\mathbb{R}(1, y)$  with irrational  $y$  is not discrete.)

► **Lemma 2.13.** *Suppose  $U \subset V$  is a rational subspace of  $V = \mathbb{R}^n$ , spanned by linearly independent rational vectors  $u_1, \dots, u_k \in \mathbb{Q}^n$  and  $\mathcal{L} \subset V$  is a lattice, generated by integral vectors. Then, the orthogonal projection of  $\mathcal{L}$  to  $U$  is a lattice.*

**Proof.** Using Gram-Schmidt orthogonalization, we may assume that the  $u_i$  are pairwise orthogonal. Then the orthogonal projection  $P : \mathbb{R}^n \rightarrow U$  is given by  $P(v) = \sum_{i=1}^k \frac{\langle u_i, v \rangle}{\langle u_i, u_i \rangle} u_i$ . This shows that  $P(\mathbb{Q}^n) \subset \mathbb{Q}^n$ . Therefore, if  $v_1, v_2, \dots, v_l \in \mathbb{Z}^n$  generate  $\mathcal{L}$ , there is a positive integer  $N$  such that  $N P(v_i) \in \mathbb{Z}^n$  for all  $i$ . Hence  $N P(\mathcal{L}) \subset \mathbb{Z}^n$ , which implies that  $P(\mathcal{L})$  is discrete. ◀

### 3 Logarithmic image of orbits

The goal of this section is to discuss the structure of the logarithmic image of orbits in the quotient space  $\mathbb{C}^n/2\pi i\mathbb{Z}$ , to study the metric  $\delta_{\log}$  defined in (1.3), and to prove that  $\delta_{\log}(\mathcal{O}_v, \mathcal{O}_w)$  is approximated by a linear form in  $\text{Log } v$  and  $\text{Log } w$ .

We assume the setting of Section 2.2:  $T = (\mathbb{C}^\times)^d$  acts on  $V = \mathbb{C}^n$  via the weight matrix  $M \in \mathbb{Z}^{d \times n}$ . The columns of  $M$  are called the weights of the action and the *weight polytope*  $P \subset \mathbb{R}^d$  is defined as the convex hull of these weights. We consider orbits of vectors  $v \in (\mathbb{C}^\times)^n$ , thus we assume that all components of  $v$  are nonzero. We note that  $\mathcal{O}_v$  is closed in  $(\mathbb{C}^\times)^n$ , see [16, Prop. 5.1]. However,  $\mathcal{O}_v$  may not be closed. It is well known that  $\mathcal{O}_v$  is closed iff 0 lies in the interior of  $P$ ; see [16, Section 3]. This will only become relevant later in Section 7.2.

#### 3.1 Structure of logarithms of orbits

The key idea is to use the group isomorphism  $\text{Log} : (\mathbb{C}^\times)^n \rightarrow \mathbb{C}^n/2\pi i\mathbb{Z}^n$  provided by the componentwise complex logarithm, compare Equation (1.2). The inverse is given by the (componentwise) exponential function  $\text{Exp}$ .

The action of the group  $T$  on  $(\mathbb{C}^\times)^n$  induces via  $\text{Log}$  an action of  $T$  on  $\mathbb{C}^n/2\pi i\mathbb{Z}^n$ , which we shall denote by  $*$ . This action is simpler to understand, since it works by *translations*. More specifically, for  $x \in \mathbb{C}^n$  and  $v \in (\mathbb{C}^\times)^n$ , we have

$$\text{Log}(e^x \cdot v) = \text{Log } v + M^T x, \quad (3.1)$$

see Figure 3. Note that, by the periodicity of  $\exp$ , the right-hand side in  $\mathbb{C}^n/2\pi i\mathbb{Z}^n$  indeed only depends on  $x \bmod 2\pi i\mathbb{Z}^n$ . (We could also consider the corresponding action of the Lie algebra  $\text{Lie}(T) = \mathbb{C}^d$ , which has the same orbits.) This leads us to the following definition.

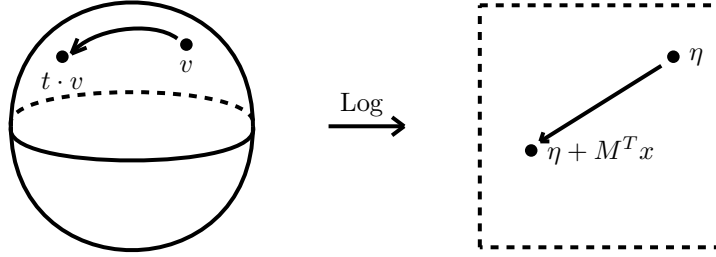
► **Definition 3.1.** We denote by  $*$  the induced action of  $T$  on  $\mathbb{C}^n/2\pi i\mathbb{Z}^n$  via translations defined by

$$e^x * \eta := \eta + M^T x, \quad \text{for } x \in \mathbb{C}^d, \eta \in \mathbb{C}^n/2\pi i\mathbb{Z}^n.$$

By construction, the orbits  $\mathcal{O}_v$  and  $\mathcal{C}_v$  of  $v \in (\mathbb{C}^\times)^n$  are mapped to the corresponding orbits of  $\eta = \text{Log } v$ . If we denote

$$T * \eta := \{t * \eta \mid t \in T\} \quad \text{and} \quad K * \eta := \{k * \eta \mid k \in K\},$$

then  $\text{Log}(\mathcal{O}_v) = T * \text{Log}(v)$  and  $\text{Log}(\mathcal{C}_v) = K * \text{Log}(v)$ .



■ **Figure 3** The logarithm *flattens* the group action. The vector  $\eta = \text{Log } v$  is translated to  $\eta + M^T x$  via the group element  $t = e^x \in T$ .

We observe that the  $*$ -orbits of  $T$  (and  $K$ ) are the *images of affine subspaces* under the canonical surjection  $\mathbb{C}^n \rightarrow \mathbb{C}^n/2\pi i\mathbb{Z}^n$ . More concretely, we denote by  $U_{\mathbb{C}} := \text{im } M^T = \{M^T x \mid x \in \mathbb{C}^d\}$  the row space of  $M$ . Then it immediately follows from the definition of the  $*$ -action that

$$T * \eta = \left( \eta + U_{\mathbb{C}} + 2\pi i\mathbb{Z}^n \right) / 2\pi i\mathbb{Z}^n. \quad (3.2)$$

We equip now  $\mathbb{C}^n/2\pi i\mathbb{Z}^n$  with the quotient metric  $\Delta$  induced by the Euclidean metric on  $\mathbb{C}^n$ , denoted by  $\text{dist}$  (by Proposition 2.12 this is indeed a metric). Explicitly,

$$\begin{aligned} \Delta(\eta, \zeta) &:= \text{dist}(\eta - \zeta, 2\pi i\mathbb{Z}^n) \\ &= \sqrt{\|\rho - \tau\|^2 + 4\pi^2 \text{dist}^2(\theta - \phi, \mathbb{Z}^n)}, \end{aligned} \quad (3.3)$$

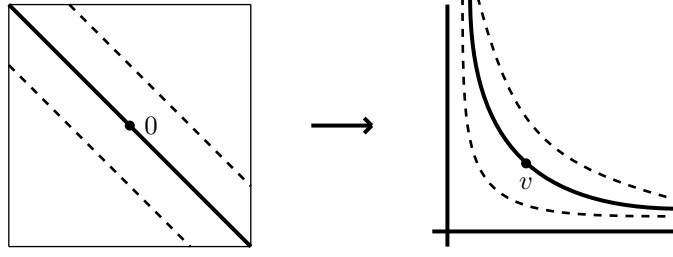
where  $\eta = \rho + 2\pi i\theta$  and  $\zeta = \tau + 2\pi i\phi$  with  $\rho, \theta, \tau, \phi \in \mathbb{R}^n$  (note that the imaginary parts  $\theta, \phi$  are only determined modulo  $\mathbb{Z}^n$ ).

► **Proposition 3.2.** The metric  $\Delta$  on  $X := \mathbb{C}^n/2\pi i\mathbb{Z}^n$  induces a metric

$$\Delta(T * \eta, T * \zeta) := \inf_{t \in T} \Delta(t * \eta, \zeta)$$

on the space  $X/T$  of  $T$ -orbits with respect to the  $*$ -action, which is compatible with the quotient topology. An analogous result holds for  $K$ -orbits.

**Proof.** Denote by  $P : \mathbb{C}^n \rightarrow U_{\mathbb{C}}^\perp$  the orthogonal projection onto the orthogonal complement of  $U_{\mathbb{C}}$ , thus  $\ker P = U_{\mathbb{C}}$ . Lemma 2.13 shows that  $2\pi i P(\mathbb{Z}^n)$  is a lattice and Proposition 2.12 implies that  $Y := U_{\mathbb{C}}^\perp / P(2\pi i\mathbb{Z}^n)$  is a metric space with respect to the quotient metric of the Euclidean metric.



■ **Figure 4** Suppose  $\mathbb{C}^\times$  acts on  $\mathbb{C}^2$  via  $t \cdot (x, y) = (tx, t^{-1}y)$  as in Figure 1. The logarithmic orbit  $T * 0$  is mapped via  $\text{Exp}$  to the orbit  $\mathcal{O}_v$  of  $v := (1, 1)$ . The image shows the  $\varepsilon$ -neighbourhood of  $T * 0$  and its image under  $\text{Exp}$ .

The projection  $P$  induces a surjective group morphism  $P': X \rightarrow Y$ . From Equation (3.2) we see that  $P'$  is constant on  $T$ -orbits. More specifically,  $P'$  induces a continuous bijection  $P'': X/T \rightarrow Y$ . By definition,  $P''$  preserves the distance  $\Delta$ :

$$\Delta(T * \eta, T * \zeta) = \Delta(\eta, \zeta).$$

Since  $\Delta$  is a metric on  $Y$ , this implies that the pseudo-metric  $\Delta$  on  $X/T$  in fact is a metric. Therefore,  $\Delta(T * \eta, T * \zeta) = 0$  implies  $T * \eta = T * \zeta$ . The claim about  $K$ -orbits follows analogously.  $\blacktriangleleft$

► **Remark 3.3.** The fact that  $\Delta$  is metric on  $X/T$  is our essential gain. Note that the orbit space  $(\mathbb{C}^\times)/T$  equipped with the induced Euclidean distance  $\text{dist}(\mathcal{O}_v, \mathcal{O}_w)$  between  $T$ -orbits is not a metric: it can be zero even though  $\mathcal{O}_v \neq \mathcal{O}_w$ , as illustrated in the case of the hyperbolas in Figure 1.

Equation (3.3) implies the important equations

$$\Delta(T * \eta, T * \zeta) = \text{dist}(\eta - \zeta + U_{\mathbb{C}}, 2\pi i \mathbb{Z}^n), \quad \Delta(K * i\theta, K * i\phi) = \text{dist}(\theta - \phi + U, \mathbb{Z}^n), \quad (3.4)$$

where  $U_{\mathbb{C}}$  and  $U$  are the complex and real row spaces of  $M$ , respectively.

The logarithmic distance of  $v, w \in (\mathbb{C}^\times)^n$ , introduced in Equation (1.3), can now be expressed as

$$\delta_{\log}(v, w) = \Delta(\text{Log } v, \text{Log } w). \quad (3.5)$$

Hence the distances of orbits in the metric  $\delta_{\log}$  are given by

$$\delta_{\log}(\mathcal{O}_v, \mathcal{O}_w) = \Delta(T * \text{Log } v, T * \text{Log } w), \quad \delta_{\log}(\mathcal{C}_v, \mathcal{C}_w) = \Delta(K * \text{Log } v, K * \text{Log } w). \quad (3.6)$$

The action  $*$  splits into a real and imaginary part: for  $x = y + iz$  with  $y, z \in \mathbb{R}^d$  and  $\eta = \rho + 2\pi i\theta$  with  $\rho \in \mathbb{R}^n$ ,  $\theta \in \mathbb{C}^n$ , we have by Definition 3.1,

$$e^x * \eta = (\rho + M^T y) + i(2\pi\theta + M^T z) = e^y * \rho + e^{iz} * (2\pi i\theta), \quad (3.7)$$

hence

$$T * \eta = ((\rho + U)) + i(2\pi\theta + U + 2\pi\mathbb{Z}^n) / 2\pi i \mathbb{Z}^n.$$

Let  $\zeta = \tau + 2\pi i\phi$ . Together with Equations (3.3) and (3.4), we obtain the following formula for the  $\Delta$ -distance between  $*$ -orbits:

$$\begin{aligned} \Delta^2(T * \eta, T * \zeta) &= \text{dist}^2(\rho - \tau, U) + 4\pi^2 \Delta^2(K * i\theta, K * i\phi) \\ &= \text{dist}^2(\rho - \tau, U) + 4\pi^2 \text{dist}^2(\theta - \phi + U, \mathbb{Z}^n) \end{aligned} \quad (3.8)$$

and similarly,

$$\begin{aligned}\Delta^2(K * \eta, K * \zeta) &= \|\rho - \tau\|^2 + 4\pi^2 \Delta^2(K * i\theta, K * i\phi) \\ &= \|\rho - \tau\|^2 + 4\pi^2 \text{dist}^2(\theta - \phi + U, \mathbb{Z}^n).\end{aligned}\tag{3.9}$$

The contributions of the real parts,  $\text{dist}^2(\rho - \tau, U)$  in the  $T$ -case and  $\|\rho - \tau\|^2$  in the  $K$ -case, are easy to compute. On the other hand, the contribution of the imaginary parts  $\text{dist}^2(\theta - \phi + U, \mathbb{Z}^n)$ , turn out to be difficult to compute. We will show in Section 6 that approximating  $\text{dist}(\theta - \phi + U, \mathbb{Z}^n)$  within a subpolynomial factor is NP-hard, by providing a polynomial time reduction from the *closest vector problem* (CVP) to it. By contrast, deciding whether this distance equals zero can be done in polynomial time, see Remark 3.4. (This is analogous to the CVP problem, see Section 6.1.)

► **Remark 3.4.** Given a subspace  $U \subset \mathbb{R}^n$  by generators  $v_1, \dots, v_m \in \mathbb{Z}^n$  and  $t \in \mathbb{Q}^n$ , we can decide  $(t + U) \cap \mathbb{Z}^n \neq \emptyset$  in polynomial time. This can be seen by putting the matrix with columns  $v_1, \dots, v_m$  into Smith normal form, which is possible in polynomial time [54].

### 3.2 Approximation of orbit distances by linear forms in logarithms: $T$ -orbits

We use now invariant theory to analyze the orbits. Let  $M \in \mathbb{Z}^{k \times n}$  be a weight matrix and denote by  $H \in \mathbb{Z}^{k \times n}$  a matrix of rational invariants of  $H \in \mathbb{Z}^{k \times n}$ , see Proposition 2.2.

For  $v, w \in (\mathbb{C}^\times)^n$ , Equation (2.3) characterizes the equality of orbits  $\mathcal{O}_v = \mathcal{O}_w$  by  $H\eta = H\zeta$ , where  $\eta = \text{Log } v$  and  $\zeta = \text{Log } w$ . Our goal is a robust version of this: to show that the closeness of the corresponding logarithmic orbits  $T * \eta$  and  $T * \zeta$ , measured in terms of the metric  $\Delta$  on  $\mathbb{C}^n / 2\pi i \mathbb{Z}^n$ , is quantitatively related to the closeness of  $H\eta$  and  $H\zeta$ , measured in terms of the metric  $\Delta$  on the quotient space  $\mathbb{C}^k / 2\pi i \mathbb{Z}^k$ . The correction factors are provided by the minimum and maximum singular values of  $H$ .

► **Proposition 3.5.** *If  $\eta, \zeta \in \mathbb{C}^n / 2\pi i \mathbb{Z}^n$ , we have for the distance of  $T$ -orbits:*

$$\sigma_{\min}(H) \Delta(T * \eta, T * \zeta) \leq \Delta(H\eta, H\zeta) \leq \sigma_{\max}(H) \Delta(T * \eta, T * \zeta).$$

For  $v, w \in (\mathbb{C}^\times)^n$  we have

$$\sigma_{\min}(H) \delta_{\log}(\mathcal{O}_v, \mathcal{O}_w) \leq \Delta(H \text{Log } v, H \text{Log } w) \leq \sigma_{\max}(H) \delta_{\log}(\mathcal{O}_v, \mathcal{O}_w).$$

If  $\eta, \zeta$  are purely imaginary, then the same bounds hold for the distances of  $K$ -orbits.

**Proof.** We have  $\Delta(T * \eta, T * \zeta) = \text{dist}(\eta - \zeta + U_{\mathbb{C}}, 2\pi i \mathbb{Z}^n)$  by Equation (3.4), where  $U_{\mathbb{C}} = \ker H = \text{im } M^T$ . The first assertion follows from Lemma 2.5 with  $A = \{\eta - \zeta\}$  and  $B = 2\pi i \mathbb{Z}^n$ . The second assertion is just a rewriting of the first, using Equation (3.6). The statement on  $K$ -orbits follows analogously. ◀

### 3.3 Approximation of orbit distances by linear forms in logarithms: $K$ -orbits

Now the task is to provide an approximation for the distance  $\text{dist}(\mathcal{C}_v, \mathcal{C}_w)$  between  $K$ -orbits, similarly to Proposition 3.5. The analogous formula in Corollary 3.8 below is more complicated. It involves a priori upper and lower bounds on the norms of  $v$  and  $w$ . For  $0 < r \leq R$  we consider the closed region defined by

$$D_{r,R} := \{v \in \mathbb{C}^n \mid \forall i \in [n] \ r \leq |v_i| \leq R\}.\tag{3.10}$$

► **Lemma 3.6.** *Let  $v, w \in D_{r,R}$  and put  $\eta = \text{Log } v$ ,  $\zeta = \text{Log } w \in \mathbb{C}^n / 2\pi i \mathbb{Z}^n$ . Then*

$$\frac{2r}{\pi} \Delta(\eta, \zeta) \leq \|v - w\| \leq R \Delta(\eta, \zeta).$$

**Proof.** It is enough to show the equality for  $n = 1$ . We write here  $\eta = \rho + i\theta$ ,  $\zeta = \tau + i\phi$  (dropping the factor  $2\pi$  to simplify notation). The cosine theorem gives

$$|v - w|^2 = |v|^2 + |w|^2 - 2|v||w| \cos(\theta - \phi) = (|v| - |w|)^2 + 2|v||w|(1 - \cos(\theta - \phi)).$$

The first contribution on the right-hand side can be upper and lower bounded by

$$r^2(\rho - \tau)^2 \leq (|v| - |w|)^2 \leq R^2(\rho - \tau)^2$$

since the mean value theorem for exp implies (w.l.o.g.  $\rho < \tau$ )

$$r \leq |v| = e^\rho \leq \left| \frac{e^\tau - e^\rho}{\tau - \rho} \right| \leq e^\tau = |w| \leq R.$$

The second contribution on the right-hand side can be upper and lower bounded by

$$\frac{4r^2}{\pi^2} \text{dist}(\theta - \phi, 2\pi\mathbb{Z})^2 \leq 2|v||w|(1 - \cos(\theta - \phi)) \leq R^2 \text{dist}(\theta - \phi, 2\pi\mathbb{Z})^2,$$

using the inequality  $\frac{4}{\pi^2}\psi^2 \leq 2 - 2\cos\psi \leq \psi^2$  for  $\psi \in [-\pi, \pi]$ . Bringing the inequalities together and observing that  $4/\pi^2 \leq 1$ , we obtain

$$\frac{4r^2}{\pi^2} ((\rho - \tau)^2 + \text{dist}(\theta - \phi, 2\pi\mathbb{Z})^2) \leq |v - w|^2 \leq R^2 ((\rho - \tau)^2 + \text{dist}(\theta - \phi, 2\pi\mathbb{Z})^2),$$

which completes the proof. ◀

We can now relate  $\text{dist}(\mathcal{C}_v, \mathcal{C}_w)$  to  $\delta_{\log}(\mathcal{C}_v, \mathcal{C}_w) = \Delta(K * \text{Log } v, K * \text{Log } w)$ .

► **Proposition 3.7.** *For  $v, w \in D_{r,R}$  we have*

$$\frac{2r}{\pi} \Delta(K * \text{Log } v, K * \text{Log } w) \leq \text{dist}(\mathcal{C}_v, \mathcal{C}_w) \leq R \Delta(K * \text{Log } v, K * \text{Log } w).$$

**Proof.** Clearly, the  $K$ -action preserves  $D_{r,R}$ . Lemma 3.6 implies that

$$\frac{2r}{\pi} \Delta(\text{Log}(k \cdot v), w) \leq \|k \cdot v - w\| \leq R \Delta(\text{Log}(k \cdot v), w)$$

for all  $k \in K$ . The claim follows from by taking the infimum over  $k \in K$ . ◀

Combining Proposition 3.7 with Equation (3.8) and Proposition 3.5, we obtain the following corollary.

► **Corollary 3.8.** *Assume that  $0 < r \leq R$  and  $v, w \in D_{r,R}$ . Suppose we have  $\text{Log } v = \rho + 2\pi i\theta$  and  $\text{Log } w = \tau + 2\pi i\phi$ . Let  $H$  be a matrix of rational invariants as in Proposition 2.2. Then,*

$$\frac{4r^2}{\pi^2} (\|\rho - \tau\|^2 + \frac{4\pi^2}{\sigma_{\max}^2(H)} \Delta^2(H\theta, H\phi)) \leq \text{dist}^2(\mathcal{C}_v, \mathcal{C}_w) \leq R^2 (\|\rho - \tau\|^2 + \frac{4\pi^2}{\sigma_{\min}^2(H)} \Delta^2(H\theta, H\phi)).$$

We finally relate the separation parameters  $\text{sep}_K(d, n, B, b)$  and  $\text{sep}_T(d, n, B, b)$  defined in Equation (1.5).

► **Corollary 3.9.** *We have  $\text{sep}_T(d, n, B, b) \leq n 2^{O(b)} \text{sep}_K(d, n, B, b)$ .*

**Proof.** Suppose that  $\text{sep}_K(d, n, B, b) = \text{dist}(\mathcal{C}_v, \mathcal{C}_w)$ . We let  $r$  be the minimum and  $R$  be the maximum of  $|v_i|, |w_i|, i = 1, 2, \dots, n$ . Clearly,  $2^{-b} \leq r$ . If moreover  $\mathcal{O}_v \neq \mathcal{O}_w$ , then Proposition 3.7 implies

$$\text{dist}(\mathcal{C}_v, \mathcal{C}_w) \geq \frac{2r}{\pi} \Delta(K * \text{Log } v, K * \text{Log } w) \geq \frac{2r}{\pi} \Delta(T * \text{Log } v, T * \text{Log } w) \geq \frac{2r}{\pi} \text{sep}_T(d, n, B, b),$$

hence  $\text{sep}_T(d, n, B, b) \leq \frac{\pi}{2r} \text{dist}(\mathcal{C}_v, \mathcal{C}_w) \leq 2^{O(b)} \text{sep}_K(d, n, B, b)$ . On the other hand, if  $\mathcal{O}_v = \mathcal{O}_w$ , then Equation (2.4) implies that  $|v_i| \neq |w_i|$  for some  $i$ . In this case,  $\text{sep}_K(d, n, B, b) \geq \left| |v_i| - |w_i| \right| \geq 2^{-b}$ . We note that since  $|v_i/w_i| \leq 2^{2b}$ , we have  $\log(|v_i|/|w_i|) \leq 2b$ . Furthermore,  $\text{dist}(\theta, 2\pi\mathbb{Z}^n) \leq \sqrt{n}\pi$  for any  $\theta \in \mathbb{R}^n$ . Hence,  $\text{sep}_T(d, n, B, b) \leq \Delta(\text{Log } v, \text{Log } w) \leq \sqrt{n}(2b + \pi\sqrt{n})$ . Altogether,

$$\text{sep}_T(d, n, B, b) \leq 2\sqrt{n}b + \pi n \leq (\pi + 2)n2^b \leq (\pi + 2)n2^{2b} \text{sep}_K(d, n, B, b),$$

which proves the assertion. ◀

## 4 The separation hypotheses and the *abc*-conjecture

The goal of this section is to explain in detail the connection between the *abc*-conjecture and Number-Theoretic Hypothesis 1.11 and to prove Theorem 1.12, i.e., to show that Separation Hypothesis 1.7 and Number-Theoretic Hypothesis 1.11 are equivalent.

### 4.1 The *abc*-conjecture and the Number-Theoretic Hypothesis 1.11

Oesterlé and Masser's *abc* conjecture [75, Conjecture 3] claims the following.

► **Conjecture 4.1** (*abc* conjecture). *For every  $\varepsilon > 0$  there exists a constant  $\kappa_\varepsilon > 0$  such that for all nonzero coprime integers  $a, b, c$  satisfying  $a + b + c = 0$ , we have*

$$\max(|a|, |b|, |c|) \leq \kappa_\varepsilon \text{rad}(abc)^{1+\varepsilon}.$$

The radical  $\text{rad}(abc)$  is defined as the product of the distinct primes dividing  $abc$ , taken with multiplicity one.

This simple looking conjecture is one of the most powerful statements in number theory. In the words of Dorian Goldfeld [36]: “The *abc* conjecture is the most important unsolved problem in Diophantine analysis” since “it provides a way of reformulating an infinite number of Diophantine problems—and, if it is true, of solving them.”

Baker [8] proposed a more precise version of the conjecture: he conjectured that there exist absolute constants  $\kappa, \kappa' > 0$  (not depending on  $\varepsilon$ ) such that for all  $\varepsilon > 0$

$$\max(|a|, |b|, |c|) \leq \kappa \varepsilon^{-\kappa' \omega(ab)} \text{rad}(abc)^{1+\varepsilon},$$

where  $\omega(ab)$  denotes the number of distinct prime factors of  $ab$ . Moreover, Baker observed that<sup>5</sup> the minimum of the right-hand side over all  $\varepsilon > 0$  occurs when  $\varepsilon = \omega(ab)/\log N$  and suggested the following  $\varepsilon$ -free version of the conjecture:

---

<sup>5</sup> Baker attributes this observation to Andrew Granville.



► **Conjecture 4.2** (Baker's refinement of the *abc* conjecture, [7, Conjecture 3]). *There exist constants  $\kappa, \kappa' > 0$  such that for all nonzero coprime integers  $a, b, c$  satisfying  $a + b + c = 0$ , we have, setting  $N := \text{rad}(abc)$ ,*

$$\max(|a|, |b|, |c|) \leq \kappa N \left( \frac{\log N}{\omega(ab)} \right)^{\kappa' \omega(ab)}. \quad (4.1)$$

Baker's refinement of the *abc*-conjecture is closely related to lower bounds for linear forms in logarithms. The following was observed in [8]. We include a proof for the sake of completeness.

► **Theorem 4.3.** *Suppose Baker's refinement of the *abc* conjecture (Conjecture 4.2) is true. Then there exists a constant  $\kappa > 0$  such that for any list of positive integers  $v_1, v_2, \dots, v_n$  and any list of integers  $e_1, e_2, \dots, e_n$  satisfying  $v^e := \prod_{i=1}^n v_i^{e_i} \neq 1$ , we have*

$$\log |v^e - 1| \geq -\kappa \log(\max_i |e_i|) \sum_{i=1}^n \log v_i.$$

**Proof.** W.l.o.g.  $e_i \neq 0$  for all  $i$ ,  $E := \max |e_i| \geq 2$ , and  $u := v_1 v_2 \dots v_n \geq 3$ . We write  $v^e = \frac{a'}{b'} = \frac{a}{b}$ , where

$$a' := \prod_{e_i > 0} v_i^{e_i}, \quad b' := \prod_{e_i < 0} v_i^{-e_i}, \quad d := \gcd(a', b'), \quad a := a'/d, \quad b := b'/d.$$

We have  $a'b' = \prod v_i^{|e_i|}$ , hence  $\text{rad}(a'b') = \text{rad}(u)$ . On the other hand,  $a'b' = d^2 ab$ , which gives  $\text{rad}(a'b') = d \text{rad}(ab)$ . Hence  $\text{rad}(ab)$  is a divisor of  $\text{rad}(u)$ , which implies  $\omega(ab) \leq \omega(u)$ .

Setting  $c := a - b$ , we have  $(-a) + b + c = 0$  and  $\gcd(a, b, c) = 1$ . Since  $a, b$  are positive, we have  $\max(a, b, |c|) = \max(a, b)$ . We claim that  $|c| \leq u^E$ , where  $E := \max |e_i|$ . Indeed,  $c \leq a \leq a' \leq u^E$ , and similarly,  $-c \leq b \leq b' \leq u^E$ .

With the above estimates, we obtain

$$N = \text{rad}(abc) = \text{rad}(c) \text{rad}(ab) \leq |c| \text{rad}(u) \leq |c| u \leq u^{E+1}, \quad \omega(ab) \leq \omega(u). \quad (4.2)$$

Conjecture 4.2 implies

$$\max(a, b) \leq \kappa N \left( \frac{\log N}{\omega(ab)} \right)^{\kappa' \omega(ab)} \leq \kappa N (\log N)^{\kappa' \omega(ab)}.$$

Using Equation (4.2), we can bound this as

$$\max(a, b) \leq \kappa |c| u ((E+1) \log u)^{2\kappa' \omega(u)}.$$

Since  $|v^e - 1| = |c/b|$  we get

$$|v^e - 1| \geq |c| / \max(a, b) \geq \kappa^{-1} u^{-1} ((E+1) \log u)^{-2\kappa' \omega(u)}.$$

Taking logarithms of both sides, we obtain

$$\log |v^e - 1| \geq -\log \kappa - \log u - 2\kappa' \omega(u) (\log(E+1) + \log \log u).$$

It is known that  $\omega(u) \log \log u = O(\log u)$ , see [44, §22.10]. Using this, we conclude that indeed  $\log |v^e - 1| \geq -\kappa'' \log E \log u$  for a suitable constant  $\kappa'' > 0$ . ◀

Given  $v_1, v_2, \dots, v_n \in \mathbb{Z}_{>0}$  and  $e_1, e_2, \dots, e_n \in \mathbb{Z}$ , we denote by  $\Lambda(v, e)$  the *linear form in logarithms*

$$\Lambda(v, e) := \log v^e = e_1 \log v_1 + e_2 \log v_2 + \dots + e_n \log v_n. \quad (4.3)$$

A bound similar to the one in the previous theorem can also be given in terms of the quantity  $\log |\Lambda(v, e)|$ , which relates the *abc*-conjecture to Number-Theoretic Hypothesis 1.11.

► **Corollary 4.4.** *Assume Baker's refinement of the abc-conjecture is true. Then there exists a constant  $\kappa > 0$  such that for all  $v_1, v_2, \dots, v_n \in \mathbb{Z}_{>0}$  and  $e_1, e_2, \dots, e_n \in \mathbb{Z}$  with  $\Lambda(v, e) \neq 0$ , we have*

$$\log |\Lambda(v, e)| \geq -\kappa \log(\max_i |e_i|) \sum_{i=1}^n \log v_i.$$

In particular, Number-Theoretic Hypothesis 1.11 is true if  $v_1, \dots, v_n \in \mathbb{Q}_{>0}$ .

**Proof.** First note that the function  $h(x) := \log(x)/(x-1)$  is monotonically decreasing and satisfies  $\frac{1}{2} \leq h(x) \leq \frac{3}{2}$  on the interval  $[\frac{1}{2}, \frac{3}{2}]$ . This implies for  $|x-1| \leq \frac{1}{2}$ ,

$$\frac{1}{2} |\log(x)| \leq |x-1| \leq 2 |\log(x)|. \quad (4.4)$$

For showing the stated bound with positive integers  $v_i$ , we may assume without loss of generality that  $|v^e - 1| \leq \frac{1}{2}$ . Then Equation (4.4) implies  $\frac{1}{2} |\Lambda(v, e)| \leq |v^e - 1|$  and Theorem 4.3 gives the desired bound.

The assertion for positive rationals  $v_i = \frac{p_i}{q_i}$  follows from the observation that  $\Lambda(v, e) = \Lambda(\tilde{v}, \tilde{e})$ , where  $\tilde{v}$  is defined by concatenating  $p$  and  $q$ , and  $\tilde{e}$  by concatenating  $e$  and  $-e$ . ◀

► **Remark 4.5.** Conversely, lower bounds similar to the one in Corollary 4.4, together with their  $p$ -adic versions imply the *abc*-conjecture. For a prime  $p$ , let  $|x|_p$  denote the  $p$ -adic absolute value of  $x \in \mathbb{Q}$ , e.g.,  $|p^\nu y|_p = p^{-\nu}$  for  $y \in \mathbb{Z}$  not divisible by  $p$  and  $\nu \in \mathbb{Z}$ . Suppose the  $v_i$  are nonzero integers,  $e_i \in \mathbb{Z}$  and write  $v_1^{e_1} v_2^{e_2} \dots v_n^{e_n} = a/b$  for coprime integers  $a, b$ . If  $p$  is a prime dividing  $c := a - b$ , then  $|b|_p = 1$  and  $|c/b|_p = |c|_p < 1$ . Hence the  $p$ -adic logarithm  $\log(a/b) = \log(1 + c/b)$  exists and  $|\log(1 + c/b)|_p = |c|_p$ . In analogy with Equation (4.3), we define

$$|\Lambda(v, e)|_p := |\log(1 + c/b)|_p = |c|_p.$$

Baker considered the product of linear forms in logarithms

$$\Xi := \min(1, |\Lambda(v, e)|) \prod_{p \text{ prime}} \min(1, p |\Lambda(v, e)|_p),$$

where the product is over all primes. He proved that a slightly stronger lower bound for  $\Xi$  than the one of Corollary 4.4 is equivalent to his refinement of the *abc*-conjecture Equation (4.1); see [8, Section 5]. Moreover, any lower bound for  $\Xi$  implies a version of the *abc*-conjecture: Stewart and Yu [81], refining an earlier work by Stewart and Tijdemann [80], used this approach to prove the inequality

$$\log c < \kappa_\varepsilon N^{\frac{1}{3} + \varepsilon},$$

using the Baker-Wüstholz bound for the linear forms in logarithms [88] and its  $p$ -adic versions by van der Poorten [84]. We refer to [9, Chapter 3.7] (see also [38]) for a summary of these results.

In Theorem 4.3 and Corollary 4.4 we assumed that the vectors  $v_i$  to be positive rationals. More generally, for our purposes, we would like to allow the  $v_i$  to be nonzero Gaussian rationals. In the following, assume that  $\log : \mathbb{C}^\times \rightarrow \mathbb{C}$  denotes the principal branch of the logarithm, so  $\Im(\log z) \in (-\pi, \pi]$ , e.g.,  $\log(-1) = i\pi$ . In analogy, we define the linear form in logarithms

$$\Lambda(v, e) := e_1 \log v_1 + e_2 \log v_2 + \cdots + e_n \log v_n. \quad (4.5)$$

How small can  $|\Lambda(v, e)|$  be, provided it is nonzero?

In the introduction, we formulated Number-Theoretic Hypothesis 1.11 and noted that it follows from famous conjectures in number theory, such as Waldschmidt's conjectures [85, Conjecture 14.25], [86, Conjecture 4.14] or the Lang-Waldschmidt conjecture for Gaussian rationals [59, Introduction to Chapters X and XI].

We also note that there are analogues of the *abc*-conjecture in different number fields [11]. Recall that  $\mathbb{Z}[i]$  is a unique factorization domain, so for an element  $x \in \mathbb{Z}[i]$  there exist distinct prime elements  $p_1, p_2, \dots, p_k$ , exponents  $\mu_1, \mu_2, \dots, \mu_k \in \mathbb{Z}_{>0}$ , and a unit  $u \in \mathbb{Z}[i]^\times$  such that  $x = u \prod_{i=1}^k p_i^{\mu_i}$ . So we can define the radical of  $x$  as  $\text{rad}(x) := \prod_{i=1}^m p_i$ . We may conjecture the Gaussian integer analogue of Baker's *abc*-conjecture<sup>6</sup>: There exist constants  $\kappa, \kappa' > 0$  such that for all nonzero  $a, b, c \in \mathbb{Z}[i]$  with  $a + b + c = 0$  and  $\gcd(a, b, c) = 1$ , the radical  $N := \text{rad}(abc)$  satisfies

$$\max(|a|, |b|, |c|) \leq \kappa |N| \left( \frac{\log |N|}{\omega(ab)} \right)^{\kappa' \omega(ab)}. \quad (4.6)$$

With essentially the same arguments as in the proof of Theorem 4.3 and Corollary 4.4, one shows that Equation (4.6) implies Number-Theoretic Hypothesis 1.11.

Number-Theoretic Hypothesis 1.11 lower bounds the Euclidean distance of  $\Lambda(v, e)$  to 0. For our purposes, we need to lower bound the distance  $\Delta(\Lambda(v, e), 0) := \text{dist}(\Lambda(v, e), 2\pi i \mathbb{Z})$  in the  $\Delta$ -metric, see Equation (3.3). This follows easily.

► **Proposition 4.6.** *Assume Number-Theoretic Hypothesis 1.11. Suppose  $v_1, \dots, v_n \in \mathbb{Q}(i)^\times$  and  $e_1, \dots, e_n \in \mathbb{Z}$ . If  $\Delta(\Lambda(v, e), 0)$  is nonzero, then  $\Delta(\Lambda(v, e), 0) \geq \exp(-\text{poly}(n, B, b))$ , where  $B$  (resp.  $b$ ) is a bound for the bit-length of  $e_i$  (resp.  $v_i$ ).*

**Proof.** We have  $\text{dist}(\Lambda(v, e), 2\pi i \mathbb{Z}) = |\Lambda(v, e) - d i \pi|$  for some  $d \in \mathbb{Z}$  satisfying  $|d| \leq O(n 2^B b)$ . Note that  $\Lambda(-1, -d) = -d \log(-1) = -d i \pi$ . Therefore  $\Lambda(v, e) - d i \pi = \Lambda(\tilde{v}, \tilde{e})$ , where  $\tilde{v}$  is obtained from  $v$  by appending  $-1$  and  $\tilde{e}$  is obtained from  $e$  by appending  $-d$ . Now we apply Number-Theoretic Hypothesis 1.11. ◀

## 4.2 Equivalence of Separation Hypotheses with the Number-Theoretic Hypothesis 1.11

We prove here Theorem 1.12 based on Proposition 3.5, which relates the logarithmic distance between the orbits  $\mathcal{O}_v$  and  $\mathcal{O}_w$  to  $\Delta(H \log v, H \log w)$ . The essential observation is that  $\Delta(H \log v, H \log w)$  is determined by linear forms in logarithms.

**Proof of Theorem 1.12.** We will only prove the first equivalence; the second is shown in the same way, but using Corollary 3.8 instead of Proposition 3.5.

<sup>6</sup> We have not seen this conjecture in the literature.

( $\Rightarrow$ ) We first assume Number-Theoretic Hypothesis 1.11 and deduce Separation Hypothesis 1.7. Suppose  $v, w \in (\mathbb{Q}(i)^\times)^n$  have bit-lengths bounded by  $b$  and that the bit-length of the weight matrix  $M$  is bounded by  $B$ . By Proposition 2.2, there is a matrix  $H$  of rational invariants of bit-length  $\text{poly}(d, n, B)$ . We denote by  $\alpha_1^T, \dots, \alpha_k^T$  the rows of  $H \in \mathbb{Z}^{k \times n}$ .

Assume that  $\delta_{\log}(\mathcal{O}_v, \mathcal{O}_w) \neq 0$ . By Proposition 3.5,  $\Delta(H \log v, H \log w) = \Delta(H(\log v - \log w), 0)$  is nonzero. Hence at least one of the one-dimensional distances  $\Delta(\alpha_j^T(\log v - \log w), 0)$  must be nonzero. By Proposition 4.6 it is lower bounded by  $\exp(-\text{poly}(d, n, B, b))$ . Proposition 3.5 also implies

$$\delta_{\log}(\mathcal{O}_v, \mathcal{O}_w) \geq \frac{1}{\sigma_{\max}(H)} \Delta(H \log v, H \log w) \geq \frac{1}{\sigma_{\max}(H)} \Delta(\alpha_j^T(\log v - \log w), 0). \quad (4.7)$$

Moreover, by Lemma 2.7,  $\sigma_{\max}^{-1}(H)$  is lower bounded by  $\exp(-\text{poly}(d, n, B))$ . This finishes the proof of the first implication.

( $\Leftarrow$ ) Now we assume Separation Hypothesis 1.7 and deduce Number-Theoretic Hypothesis 1.11. Suppose  $v_i \in (\mathbb{Q}(i)^\times)^n$  have bit-length bounded by  $b$ , and  $e_i \in \mathbb{Z}$  have bit-lengths bounded by  $B$  such that  $\Lambda(v, e) \neq 0$ . We assume without loss of generality that  $\gcd(e_1, e_2, \dots, e_n) = 1$ . Put  $e := (e_1, e_2, \dots, e_n) \in \mathbb{Z}^n$  and consider the rational hyperplane  $e^\perp$  of  $\mathbb{R}^n$ . Note  $\Lambda(v, e) \neq 0$  expresses that  $\log v \notin e^\perp$ . There is a lattice basis  $\alpha_1, \alpha_2, \dots, \alpha_{n-1} \in \mathbb{Z}^n$  of  $e^\perp$  having bit-length  $\text{poly}(n, B)$ ; see Theorem 2.1 (a). Let  $M \in \mathbb{Z}^{d \times n}$  denote the matrix with rows  $\alpha_i$ , where  $d := n - 1$ .

We let  $(\mathbb{C}^\times)^d$  act on  $\mathbb{C}^n$  via the weight matrix  $M$ . The vector  $e$  generates the lattice of invariant Laurent monomials (recall  $\gcd(e_1, e_2, \dots, e_n) = 1$ ). In this case the matrix  $H \in \mathbb{Z}^{1 \times n}$  of rational invariants is just  $H = e^T$ .

Consider  $v := (v_1, v_2, \dots, v_n)$  and  $w := (1, \dots, 1) \in \mathbb{C}^n$ , so  $\log w = 0$ . Since  $H$  has only one row, we have  $\sigma_{\max}(H) = \sigma_{\min}(H) = \|e\|$ , which implies that the inequality from Proposition 3.5 is actually an equality:

$$\Delta(e^T \log v, 0) = \|e\| \delta_{\log}(\mathcal{O}_v, \mathcal{O}_w). \quad (4.8)$$

If  $\Delta(e^T \log v, 0) \neq 0$ , then by Separation Hypothesis 1.7,  $\delta_{\log}(\mathcal{O}_v, \mathcal{O}_w) \geq \exp(-\text{poly}(n, b, B))$ . Hence  $|\Lambda(v, e)| \geq \Delta(e^T \log v, 0) \geq \exp(-\text{poly}(n, b, B))$ .

When  $\Delta(e^T \log v, 0) = 0$ , we have  $0 \neq \Lambda(v, e) = e^T \log v \in 2\pi\mathbb{Z}$  so  $|\Lambda(v, e)| \geq 2\pi$ .  $\blacktriangleleft$

► **Remark 4.7.** Matveev's lower bound (1.8) combined with inequality Equation (4.7) shows that unconditionally

$$\text{sep}_T(d, n, B, b) \geq \exp(-B^{O(1)} b^{O(n)}).$$

The same lower bound holds for  $\text{sep}_K(d, n, B, b)$ .

## 5 Approximations of orbit distances

In this section, we provide polynomial time approximation algorithms for the orbit distance approximation problems. We solve Problem 1.2 in polynomial time and prove Theorem 1.8. The main theorem of this section is Theorem 5.2.

### 5.1 Approximate orbit distance problems

So far, we discussed four *group actions*: the actions of  $T = (\mathbb{C}^\times)^d$  or  $K = (S^1)^d$  on the spaces  $V = \mathbb{C}^n$  and  $\mathbb{C}^n/2\pi i\mathbb{Z}^n$  given by a weight matrix  $M \in \mathbb{Z}^{d \times n}$ . We also defined a *metric*  $\delta$  on

the resulting spaces of orbits. Let us list all these cases:

$$(T, \delta_{\log}) \quad \delta_{\log}(\mathcal{O}_v, \mathcal{O}_w) := \Delta(\text{Log}(\mathcal{O}_v), \text{Log}(\mathcal{O}_w))$$

$$(K, \text{dist}) \quad \text{dist}(\mathcal{C}_v, \mathcal{C}_w) := \inf_{k \in K} \|k \cdot v - w\|$$

$$(T, \Delta) \quad \Delta(T * \eta, T * \zeta) := \inf_{t \in T} \Delta(t * \eta, \zeta)$$

$$(K, \Delta) \quad \Delta(K * \eta, K * \zeta) := \inf_{k \in K} \Delta(k * \eta, \zeta),$$

see Definition 3.1 for the  $*$ -action and Equation (3.3) for the  $\Delta$ -metric. We define an orbit distance approximation problem in each setting.

► **Definition 5.1.** *In each of above case, abbreviated  $(G, \delta)$ , the orbit distance approximation problem  $\text{ROP}(G, \delta)_\gamma$  with approximation factor  $\gamma \geq 1$  is the following: on input a weight matrix  $M \in \mathbb{Z}^{d \times n}$  and  $v, w \in (\mathbb{Q}(\mathbf{i})^\times)^n$ , resp.  $\eta, \zeta \in \mathbb{Q}^n + 2\pi \mathbf{i} \mathbb{Q}^n$ , compute a number  $D \in \mathbb{Q}_{\geq 0}$  such that*

$$\delta(G \cdot v, G \cdot w) \leq D \leq \gamma \delta(G \cdot v, G \cdot w).$$

We think of  $\gamma$  as a function of the input bit-length, determined by the usual parameters  $d, n, B, b$ .

We can now precisely state the main algorithmic result of this section. Recall from Equation (3.10) the regions  $D_{r,R}$  defined for  $0 < r \leq R$ . Note that Theorem 1.8 follows from the second and third part of the next result.

► **Theorem 5.2.** *There are approximation factors  $\gamma_i = \gamma_i(d, n, B)$  bounded  $\exp(\text{poly}(d, n, B))$  such that:*

- (a)  $\text{ROP}(G, \delta)_{\gamma_1}$  admits a polynomial time algorithm in the cases  $(T, \Delta)$  and  $(K, \Delta)$ .
- (b)  $\text{ROP}(T, \delta_{\log})_{\gamma_2}$  admits a polynomial time algorithm, if Separation Hypothesis 1.7 holds.
- (c)  $\text{ROP}(K, \text{dist})_\gamma$ , when restricted to inputs in  $D_{r,R}$ , admits a polynomial time algorithm with approximation factor  $\gamma = \frac{R}{r} \gamma_3(d, n, B)$ , if Separation Hypothesis 1.6 holds.

One may wonder why the first part of Theorem 5.2 holds unconditionally. This is explained by the next result, which gives an unconditional separation of orbits with respect to the  $\Delta$ -metric.

► **Proposition 5.3.** *If  $\eta, \zeta \in \mathbb{Q}^n + 2\pi \mathbf{i} \mathbb{Q}^n$  with  $T * \eta \neq T * \zeta$ , then*

$$\Delta(T * \eta, T * \zeta) \geq \exp(-\text{poly}(d, n, b, B)).$$

The analogous statement holds for the action of  $K$ .

**Proof.** We write  $\eta = \rho + 2\pi \mathbf{i} \theta$ ,  $\zeta = \tau + 2\pi \mathbf{i} \phi$  and recall from Equation (3.8) that

$$\Delta^2(T * \eta, T * \zeta) = \text{dist}^2(\rho - \tau, U) + 4\pi^2 \Delta^2(K * \mathbf{i} \theta, K * \mathbf{i} \phi).$$

The first contribution  $\text{dist}^2(\rho - \tau, U)$ , if nonzero, is easily lower bounded using the Gram-Schmidt orthogonalization. So we may assume  $K * \mathbf{i} \theta \neq K * \mathbf{i} \phi$ . In this case, Proposition 3.5 implies  $\text{dist}(H(\theta - \phi), \mathbb{Z}^k) \neq 0$  and

$$\Delta(K * \mathbf{i} \theta, K * \mathbf{i} \phi) \geq \frac{1}{\sigma_{\max}(H)} \text{dist}(H(\theta - \phi), \mathbb{Z}^k),$$

where  $H$  is a matrix of rational invariants whose bit-length is bounded by Proposition 2.2. We can upper bound  $\sigma_{\max}(H)$  by Proposition 2.2 and Lemma 2.7, which provides the desired lower bound for the  $T$ -action. The distance of  $H(\theta - \phi)$  to  $\mathbb{Z}^k$  is also at most exponentially small since it is lower bounded by the reciprocal of the largest denominator of the entries of the rational vector  $H(\theta - \phi)$ . The proof for the  $K$ -action is analogous.  $\blacktriangleleft$

We also show that our algorithms for approximating the distance between orbits can be modified to produce a group element witnessing the  $\gamma$ -proximity of orbits. To avoid repetition, we only state this for the case  $(T, \delta_{\log})$ , even though this can also be achieved in the other settings in a similar way.

► **Theorem 5.4.** *Assume Separation Hypothesis 1.7. On input  $M, v, w$ , one can compute in polynomial time a vector  $x \in \mathbb{C}^d$  that satisfies*

$$\delta_{\log}(e^x \cdot v, w) \leq \gamma \delta_{\log}(\mathcal{O}_v, \mathcal{O}_w) \quad (5.1)$$

if  $\mathcal{O}_v \neq \mathcal{O}_w$ . If  $\mathcal{O}_v = \mathcal{O}_w$ , then the algorithm correctly identifies this case instead of returning  $x$ . The approximation parameter is exponentially bounded in the bit-length of the input.

In the next section, we introduce the distance computation problem SLDP and exhibit a polynomial time algorithm for it. Theorems 5.2 and 5.4 are then proved in Section 5.3 by reducing the orbit distance approximation problems to SLDP.

## 5.2 The subspace-to-cubic-lattice distance problem

The orbit distance approximation problems are closely related to the following problem.

► **Definition 5.5.** *The subspace-to-cubic-lattice distance approximation problem  $SLDP_\gamma$  with approximation factor  $\gamma$  is the task of computing on input*

- a target vector  $t \in \mathbb{Q}^n$ ,
- a subspace  $U \subset \mathbb{R}^n$ , spanned by given linearly independent input vectors  $u_1, \dots, u_{n-k} \in \mathbb{Z}^n$ ,
- a number  $D \in \mathbb{Q}_{\geq 0}$  such that

$$\text{dist}(t + U, \mathbb{Z}^n) \leq D \leq \gamma \text{dist}(t + U, \mathbb{Z}^n).$$

We think of  $\gamma$  a function of the input bit-length.

► **Remark 5.6.** The problem  $SLDP_\gamma$  is easy if  $U = 0$ . Indeed,  $\text{dist}^2(t, \mathbb{Z}^n)$  can be exactly and efficiently computed by rounding the coordinates of  $t$  to integers. At the other extreme, the problem is trivial if  $U = \mathbb{R}^n$ . However, we will show in Section 6, that  $SLDP_\gamma$  is NP-hard for constant  $\gamma$  (or more generally, an almost polynomial  $\gamma$ ), by providing a reduction from the *closest vector problem* to it.

► **Proposition 5.7.** *There is an approximation factor  $\gamma$  bounded exponentially in the input bit-length such that  $SLDP_\gamma$  admits a polynomial time algorithm. This algorithm can be modified to compute on input  $(t, U)$  a witness of proximity  $(u, \alpha) \in U \times \mathbb{Z}^n$  such that*

$$\text{dist}(t + u, \alpha) \leq \gamma \text{dist}(t + U, \mathbb{Z}^n).$$

**Proof.** For given  $M \in \mathbb{Z}^{(n-k) \times n}$  we compute  $H \in \mathbb{Z}^{k \times n}$  such that  $U := \text{im } M^T = \ker H$  in polynomial time, see Proposition 2.2. When applying Lemma 2.5 to  $A = \{t\}$  and  $B = \mathbb{Z}^n$ , we get

$$\text{dist}(t + U, \mathbb{Z}^n) \leq D_1 := \frac{\text{dist}(Ht, \mathbb{Z}^k)}{\sigma_{\min}(H)} \leq \frac{\sigma_{\max}(H)}{\sigma_{\min}(H)} \text{dist}(t + U, \mathbb{Z}^n). \quad (5.2)$$

Hence,  $D_1$  approximates  $\text{dist}(t + U, \mathbb{Z}^n)$  within a factor of  $\sigma_{\max}(H)/\sigma_{\min}(H)$ . By Lemma 2.7 we can compute a rational approximation  $D$  of  $D_1$  within a factor of 2 in polynomial time. Moreover  $\sigma_{\max}(H)/\sigma_{\min}(H)$  is exponentially upper bounded in the input bit-length.

We show now how to compute the witness. By rounding the entries of  $Ht$  to nearest integers, we can compute an integral vector  $\beta \in \mathbb{Z}^k$  such that  $\text{dist}(Ht, \mathbb{Z}^k) = \text{dist}(Ht, \beta)$ . Recall that  $H(\mathbb{Z}^n) = \mathbb{Z}^k$ . We can further efficiently compute  $\alpha \in \mathbb{Z}^n$  such that  $H\alpha = \beta$ . By projecting  $t - \alpha$  orthogonally onto  $U$ , we can compute in polynomial time  $u \in U$  such that  $\text{dist}(t + u, \alpha) = \text{dist}(t + U, \alpha)$ . ◀

► **Remark 5.8.** In Section 6.1, we analyse an alternative algorithm (based on the LLL-algorithm), which solves  $\text{SLDP}_\gamma$  for  $\gamma = 2^{O(n)}$  in polynomial time (Corollary 6.4). In comparison, the algorithm given in Proposition 5.7 provides an approximation factor of  $\gamma = \kappa(H) := \sigma_{\max}(H)/\sigma_{\min}(H)$  which is bounded by  $2^{n(B + \log_2 n)}$  in the worst case.

### 5.3 Proof of Theorem 5.2

The first part of this theorem follows from the following reduction, jointly with Proposition 5.7.

► **Lemma 5.9.** *Both  $\text{ROP}(T, \Delta)_{2\gamma}$  and  $\text{ROP}(K, \Delta)_{2\gamma}$  reduce to  $\text{SLDP}_\gamma$  in polynomial time, for any  $\gamma \geq 1$ . Hereby the ambient dimension does not change.*

**Proof.** We will only prove the reduction for  $G = K$  since the  $G = T$  case is analogous. Suppose that  $M, \eta, \zeta$  are given as input, write  $\eta = \rho + 2\pi i \theta, \zeta = \tau + 2\pi i \phi$ , set  $t := \theta - \phi$ , and  $U := \text{im } M^T$ . We can compute an integral basis of  $U$  in polynomial time. Equation (3.4) and Equation (3.9) imply

$$\Delta^2(K * \eta, K * \zeta) = \|\rho - \tau\|^2 + 4\pi^2 \text{dist}^2(t + U, \mathbb{Z}^n).$$

If  $\text{dist}(t + U, \mathbb{Z}^n) \leq D \leq \gamma \text{dist}(t + U, \mathbb{Z}^n)$ , then

$$\Delta^2(K * \eta, K * \zeta) \leq \|\rho - \tau\|^2 + 4\pi^2 D^2 \leq \gamma^2 \|\rho - \tau\|^2 + 4\pi^2 D^2 \leq \gamma^2 \Delta^2(K * \eta, K * \zeta),$$

hence  $D' := \sqrt{\|\rho - \tau\|^2 + 4\pi^2 D^2}$  is a  $\gamma$ -approximate solution of  $\text{ROP}(K, \Delta)$ . We compute a rational number  $D''$  such that  $D' \leq D'' \leq 2D'$ . Then  $\Delta(K * \eta, K * \zeta) \leq D'' \leq 2\gamma \Delta(K * \eta, K * \zeta)$ , ◀

A reduction in the reverse direction will be needed later for the hardness proof, see Lemma 5.12. The next lemma studies that happens with  $\text{ROP}$ , when the metric  $\delta_{\log}$ , resp.  $\text{dist}$ , are replaced by  $\Delta$ . This lemma proves the second and third part of Theorem 5.2 by reducing it to its first part.

► **Lemma 5.10.** *Let  $\gamma \geq 1$  be any function of the inputs, let  $0 < r \leq R$  and put  $\gamma_1 := \frac{2R}{r} \gamma$ .*

- (a) *There is a polynomial time reduction from  $\text{ROP}(T, \delta_{\log})_{2\gamma}$  to  $\text{ROP}(T, \Delta)_\gamma$ , provided Separation Hypothesis 1.7 holds.*
- (b) *There is a polynomial time reduction from  $\text{ROP}(K, \text{dist})_{\gamma_1}$ , when restricted to inputs in  $D_{r,R}$ , to  $\text{ROP}(K, \Delta)_\gamma$ , provided Separation Hypothesis 1.6 holds.*

**Proof.** We only give the proof for part two since the proof of part one is analogous. The only difference between the two cases is that in the second we need to use Proposition 3.7 which gives an equivalence between the metrics  $\text{dist}$  and  $\delta_{\log}$  within an  $O(R/r)$ -factor, and in the first we do not need it.



Given  $v, w \in D_{r,R}$ , we test in polynomial time whether  $\mathcal{C}_v = \mathcal{C}_w$  (see Section 2.2). If  $\mathcal{C}_v = \mathcal{C}_w$  we return  $D = 0$ .<sup>7</sup> Otherwise, the reduction just consists of replacing the exact values  $\eta := \text{Log } v$  and  $\zeta := \text{Log } w$  by approximations

$$\|\tilde{\eta} - \eta\| < \kappa/2, \quad \|\tilde{\zeta} - \zeta\| < \kappa/2, \quad (5.3)$$

where the accuracy is prescribed by the separation parameter. For  $\kappa := \frac{1}{9R} \text{sep}_K(d, n, B, b)$ , this can be achieved in polynomial time assuming Separation Hypothesis 1.6. For convenience, we denote the orbit distances of the exact vectors and the approximations computed by

$$\Delta := \Delta(K * \eta, K * \zeta), \quad \tilde{\Delta} := \Delta(K * \tilde{\eta}, K * \tilde{\zeta}).$$

It suffices to prove that  $\tilde{\Delta} \leq D \leq \gamma \tilde{\Delta}$  implies

$$\text{dist}(\mathcal{C}_v, \mathcal{C}_w) \leq \frac{9RD}{8} \leq \gamma_1 \text{dist}(\mathcal{C}_v, \mathcal{C}_w),$$

where  $\gamma_1 = 2R\gamma/r$ . We now prove this implication.

By Proposition 3.7, we have

$$\frac{2r}{\pi} \Delta \leq \text{dist}(\mathcal{C}_v, \mathcal{C}_w) \leq R \Delta, \quad (5.4)$$

which implies  $\Delta \geq 9\kappa$  by our choice of  $\kappa$ . With Equation (5.3) and the triangle inequality, we get

$$|\tilde{\Delta} - \Delta| < \kappa, \quad \frac{\kappa}{\Delta} \leq \frac{1}{9}.$$

By dividing through  $\Delta$ , this implies  $\frac{8}{9} < \frac{\tilde{\Delta}}{\Delta} < \frac{10}{9}$ . From the assumption  $\tilde{\Delta} \leq D \leq \gamma \tilde{\Delta}$ , we infer

$$\Delta \leq \frac{9}{8} D \leq \frac{10}{8} \gamma \Delta.$$

Equation (5.4) implies

$$\text{dist}(\mathcal{C}_v, \mathcal{C}_w) \leq \frac{9RD}{8} \leq \frac{10\pi R\gamma}{16r} \text{dist}(\mathcal{C}_v, \mathcal{C}_w) < \frac{2R\gamma}{r} \text{dist}(\mathcal{C}_v, \mathcal{C}_w),$$

which completes the proof.  $\blacktriangleleft$

We can prove Theorem 5.4 in a similar way.

**Proof of Theorem 5.4.** On input  $v, w \in (\mathbb{C}^\times)^n$ , we first test in polynomial time whether  $\mathcal{O}_v = \mathcal{O}_w$  (see Section 2.2). If this is the case, we return  $D = 0$ . Otherwise, as in Equation (5.3), we compute approximations  $\tilde{\eta}$  and  $\tilde{\zeta}$  of the exact values  $\eta := \text{Log } v$  and  $\zeta := \text{Log } w$  with accuracy  $\kappa = \text{sep}_T(d, n, B, b)/2$ : Separation Hypothesis 1.7 guarantees that this can be done in polynomial time. We have for all  $x \in \mathbb{C}^d$

$$\Delta(e^x * \eta, \zeta) \geq \Delta(T * \eta, T * \zeta) = \delta_{\log}(\mathcal{O}_v, \mathcal{O}_w) \geq \text{sep}_T(d, n, B, b) = 2\kappa.$$

Hence, using the triangle inequality, we get for all  $x \in \mathbb{C}^d$

$$\left| \Delta(e^x * \tilde{\eta}, \tilde{\zeta}) - \Delta(e^x * \eta, \zeta) \right| < \kappa \leq \frac{1}{2} \Delta(e^x * \eta, \zeta),$$

<sup>7</sup> To get a Karp reduction, we can simply return the  $\text{ROP}(K, \Delta)$  instance  $(\eta, \zeta) = (0, 0)$ .

therefore

$$\frac{1}{2} \Delta(e^x * \eta, \zeta) < \Delta(e^x * \tilde{\eta}, \tilde{\zeta}) < \frac{3}{2} \Delta(e^x * \eta, \zeta). \quad (5.5)$$

Write  $\tilde{\eta} = \tilde{\rho} + 2\pi i \tilde{\theta}$  and  $\tilde{\zeta} = \tilde{\tau} + 2\pi i \tilde{\phi}$  and recall from Equation (3.8) that

$$\Delta^2(T * \tilde{\eta}, T * \tilde{\zeta}) = \text{dist}^2(\tilde{\rho} - \tilde{\tau}, U) + 4\pi^2 \text{dist}^2(\tilde{\theta} - \tilde{\phi} + U, \mathbb{Z}^n).$$

For the first summand, we compute the orthogonal projection  $u_1$  of  $\tilde{\rho} - \tilde{\tau}$  onto  $U$ . For the second summand, we use Proposition 5.7 to compute a witness  $(u_2, \alpha)$  such that

$$\text{dist}(\tilde{\theta} - \tilde{\phi} + u_2, \alpha) \leq \gamma \text{dist}(\tilde{\theta} - \tilde{\phi} + U, \mathbb{Z}^n) \quad (5.6)$$

for some  $\gamma$  satisfying  $\gamma = \exp(\text{poly}(B, n))$ . We then compute  $y, z \in \mathbb{R}^d$  such that  $M^T y = u_1$  and  $M^T z = u_2$  and compute the group element  $x := y + 2\pi i z$ . Then we have by Equation (3.3)

$$\begin{aligned} \Delta^2(e^x * \tilde{\eta}, \tilde{\zeta}) &= \text{dist}^2(\tilde{\rho} - \tilde{\tau}, M^T y) + 4\pi^2 \text{dist}^2(\tilde{\theta} - \tilde{\phi} + u_2, \alpha) \\ &\leq \text{dist}^2(\tilde{\rho} - \tilde{\tau}, U) + 4\pi^2 \gamma^2 \text{dist}^2(\tilde{\theta} - \tilde{\phi} + U, \mathbb{Z}^n) \\ &\leq \gamma^2 \Delta^2(T * \tilde{\eta}, T * \tilde{\zeta}), \end{aligned}$$

where we used Equation (5.6) for the second inequality. Combining with Equation (5.5), we obtain

$$\Delta(e^x * \eta, \zeta) \leq 2\Delta(e^x * \tilde{\eta}, \tilde{\zeta}) \leq 2\gamma \Delta(T * \eta, T * \zeta),$$

which finishes the proof.  $\blacktriangleleft$

► **Remark 5.11.** The algorithm underlying the reduction in the proof of Lemma 5.10 runs in time polynomial in the input bit-length and  $\log \text{sep}_T^{-1}(d, n, B, b)$ . Thus it is a polynomial time algorithm if and only if the separation hypothesis is true.

## 5.4 Reductions from SLDP to ROP

The following reductions are needed in the next section. While the proofs are straightforward perturbation arguments, dealing with the relative errors requires some care, so that we write down the detailed arguments at least in one case.

► **Lemma 5.12.** *Let  $(G, \delta)$  denote any of the four cases in Definition 5.1. Then there is a polynomial-time reduction from  $\text{SLDP}_{2\gamma}$  to  $\text{ROP}(G, \delta)_\gamma$ , for any  $\gamma \geq 1$ . Hereby the ambient dimension is preserved.*

**Proof.** We only provide the argument for  $(G, \delta) = (K, \text{dist})$ , the case  $(T, \delta_{\log})$  being similar and the remaining two cases being trivial. Consider an instance  $U, t$  of SLDP, where  $U \subset \mathbb{R}^n$  is a subspace of dimension  $d := n - k$  given as the row span of  $M \in \mathbb{Z}^{(n-k) \times n}$ , and  $t \in \mathbb{Q}^n$ . The matrix  $M$  defines actions of  $T = (\mathbb{C}^\times)^d$  and  $K = (S^1)^d$  on  $\mathbb{C}^n$ . The essential connection is Equation (3.4), which implies that

$$\text{dist}(t + U, \mathbb{Z}^n) = \Delta(K * i t, K * 0) = \Delta(T * i t, T * 0). \quad (5.7)$$

By Proposition 5.3, there is a separation function  $0 < \epsilon \leq 1$  such that  $\log \epsilon^{-1}$  polynomially bounded in the parameters  $d, n, B, b$  and  $\Delta(K * 2\pi i t, K * 0) \geq \epsilon$  when the orbits are different.

By Lemma 2.9 we compute  $v \in (\mathbb{Q}(i)^\times)^n$  such that  $\|v - \text{Exp}(2\pi i t)\| < \kappa := \epsilon/18$ . Hence  $v \in D_{r,R}$  for  $r = 17/18$  and  $R = 19/18$ . Lemma 5.13 below expresses a Lipschitz property of Log and gives

$$\|\text{Log } v - 2\pi i t\|_\infty < \frac{\kappa}{1-\kappa} < \frac{18}{17} \kappa.$$

Thus the distances  $\Delta_1 := \Delta(K * \text{Log } v, 0)$  and  $\Delta_2 := \Delta(K * 2\pi i t, 0)$  of the corresponding  $K$ -orbits to the zero orbit are close:  $|\Delta_1 - \Delta_2| < \frac{18}{17} \kappa$ . Using  $\Delta_2 \geq \epsilon = 18\kappa$ , this bounds the relative errors as  $\frac{|\Delta_1 - \Delta_2|}{\Delta_2} \leq \frac{1}{17}$ , and hence

$$\Delta_2 \leq \frac{17}{16} \Delta_1 \leq \frac{18}{17} \Delta_2.$$

Note that  $\text{dist}(t + U, \mathbb{Z}^n) = \frac{1}{2\pi} \Delta_2$  by Equation (5.7). Consider the vector  $w := (1, \dots, 1)$ . Proposition 3.7 implies, using  $v \in D_{r,R}$  with  $r = 17/18$  and  $R = 19/18$ , that

$$\frac{2}{\pi} \frac{17}{18} \Delta_1 \leq \text{dist}(\mathcal{C}_v, \mathcal{C}_w) \leq \frac{19}{18} \Delta_1. \quad (5.8)$$

Now assume that  $\text{dist}(\mathcal{C}_v, \mathcal{C}_w) \leq D \leq \gamma \text{dist}(\mathcal{C}_v, \mathcal{C}_w)$ . Then,

$$\Delta_2 \leq \frac{17}{16} \Delta_1 \leq \frac{17}{16} \frac{\pi}{2} \frac{18}{17} D \leq \frac{17}{16} \frac{\pi}{2} \frac{18}{17} \gamma \frac{19}{18} \frac{18}{17} \Delta_2 \leq 2 \frac{\Delta_2}{2\pi}.$$

This means that

$$\text{dist}(t + U, \mathbb{Z}^n) \leq \frac{9}{32} D \leq 2\gamma \text{dist}(t + U, \mathbb{Z}^n),$$

which completes the proof of the reduction.  $\blacktriangleleft$

► **Lemma 5.13.** *For any  $t \in \mathbb{R}^n$  and  $z \in \mathbb{C}^n$ ,  $\|z - \text{Exp}(2\pi i t)\| < \kappa < 1$  implies that  $\|\text{Log}(z) - 2\pi i t\| < \kappa/(1-\kappa)$ .*

**Proof.** It is sufficient to verify the claim for  $n = 1$ , in which case the claim follows using  $|\text{Log}' z| = |\frac{1}{z}| \leq \frac{1}{1-\kappa}$ .  $\blacktriangleleft$

## 6 Hardness of robust orbit problems

The main goal of this section is to prove Theorem 1.4. Actually, we prove the following, slightly more general result, that covers all four settings introduced in Section 5.1.

► **Theorem 6.1.** *There is  $c > 0$  such that the robust orbit distance approximation problem  $\text{ROP}(G, \delta)_\gamma$  is NP-hard for the approximation factor  $\gamma(n) = n^{c/\log \log n}$ , for each of the four combinations of group actions and metrics considered in Definition 5.1.*

Of course, this implies the hardness of the decisional version of the considered orbit distance approximation problems, namely:

► **Corollary 6.2.** *There is no polynomial time algorithm that on input  $M, v, w$  and  $\epsilon > 0$  decides*

$$\text{dist}(\mathcal{C}_v, \mathcal{C}_w) \leq \epsilon,$$

*unless  $P = NP$ . The analogous result also holds for the logarithmic distance  $\delta_{\log}(\mathcal{O}_v, \mathcal{O}_w)$  of the orbits of the  $T$ -action.*

**Proof.** By binary search, making oracle calls to a hypothetical polynomial time algorithm for the above decision problem, we can compute an approximation to  $\text{dist}(\mathcal{C}_v, \mathcal{C}_w)$  within any constant factor, say two. This provides a polynomial time algorithm for the robust orbit distance approximation problem with approximation factor  $\gamma = 2$ . By Theorem 6.1, this implies  $P = NP$ .  $\blacktriangleleft$

The proof of Theorem 6.1 relies on the NP-hardness of the closest vector problem  $\text{CVP}_\gamma$  of algorithmic lattice theory due to Dinur et al. [29]. We discuss the closest vector problem in Section 6.1. The proof of Theorem 6.1 then proceeds by exhibiting a polynomial time reduction from  $\text{CVP}_{2\gamma}$  to  $\text{ROP}(G, \delta)_\gamma$ . This goes in two steps: from Lemma 5.12 we already know that  $\text{SLDP}_{2\gamma}$  can be reduced to  $\text{ROP}(G, \delta)_\gamma$  in polynomial time. The main ingredient of the proof is a polynomial time reduction of  $\text{CVP}_\gamma$  to  $\text{SLDP}_\gamma$ , which is presented in Section 6.2. This relies on Theorem 1.5 on lattice lifting, whose proof is given in Section 6.3.

## 6.1 The closest vector problem

This problem attracted a lot of research due to lattice based cryptography, such as GGH, NTRU and homomorphic encryption [37, 47, 48, 35]. These cryptosystems are conjectured to be secure against quantum computers and rely on a (conjectured) hardness of CVP.

Throughout the section,  $m$  will always denote the dimension of a CVP instance and  $n$  will always denote the dimension of a SLDP instance.

► **Definition 6.3.** The closest vector problem *with approximation factor*  $\gamma \geq 1$ , denoted  $\text{CVP}_\gamma$ , is the task of computing on input

- a target vector  $t \in \mathbb{Q}^m$ ,
- a lattice  $\mathcal{L}$  spanned by the columns of generator matrix  $G \in \mathbb{Z}^{m \times m}$  with  $\det G \neq 0$ ,  
a number  $D \in \mathbb{Q}_{\geq 0}$  such that  $\text{dist}(t, \mathcal{L}) \leq D \leq \gamma \text{dist}(t, \mathcal{L})$ .

The first observation is that  $\text{SLDP}_\gamma$  can be easily reduced to  $\text{CVP}_\gamma$ . Indeed, if  $P : \mathbb{R}^n \rightarrow U^\perp$  denotes the orthogonal projection along  $U$ , then  $\text{dist}(t + U, \mathbb{Z}^n) = \text{dist}(P(t), \mathcal{L})$  by Corollary 2.6. Here,  $\mathcal{L} := P(\mathbb{Z}^n)$  is a lattice by Lemma 2.13, and we can compute a lattice basis of  $\mathcal{L}$  in polynomial time. This reduces  $\text{SLDP}_\gamma$  to  $\text{CVP}_\gamma$ . We note that this reduction preserves  $\gamma$ . Moreover, the ambient dimension  $n$  of the SLDP instance upper bounds the lattice dimension  $m$  of the constructed CVP instance.

The key contribution in the proof of Theorem 6.1 is a polynomial time reduction in the reverse direction, see Theorem 6.7.

$\text{CVP}_\gamma$  is known to be NP-hard for a nearly polynomial approximation ratio,  $\gamma = m^{c/\log \log m}$  for some  $c > 0$ , see [29]. On the other hand, it is well known that  $\text{CVP}_\gamma$  can be solved in polynomial time with approximation factor  $\gamma(m) = 2^{O(m)}$  by the well-known LLL-basis reduction algorithm of [61]. This implies the following.

► **Corollary 6.4.**  $\text{SLDP}_\gamma$  admits a polynomial time algorithm for the approximation factor  $\gamma(n) = 2^{O(n)}$ .

If we do not insist on polynomial time algorithms, then the lattice element  $\alpha \in \mathcal{L}$  closest to the target vector  $t$  can be computed exactly. Kannan [53] showed that there is an algorithm for doing so, that runs in time  $2^{O(m \log m)}$  times the input size. Combining this with the above reduction, we obtain:

► **Corollary 6.5.**  $\text{SLDP}_1$  can be solved by a polynomial time algorithm when  $n$  is fixed, if we allow exact computation of square roots.

We use this to show polynomial time feasibility of Problem 1.2 and Problem 1.3 when  $n$  is fixed.

**Proof of Theorem 1.9 and Theorem 1.10.** Matveev's bound (1.8) implies Number-Theoretic Hypothesis 1.11 if  $n$  is fixed. Then also Separation Hypothesis 1.7 holds if  $n$  is fixed, as shown by the proof of Theorem 1.12. Note that the reductions in Lemma 5.10 preserve the ambient dimension. The assertion follows now with Corollary 6.5. ◀

► **Remark 6.6.** The interesting regime of  $\text{CVP}_\gamma$  is when  $\gamma$  is polynomial in  $m$ . Lattice based cryptosystems rely on the hardness of  $\text{CVP}_\gamma$  in this regime, but as of now, proving this is a major open problem in the area. In fact it is known that  $\text{CVP}_\gamma$  is in  $\text{NP} \cap \text{coNP}$  [2], for  $\gamma = 100\sqrt{m}$ , which implies that  $\text{CVP}_\gamma$  cannot be NP-hard in this regime, unless  $\text{NP} \subset \text{coNP}$  and the polynomial hierarchy collapses.

## 6.2 The reduction from CVP to SLDP via lattice liftings

We now establish the key reduction from CVP to SLDP by relying on the lattice lifting result Theorem 1.5.

► **Theorem 6.7.** *There is a polynomial time (Turing) reduction from  $\text{CVP}_\gamma$  to  $\text{SLDP}_\gamma$  (with the same  $\gamma$ ). In more detail, given as input a CVP instance  $(t, \mathcal{L})$  with ambient dimension  $m$ , the reduction either solves CVP exactly or it computes a scaling factor  $s \in \mathbb{Z}_{>0}$  and an SLDP instance  $(\tilde{t}, U)$  of ambient dimension  $n$  where  $\text{dist}(t, \mathcal{L}) = s \text{dist}(\tilde{t} + U, \mathbb{Z}^n)$  and  $m \leq n \leq O(m^2 \log m)$ .*

**Proof of Theorem 6.7.** An instance of  $\text{CVP}_\gamma$  consists of a rank  $m$  lattice  $\mathcal{L}$  in  $\mathbb{R}^m$ , given by a generator matrix, and a target vector  $t \in \mathbb{Q}^m$ . We apply Theorem 1.5. Thus, for a given such instance, we can compute in polynomial time an orthonormal basis  $v_1, v_2, \dots, v_n \in \mathbb{Q}^n$ , where  $n \geq m$ , and a scaling factor  $s \in \mathbb{Z}_{>0}$ , such that the lattice  $\mathcal{L}'$  generated by  $v_1, v_2, \dots, v_n$  satisfies  $\mathcal{L} = sP(\mathcal{L}')$ , where  $P: \mathbb{R}^n \rightarrow \mathbb{R}^m$  denotes the orthogonal projection onto the first  $m$  coordinates. We view  $\mathbb{R}^m$  as the subspace of  $\mathbb{R}^n$  whose last  $n - m$  coordinates are zero.

Consider the orthogonal matrix  $Q \in \mathbb{Q}^{n \times n}$  with columns  $v_1, v_2, \dots, v_n$ , thus  $Qe_i = v_i$  if the  $e_i$  denote the standard basis vectors. This implies that  $Q(\mathbb{Z}^n) = \mathcal{L}'$ . Let  $w_i$  be the rows of  $Q$ , thus  $w_i = Q^T e_i$  and  $Qw_i = e_i$ . We define the subspace  $U$  as the span of  $w_{m+1}, \dots, w_n$ , thus  $Q(U) = \langle e_{m+1}, \dots, e_n \rangle = \ker P$ . Moreover, we define  $t' := s^{-1}Q^T t$ . Then  $Qt' = s^{-1}t$ . Summarizing,

$$sP(\mathcal{L}') = \mathcal{L}, \quad Q(\mathbb{Z}^n) = \mathcal{L}', \quad Q(U) = \ker P, \quad Qt' = s^{-1}t.$$

Let us verify that

$$\text{dist}(t, \mathcal{L}) = s \text{dist}(t' + U, \mathbb{Z}^n). \tag{6.1}$$

Indeed,  $\text{dist}(t' + U, \mathbb{Z}^n) = \text{dist}(Q(t' + U), Q(\mathbb{Z}^n))$  by the orthogonality of  $Q$ . Moreover,

$$\text{dist}(Q(t') + Q(U), Q(\mathbb{Z}^n)) = \text{dist}(s^{-1}t + \ker P, \mathcal{L}') = \text{dist}(s^{-1}P(t), P(\mathcal{L}')) = s^{-1} \text{dist}(t, \mathcal{L})$$

where we used Corollary 2.6 in the second and  $P(t) = t$  in the last equality.

Clearly, this defines a polynomial time reduction of  $\text{CVP}_\gamma$  to  $\text{SLDP}_\gamma$  that does not change the approximation factor  $\gamma$ .

Suppose the given instance of CVP has bit-length  $l$ . Theorem 1.5 implies that we can assume  $n = O(m(\log m + \log l))$ . We can bound  $l$  in terms of  $m$  by making a case

distinction. If  $l$  is so large that  $l \geq 2^{\Omega(m \log m)}$ , then we apply the algorithm in [53], which solves CVP exactly, in time  $l 2^{O(m \log m)}$  which is polynomially bounded in  $l$  in this case. Otherwise, if  $l \leq 2^{O(m \log m)}$ , we perform the above described reduction to SLDP. Then we have  $m(\log m + \log l) \leq O(m^2 \log m)$  and hence we can upper bound  $n = O(m^2 \log m)$ . This completes the proof.  $\blacktriangleleft$

**Proof of Theorem 6.1.** We compose the polynomial time reduction  $\text{CVP}_{2\gamma}$  to  $\text{SLDP}_{2\gamma}$  of Theorem 6.7 with the polynomial time reductions  $\text{SLDP}_{2\gamma}$  to  $\text{ROP}(G, \delta)_\gamma$  in Lemma 5.12. If  $m$  denotes the dimension of the given instance of CVP, then the dimension  $n$  of the constructed instance of ROP satisfies  $m \leq n \leq O(m^2 \log m)$ .

Using [29], we choose  $c$  such that  $\text{CVP}_{2\gamma}$  is NP-hard with the approximation factor  $2\gamma(m) = m^{c/\log \log m}$ . Since  $m \leq n \leq m^3$  for sufficiently large  $m$ , we have

$$\gamma(m) = \frac{1}{2} m^{c/\log \log m} \geq \frac{1}{2} n^{c/3 \log \log n} \geq n^{c'/\log \log n}$$

for a suitably chosen  $c'$ . Thus we see that  $\text{ROP}(G, \delta)_\gamma$  is NP-hard for the approximation factor  $n^{c'/\log \log n}$ , which completes the proof.  $\blacktriangleleft$

### 6.3 Proof of Theorem 1.5

Recall that  $L \in \mathbb{Z}^{m \times n}$  is called *right-invertible* over  $\mathbb{Z}$  if there exists  $R \in \mathbb{Z}^{n \times m}$  such that  $LR = I_m$ . This means that the lattice generated by the columns of  $L$  equals  $\mathbb{Z}^m$ , that is,  $L(\mathbb{Z}^n) = \mathbb{Z}^m$ .

► **Proposition 6.8.** *Suppose  $\mathcal{L} \subset \mathbb{R}^m$  is a lattice generated by the columns of  $G \in \mathbb{Z}^{m \times m}$  with  $\det G \neq 0$ . Then the following are equivalent for integers  $n \geq m$  and  $s \in \mathbb{Z}_{>0}$ :*

(1) *There exists a lattice  $\mathcal{L}' \subset \mathbb{R}^n$  generated by an orthonormal basis such that*

$$\mathcal{L} = sP(\mathcal{L}')$$

*where  $P : \mathbb{R}^n \rightarrow \mathbb{R}^m$  denotes the orthogonal projection onto the first  $m$ -coordinates.*

(2) *There exists a matrix  $L \in \mathbb{Z}^{m \times n}$  that is right-invertible over  $\mathbb{Z}$ , such that*

$$(GL)(GL)^T = s^2 I_m.$$

We need the following observation, which easily follows with Gram-Schmidt orthogonalization.

► **Lemma 6.9.** *Suppose  $m \leq n$  and  $X \in \mathbb{R}^{m \times n}$ . Then  $X$  can be completed to an orthogonal matrix if and only if  $XX^T = I_m$ .*

**Proof of Proposition 6.8.** (1)  $\Rightarrow$  (2): Suppose  $v_1, v_2, \dots, v_n \in \mathbb{R}^n$  is an orthonormal basis generating  $\mathcal{L}'$  and put  $u_i := P(v_i)$ . The matrix  $U \in \mathbb{R}^{m \times n}$  with columns  $u_1, \dots, u_n$  has orthogonal rows, that is,  $UU^T = I_m$ . By assumption, the vectors  $su_i$  are in  $\mathcal{L}$  and hence can be written as integer linear combinations of the columns of  $G$ . This means that there exists an integer matrix  $L \in \mathbb{Z}^{m \times n}$  such that  $GL = sU$ . The rows of this matrix are orthogonal, hence  $(GL)(GL)^T = s^2 I_m$ . It remains to show that  $L$  is right-invertible over  $\mathbb{Z}$ . Also by assumption, since the  $su_i$  generate  $\mathcal{L}$ , the columns of  $G$  can be written as integer linear combinations of  $u_i$ . Thus there exists an integer matrix  $R \in \mathbb{Z}^{n \times m}$  such that  $(GL)R = sUR = G$ . Hence  $LR = I_m$ , since  $\det G \neq 0$ , so that  $L$  is indeed right-invertible.

(2)  $\Rightarrow$  (1): Suppose  $s^2 I_m = (GL)(GL)^T$ . Thus the matrix  $X := \frac{1}{s} GL$  satisfies  $XX^T = I_m$ . By Lemma 6.9 we can complete  $X$  to an orthogonal matrix: thus there exists an orthogonal

matrix  $Y \in O_n$  whose first  $m$  rows constitute the matrix  $\frac{1}{s}GL$ . Therefore  $sPY = GL$ . Hence, if  $\mathcal{L}'$  denotes the lattice generated by the columns of  $Y$ , we have  $sP(\mathcal{L}') \subset \mathcal{L}$ . Equality holds since  $L$  is right-invertible over  $\mathbb{Z}$ .  $\blacktriangleleft$

Proposition 6.8 indicates a strategy for proving Theorem 1.5. We need two ingredients:

- A polynomial time algorithm for computing on input  $G$  an integer  $n \geq m$  and a right invertible matrix  $L \in \mathbb{Z}^{m \times n}$  such that  $(GL)(GL)^T = s^2 I_m$ .
- An efficient version of Lemma 6.9.

We prove now that both requirements can be satisfied. For the first step we will use the following result.

► **Theorem 6.10** (Efficient Waring decomposition for quadratic forms). *Assume  $A \in \mathbb{Q}^{m \times m}$  is a positive definite, symmetric matrix of bit-length  $b$ . Then, in  $\text{poly}(m, b)$ -time, we can compute  $N$  vectors  $l_1, l_2, \dots, l_N \in \mathbb{Q}^m$ , where  $m \leq N \leq O(m(\log m + \log b))$ , such that*

$$A = \sum_{i=1}^N l_i l_i^T.$$

► **Remark 6.11.** In 1932, Mordell [69] considered what he called the *Waring's problem for quadratic forms*, the problem of writing a given positive definite quadratic form  $f(x) := x^T A x$  as sum of squares of linear forms over  $\mathbb{Q}$ . Note that this equivalent to writing the matrix  $A$  as in Theorem 6.10. Mordell proved that a Waring decomposition with  $N = m + 3$  is always possible, and described a method for computing the linear forms. Unfortunately, his method relies on Lagrange's four squares theorem that every positive integer  $D$  can be written as the sum of four squares  $D = a^2 + b^2 + c^2 + d^2$ . Randomized polynomial time algorithms to compute  $a, b, c, d$  are available [77], but to the best of our knowledge, it is still an open problem whether a polynomial time deterministic algorithm exists. Instead, we will use the lemma below, which uses more squares but can easily be shown to run in deterministic polynomial time.

The proof of Theorem 6.10 requires some preparations.

► **Lemma 6.12.** *Given a positive integer  $D$ , we can in polynomial time compute  $k = O(\log \log D)$  integers  $a_1, a_2, \dots, a_k$  such that*

$$D = a_1^2 + a_2^2 + \dots + a_k^2.$$

**Proof.** We first compute  $a_1 := \lfloor \sqrt{D} \rfloor$ . Replacing  $D$  with  $D' := D - a_1^2$  and repeating the process, we compute integers  $a_1, a_2, \dots, a_k$  such that  $D = a_1^2 + \dots + a_k^2$ . Let us bound  $k$ : since  $a_1^2 \leq D < (a_1 + 1)^2$ , we have  $D' = D - a_1^2 \leq (a_1 + 1)^2 - 1 - a_1^2 \leq 2\sqrt{D}$ . This implies that if  $D_j$  is computed in the  $j$ -th step, then  $D_j \leq 2\sqrt{D_{j-1}} \leq \dots \leq 2^{1+\frac{1}{2}+\dots+\frac{1}{2^{j-1}}} D^{2^{-j}} \leq 4D^{2^{-j}}$ . We deduce that after  $k := \log_2 \log_2 D$  steps we have  $D_k \leq 8$ . The algorithm terminates after at most 4 more steps.  $\blacktriangleleft$

The following result can be found in [62, Algorithm 12.1].

► **Lemma 6.13** (Lagrange's method for congruence diagonalization). *Suppose  $X \in \mathbb{Z}^{m \times m}$  is a symmetric matrix. Then we can compute in polynomial time a matrix  $Q \in \mathbb{Z}^{m \times m}$  with  $\det Q \neq 0$  such that*

$$QXQ^T = \text{diag}(d_1, d_2, \dots, d_m) \in \mathbb{Z}^{m \times m}$$

*is diagonal with integer entries.*



**Proof of Theorem 6.10.** By multiplying  $A$  with the square of the least common multiple of the denominators of the entries of  $A$ , we may assume without loss of generality that  $A$  is integral. We compute a matrix  $Q$  such that  $QAQ^T = \text{diag}(d_1, d_2, \dots, d_m)$  by Lemma 6.13. Since  $A$  is positive definite, all  $d_i$  are positive. Using Lemma 6.12, for each  $i = 1, 2, \dots, m$ , we compute  $e_{ij}$  such that  $d_i = \sum_{j=1}^{k_i} e_{ij}^2$ . Denote by  $L$  the matrix

$$L := \begin{bmatrix} e_{11} & \dots & e_{1k_1} & & & \\ & & & e_{21} & \dots & e_{2k_2} \\ & & & & & \ddots \\ & & & & & & e_{m1} & \dots & e_{mk_m} \end{bmatrix} \in \mathbb{Z}^{m \times \sum k_i}.$$

Then  $LL^T = \text{diag}(d_1, d_2, \dots, d_m) = QAQ^T$ , so  $A = (Q^{-1}L)(Q^{-1}L)^T$ . Defining  $l_i \in \mathbb{Q}^m$  to be the  $i$ -th column of  $Q^{-1}L$ , we have  $A = \sum_{i=1}^N l_i l_i^T$  where  $N := \sum k_i$  and this proves the claim. For the bound on  $N$ , we first note that since the algorithm in Lemma 6.13 runs in polynomial time, we have  $\log d_i = (bm)^{O(1)}$ . Lemma 6.12 then implies that  $k_i = O(\log \log d_i) = O(\log(bm))$  and  $N = O(m \log(bm))$ .  $\blacktriangleleft$

The efficient Waring decomposition is the first ingredient in our proof of Theorem 1.5. The second ingredient is an efficient version of Lemma 6.9.

► **Lemma 6.14.** *Suppose  $m \leq n$  and  $X \in \mathbb{Q}^{m \times n}$  satisfies  $XX^T = I_m$ . Then, we can compute in polynomial time an orthogonal matrix  $Y \in \mathbb{Q}^{n \times n}$  such that the first  $m$  rows of  $Y$  are the rows of  $X$ .*

**Proof.** We construct  $Y$  as a product of reflections at rational hyperplanes.

First note that for any  $v, w \in \mathbb{Q}^n$  of norm one, there is an orthogonal matrix  $Y \in \mathbb{Q}^{n \times n}$  such that  $Yv = w$ . Indeed, w.l.o.g.  $v \neq -w$ , and take for  $Y$  the reflection at the hyperplane orthogonal to  $v + w$ , which is given by the linear map

$$Yx := 2 \frac{\langle x, v + w \rangle}{\|v + w\|^2} (v + w) - x.$$

Note that  $Y$  defines a rational orthogonal matrix and indeed  $Yv = w$ . Denote by  $e_i$  the standard basis vectors.

Given  $X$  as in the lemma, we compute an orthogonal matrix  $Y_1 \in \mathbb{Q}^{n \times n}$  which maps  $e_1$  to the first row of  $X$ . Since  $Y$  is orthogonal and  $Y^2 = I_n$ , the first row of  $Y$  equals the first row of  $X$ . We construct  $Y$  by iterating this process.  $\blacktriangleleft$

► **Remark 6.15.** M. Hall [41, 42] gave necessary conditions for an integer matrix to be completable to a scalar multiple of an orthogonal matrix in the sense of Lemma 6.14.

**Proof of Theorem 1.5.** Given  $G \in \mathbb{Z}^{m \times m}$  with  $\det G \neq 0$ , we compute  $s := m\|G\|_{\max} + 1$ . By Lemma 2.7 we have  $\sigma_{\max}(G) < s$ . The minimum eigenvalue of  $G^{-1}G^{-T}$  satisfies

$$\lambda_{\min}(G^{-1}G^{-T}) = \sigma_{\min}^2(G^{-1}) = \frac{1}{\sigma_{\max}^2(G)} > \frac{1}{s^2}.$$

Consequently, the matrix  $A := s^2(G^{-1}G^{-T}) - I_m$  is positive definite.

Using Theorem 6.10, we compute in polynomial time a number  $N$  and a matrix  $L \in \mathbb{Q}^{m \times N}$  such that  $A = LL^T$ . Let  $f \in \mathbb{Z}_{>0}$  denote the least common multiple of the denominators of the entries of  $L$ , so that  $L' := fL$  is integral. Then  $(sf)^2 G^{-1}G^{-T} - f^2 I_m = L'(L')^T$ .

Using Lemma 6.12, we compute integers  $b_1, \dots, b_p$  such that  $f^2 - 1 = b_1^2 + b_2^2 + \dots + b_p^2$ . Since  $L$  is computed from  $A$  in polynomial time, the bit-length of  $f$  is  $\text{poly}(b, m)$  so Lemma 6.12 implies  $p = O(\log(bm))$ . Considering the integer matrix  $L'' = [b_1 I_m \mid b_2 I_m \mid \dots \mid b_p I_m \mid L']$  of format  $m \times (pm + N)$ , we get

$$(sf)^2 G^{-1} G^{-T} - I_m = (b_1^2 + b_2^2 + \dots + b_p^2) I_m + L'(L')^T = L''(L'')^T. \quad (6.2)$$

We now consider the matrix  $X := [G \mid GL'']$ . By Equation (6.2), we have  $XX^T = (sf)^2 I_m$ . Using Lemma 6.14 with input  $(sf)^{-1}X$ , we compute a rational, orthogonal matrix  $Y \in \mathbb{Q}^{n \times n}$  such that the first  $m$  rows of  $Y$  are the rows of  $(sf)^{-1}X$  where  $n := m + (pm + N)$ .

The columns  $v_1, \dots, v_n$  of  $Y$  form an orthonormal basis. Moreover, the orthogonal projections  $P(v_i)$  onto the first  $m$  coordinates equal the columns of  $(sf)^{-1}X$ . Thus if we denote by  $\mathcal{L}'$  the lattice generated by  $v_1, \dots, v_n$ , then  $sfP(\mathcal{L}')$  is the lattice spanned by the columns of  $X = [G \mid GL'']$ , which is equal to  $\mathcal{L} = G(\mathbb{Z}^m)$ , because  $L''$  is an integer matrix. We conclude that  $sfP(\mathcal{L}') = \mathcal{L}$ .  $\blacktriangleleft$

## 7 Orbit Problems and the Kempf-Ness Approach

In Section 1.9 we outlined a general approach for the Orbit Equality Problem, based on the Kempf-Ness theorem [55]. Here we analyze this approach for torus actions. We formulate a conjecture on the complexity of computing approximate solutions to unconstrained geometric programs. A positive answer to the conjecture leads to a numerical algorithm for the Orbit Equality Problem that runs in polynomial time, provided Separation Hypothesis 1.7 is true.

### 7.1 The general Kempf-Ness theorem

This theorem holds for any reductive group  $T$  (not just a torus) with a rational action on a finite dimensional vector space  $V$ . We denote by  $K$  a maximal compact subgroup of  $T$  and assume that  $V$  is endowed with a Hermitian inner product such that  $K$  acts by isometries. The inner product defines a norm and Euclidean metric on  $V$ . In the special case  $T = (\mathbb{C}^\times)^d$  of a torus we have  $K = (S^1)^d$ .

It is well known [24, Thm. 2.3.6] that each orbit closure  $\overline{\mathcal{O}_v}$  contains a unique closed orbit. Therefore, we have  $\overline{\mathcal{O}_v} \cap \overline{\mathcal{O}_w} \neq \emptyset$  iff the orbit closures  $\overline{\mathcal{O}_v}$  and  $\overline{\mathcal{O}_w}$  share the same closed orbit. We therefore focus on closed orbits: let us call a vector  $v \in V$  *polystable* if its orbit  $\mathcal{O}_v$  is closed. We denote by  $V^{ps}$  the set of polystable vectors in  $V$ . As in Section 2.5, we can endow the space  $V^{ps}/T$  of closed orbits with the quotient topology, which is Hausdorff. The Kempf-Ness Theorem [55] states that the space  $V^{ps}/T$  is homeomorphic to a “smaller object”, which is defined in analytic terms, namely the space  $\text{Crit}(V)/K$  of  $K$ -orbits of the closed and  $K$ -invariant subset  $\text{Crit}(V)$  of critical points of  $V$ .

The critical points are defined in terms of the *Kempf-Ness function*:

$$F_v : T \rightarrow \mathbb{R}, \quad F_v(g) := \log \|g \cdot v\| = \frac{1}{2} \log \|g \cdot v\|^2. \quad (7.1)$$

Let us denote the induced action of  $\text{Lie}(T)$  on  $V$  by  $x \cdot v := \left. \frac{d}{dt} \right|_{t=0} (e^{tx} \cdot v)$ , for  $x \in \text{Lie}(T)$  and  $v \in V$ . Note that  $x \cdot v = v + M^T x$  for the action Equation (1.1) of a torus. One checks that the derivative of  $F_v$  at  $I$  is given by  $D_I F_v x = \|v\|^{-2} \langle x \cdot v, v \rangle$ . Thus we define  $v \in V$  to be *critical* iff  $\langle x \cdot v, v \rangle = 0$  for all  $x \in \text{Lie}(T)$ . By definition, the set  $\text{Crit}(V)$  of critical vectors in  $V$  is a closed real algebraic set in  $V$ , cut out by real quadratic polynomials. Moreover,  $\text{Crit}(V)$  is  $K$ -invariant, since we assume the Hermitian inner product to be  $K$ -invariant.

It is easy to see that closed orbits contain a critical point. The celebrated Kempf-Ness theorem [55] states the converse: the orbits of critical points are closed, thus  $\text{Crit}(V) \subset V^{ps}$ . Even more is known:

- **Theorem 7.1** (The Kempf-Ness theorem). (a) *The orbit  $\mathcal{O}_v$  is closed if and only if  $\mathcal{O}_v \cap \text{Crit}(V) \neq \emptyset$ .*  
 (b) *The intersection  $\mathcal{C}_{v^*} := \overline{\mathcal{O}_v} \cap \text{Crit}(V)$  is a single  $K$ -orbit, that we call the Kempf-Ness orbit of  $v$ .*  
 (c) *We have  $\overline{\mathcal{O}_v} \cap \overline{\mathcal{O}_w} \neq \emptyset$  if and only if  $\mathcal{C}_{v^*} = \mathcal{C}_{w^*}$ .*

The proof of Theorem 7.1 relies on convexity properties of the function  $F_v$ . One observes that  $t \mapsto F_v(\gamma(t))$  is convex for all curves of the form  $\gamma(t) := \text{Exp}(tx) \cdot v$ . A more conceptual view of the situation is as follows (see [17]). One can view  $F_v$  as a function on  $T/K = \{Kg \mid g \in T\}$ , since  $F_v$  is constant on the cosets  $Kg$ . Moreover,  $T/K$  is a symmetric space with respect to a  $K$ -invariant Riemannian metric, whose geodesics are exactly the curves arising from the  $\gamma$ . For this reason, one calls  $F_v$  a *geodesically convex* function on  $T/K$ . In the special case where  $T = (\mathbb{C}^\times)^d$  is the torus, we have the isometry

$$\mathbb{R}^d \xrightarrow{\sim} T/K, \quad x \mapsto K \text{Exp}(x/2) \quad (7.2)$$

given by the exponential map. In general, the convexity implies that  $v^* \in \overline{\mathcal{O}_v}$  is critical iff

$$\|v^*\| = \inf_{v' \in \mathcal{O}_v} \|v'\|.$$

► **Remark 7.2.** The second property in Theorem 7.1 expresses that the continuous map  $\text{Crit}(V)/K \rightarrow V^{ps}/T$ , induced by the inclusion, is a bijection. Equivalently, this map is a homeomorphism.<sup>8</sup>

## 7.2 Efficient approximation of the Kempf-Ness orbit

We return to the situation of the action of a torus  $T$  on  $V = \mathbb{C}^n$ . Let  $\omega_1, \dots, \omega_n \in \mathbb{Z}^d$  denote the corresponding weights, i.e., the columns of the weight matrix  $M$ . We will always assume that the affine span of these weights is  $\mathbb{R}^d$ . We fix a vector  $v \in V$  and assume for simplicity that  $q_i := |v_i|^2 > 0$  for all  $i$ . Using the parametrization (7.2) and multiplying by 2 to simplify the presentation, the Kempf-Ness function (7.1) becomes

$$f: \mathbb{R}^d \rightarrow \mathbb{R}, \quad f(x) := 2 \log \|e^{x/2} \cdot v\| = \log \left( \sum_{i=1}^n q_i e^{\omega_i^T x} \right). \quad (7.3)$$

The gradient  $\nabla f(x)$  and the Hessian of  $f$  at  $x$  are given by

$$\nabla f(x) = \frac{\sum_{i=1}^n q_i e^{\omega_i^T x} \omega_i}{\sum_{i=1}^n q_i e^{\omega_i^T x}}, \quad \nabla^2 f(x) = \frac{\sum_{i=1}^n q_i e^{\omega_i^T x} \omega_i \omega_i^T}{\sum_{i=1}^n q_i e^{\omega_i^T x}} - \nabla f(x) \nabla f(x)^T. \quad (7.4)$$

This shows that  $\nabla f(x)$  lies in the interior of the *weight polytope*

$$P := \text{conv}(\omega_1, \dots, \omega_n).$$

<sup>8</sup> Pass to projective spaces to see this.

Moreover, the Cauchy-Schwarz inequality implies for all nonzero  $y \in \mathbb{R}^d$ ,

$$y^T \nabla^2 f(x) y = \frac{\sum_{i=1}^n q_i e^{\omega_i^T x} (y^T \omega_i)^2}{\sum_{i=1}^n q_i e^{\omega_i^T x}} - (y^T \nabla f(x))^2 > 0, \quad (7.5)$$

which directly proves that  $f$  is a strictly convex function.<sup>9</sup>

For the following equivalences, see [20, 16]. The function  $f$  has a finite infimum iff  $0 \in P$ . This is equivalent to  $0 \notin \overline{\mathcal{O}_v}$ . Moreover, the infimum  $f_\star$  is attained iff  $0$  lies in the interior of  $P$ , which is equivalent to the orbit  $\mathcal{O}_v$  being closed. By the strict convexity, the minimum is attained at a unique point  $x^\star \in \mathbb{R}^d$ . By Theorem 7.1,  $v^\star := e^{x^\star} \cdot v$  defines the  $K$ -orbit  $\mathcal{C}_{v^\star}$ .

There is a vast amount of literature on the problem of minimizing  $f(x)$  using the ellipsoid or the interior point methods, see [20, 14, 74, 79, 82, 60] and the references therein. On input  $\varepsilon > 0$ , these algorithms return a point  $x \in \mathbb{R}^d$  with  $f(x) - f_\star < \varepsilon$  in polynomial time. However, surprisingly, we are not aware of any result on computing a certified approximation to  $x^\star$  in polynomial time, and currently this seems to be an open problem. We conjecture that this can be achieved in polynomial time:

► **Conjecture 7.3.** *There exists an algorithm that on input  $M \in \mathbb{Z}^{d \times n}$  such that  $0 \in \text{int}(P)$ , a vector  $v \in (\mathbb{Q}(i)^\times)^n$ , and  $\varepsilon \in \mathbb{Q}_{>0}$ , computes  $x \in \mathbb{Q}^d$  such that*

$$\|x - x^\star\| < \varepsilon,$$

*in time  $\text{poly}(d, n, B, b, \log \frac{1}{\varepsilon})$ , where  $B$  is the bit-length of  $M$  and  $b$  is the bit-length of  $v$ .*

We will now explain why computing an approximation to  $x^\star$  seems to be more challenging than approximately minimizing  $f(x)$ . In Example 7.4 below, we provide a family of Kempf-Ness functions  $f$  and points  $x$  such that  $f(x) - f_\star$  becomes doubly-exponentially small in the input bit-length, while the distance to the true minimizer  $x^\star$  remains constant:  $\|x - x^\star\| = 1$ . This implies that using known minimization algorithms in a black box will not be sufficient to solve Conjecture 7.3 and more specific algorithms appear to be required.

► **Example 7.4.** We consider the action of  $T := (\mathbb{C}^\times)^2$  on  $V := \mathbb{C}^4$  with the following weights:  $\omega_1 := (1, 0)$ ,  $\omega_2 := (-2, 0)$ ,  $\omega_3 := (-N, 1)$ ,  $\omega_4 := (-N, -1)$  where  $N > 2$  is a positive integer. See Figure 4 for illustration. The Kempf-Ness function (7.3) for  $v := (1, 1, 1, 1)$  reads

$$f(x_1, x_2) = \log \left( e^{x_1} + e^{-2x_1} + e^{-Nx_1+x_2} + e^{-Nx_1-x_2} \right).$$

We note that  $f(x) \geq 0$  for every  $x \in \mathbb{R}^d$  since at least one of the exponents  $x_1, -2x_1$  is non-negative.

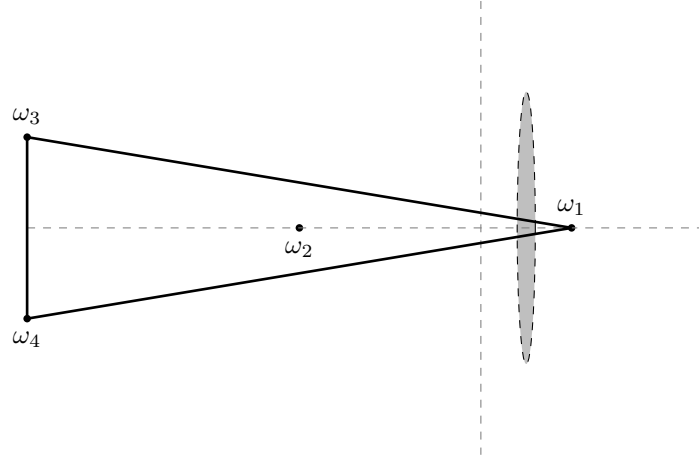
We now focus on the unique minimizer  $x^\star = (x_1^\star, x_2^\star)$  of the convex program  $f_\star := \min_x f(x)$ , which is characterized by the property that the gradient vanishes:  $\nabla f(x^\star) = 0$ . By Equation (7.4) this is equivalent to

$$e^{x_1^\star} - 2e^{-2x_1^\star} - Ne^{-Nx_1^\star+x_2^\star} - Ne^{-Nx_1^\star-x_2^\star} = 0 \quad \text{and} \quad e^{-Nx_1^\star+x_2^\star} - e^{-Nx_1^\star-x_2^\star} = 0. \quad (7.6)$$

The second equation implies  $x_2^\star = 0$ . The first equation then reduces to

$$e^{x_1^\star} - 2e^{-2x_1^\star} - 2Ne^{-Nx_1^\star} = 0, \quad (7.7)$$

<sup>9</sup> To see this, we split  $\sum_i q_i e^{\omega_i^T x} y^T \omega_i$  as  $\sum_i q_i^{\frac{1}{2}} e^{\frac{1}{2} \omega_i^T x} y^T \omega_i \cdot q_i^{\frac{1}{2}} e^{\frac{1}{2} \omega_i^T x}$  and apply the Cauchy-Schwarz inequality. Since the vectors  $\omega_i$  affinely span  $\mathbb{R}^d$  we have  $y^T \omega_i \neq y^T \omega_j$  for some  $i \neq j$ . This shows that the vectors  $(q_i^{\frac{1}{2}} e^{\frac{1}{2} \omega_i^T x} y^T \omega_i \mid i \in [n])$  and  $(q_i^{\frac{1}{2}} e^{\frac{1}{2} \omega_i^T x} \mid i \in [n])$  cannot be proportional and hence the Cauchy-Schwarz inequality must be strict.



■ **Figure 5** An example of a torus action with weights  $(1, 0), (-2, 0), (-N, 1), (-N, -1)$ . The points in the shaded region all have small  $f(x) - f_*$  but they can be far away from the optimal solution  $x^*$ .

which implies that  $x_1^*$  is positive and decreases with  $N$ . In the limit  $N \rightarrow \infty$ , the first equation degenerates to  $e^{x_1^*} - 2e^{-2x_1^*} = 0$ . We deduce that  $e^{x_1^*} > \sqrt[3]{2}$  for every  $N > 2$ . In fact, we have:

$$\sqrt[3]{2} < e^{x_1^*} < \left(1 + \frac{N}{2^{N/3}}\right) \sqrt[3]{2}. \quad (7.8)$$

To see this, we write  $e^{x_1^*} = (1 + \varepsilon) \sqrt[3]{2}$  where  $\varepsilon > 0$ . Then Equation (7.7) implies

$$(1 + \varepsilon) \sqrt[3]{2} - \frac{\sqrt[3]{2}}{(1 + \varepsilon)^2} = \frac{2N}{(1 + \varepsilon)^2 2^{N/3}}.$$

Multiplying this equation by  $(1 + \varepsilon)^N / \sqrt[3]{2}$ , we get  $(1 + \varepsilon)^{N-2} (\varepsilon^2 + 3\varepsilon + 3) \varepsilon = 2N \times 2^{-(N+1)/3}$ . Since  $\varepsilon > 0$ , we have  $\varepsilon^2 + 3\varepsilon + 3 > 3$  and  $3\varepsilon < 2N \times 2^{-(N+1)/3}$ , which implies Equation (7.8).

We now consider the point  $x := (x_1^*, 1)$ , whose first coordinate agrees with the first coordinate of  $x^* = (x_1^*, 0)$  but has a constant distance away from  $x^*$  in the second coordinate. We will show that  $f(x) - f_*$  is exponentially small in  $N$ :

$$e^{f(x)-f_*} = 1 + \frac{e^{-Nx_1^*}(e + e^{-1} - 2)}{e^{f_*}} \leq 1 + e^{-Nx_1^*}(e + e^{-1} - 2)$$

where the inequality holds since  $f_* \geq 0$ . Taking logarithms we get

$$f(x) - f_* \leq \log(1 + e^{-Nx_1^*}(e + e^{-1} - 2)) \leq e^{-Nx_1^*}(e + e^{-1} - 2) < 2^{-N/3}(e + e^{-1} - 2)$$

since  $e^{x_1^*} > \sqrt[3]{2}$ . This shows that  $f(x) - f_*$  is exponentially small in  $N$  and hence doubly-exponentially small in the input bit-length, and yet  $\|x - x^*\| = 1$ .

► **Remark 7.5.** In the setting of Example 7.4, consider the vectors  $v := (1, 1, 1, 1)$  and  $w := (1, 1, 2, 2)$ , and denote by  $x^*, y^*$  the minimizers of the Kempf-Ness functions of  $v, w$ , respectively, i.e.,  $e^{x^*} = v^*$  and  $e^{y^*} = w^*$ . By Example 7.4, we have  $x^* = (x_1^*, 0)$ , where  $x_1^*$  satisfies Equation (7.8). By a similar reasoning, we also have  $y^* = (y_1^*, 0)$ , where  $y_1^*$  satisfies

$$\sqrt[3]{2} < e^{y_1^*} < \left(1 + \frac{2N}{2^{N/3}}\right) \sqrt[3]{2}.$$

We will now show that the coordinates of  $v^*, w^*$  are exponentially close in  $N$ . To this end, denote  $\varepsilon_N := 2N \times 2^{-N/3}$ . We have  $|(v^*)_1 - (w^*)_1| = |e^{x_1^*} - e^{y_1^*}| < 2^{1/3} \varepsilon_N$ . For the second coordinate we have  $|(v^*)_2 - (w^*)_2| = |e^{-2x_1^*} - e^{-2y_1^*}| < 2^{-2/3}((1 + \varepsilon_N)^2 - 1) < 2^{4/3} \varepsilon_N$ , using the crude bound  $(1 + \varepsilon_N)^2 - 1 < 4\varepsilon_N$ , which holds for  $\varepsilon_N < 2$  and hence for  $N$  large enough. For the third and fourth coordinate we have  $|(v^*)_3 - (w^*)_3| = |(v^*)_4 - (w^*)_4| = |e^{-Nx_1^*} - 2e^{-Ny_1^*}| < |e^{-Nx_1^*}| + 2|e^{-Ny_1^*}| < 3 \times 2^{-N/3} < \varepsilon_N$ . Hence  $\|v^* - w^*\|_\infty < 2^{4/3} \varepsilon_N$  and we conclude

$$\text{dist}(\mathcal{C}_{v^*}, \mathcal{C}_{w^*}) \leq \|v^* - w^*\| \leq 2\|v^* - w^*\|_\infty < 2^{7/3} \varepsilon_N$$

where the second inequality follows from the inequality  $\|\cdot\|_2 \leq \sqrt{n} \|\cdot\|_\infty$  which holds in the  $n$ -dimensional Euclidean space. This shows that the Euclidean distance between  $\mathcal{C}_{v^*}$  and  $\mathcal{C}_{w^*}$  can be exponentially small in  $N$  and doubly exponentially small in the bit-length of  $N$ . On the other hand, it is easy to lower bound the logarithmic distance: It is easy to verify that

$$H = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 2N & 0 & 1 & 1 \end{bmatrix}$$

is the matrix of rational invariants and  $H \text{Log } v = (0, 0)$  and  $H \text{Log } w = (0, 2 \log 2)$ . Proposition 3.5 then implies

$$\delta_{\log}(\mathcal{C}_{v^*}, \mathcal{C}_{w^*}) \geq \delta_{\log}(\mathcal{O}_v, \mathcal{O}_w) \geq \frac{1}{\sigma_{\max}(H)} \Delta(H \text{Log } v, H \text{Log } w) \geq \frac{\log 2}{2N}$$

where in the last inequality we use Lemma 2.7 to upper bound  $\sigma_{\max}(H) \leq 4N$ .

More generally, if Separation Hypothesis 1.7 holds then it is true for general torus actions that  $\delta_{\log}(\mathcal{C}_{v^*}, \mathcal{C}_{w^*})$  is at most singly exponentially small: If  $\mathcal{C}_{v^*} \neq \mathcal{C}_{w^*}$ , then, in general,  $\mathcal{O}_v \neq \mathcal{O}_w$ , hence,  $\delta_{\log}(\mathcal{C}_{v^*}, \mathcal{C}_{w^*}) \geq \delta_{\log}(\mathcal{O}_v, \mathcal{O}_w) \geq \text{sep}_T(d, n, B, b)$ . This is one of the main reasons why we work with the logarithmic distance instead of the Euclidean distance.

### 7.3 Deciding the equality of the Kempf-Ness orbits

We now show that, assuming Separation Hypothesis 1.7 and Conjecture 7.3, that for given  $v, w \in (\mathbb{C}^\times)^n$ , it is possible to decide in polynomial time whether the orbits  $\mathcal{O}_v$  for  $\mathcal{O}_w$  are equal. By Theorem 7.1, this is equivalent to the equality of the Kempf-Ness orbits  $\mathcal{C}_{v^*}$  and  $\mathcal{C}_{w^*}$ . We decide this by computing the approximate distance of the corresponding orbits  $\mathcal{C}_{v^*}$  and  $\mathcal{C}_{w^*}$  with sufficient accuracy. This has two ingredients: first we compute  $x'$  and  $y'$  that are  $\varepsilon$ -close approximations of  $x^*$  and  $y^*$ , respectively. This is possible by Conjecture 7.3. In a second step, we compute an approximation  $D$  of the distance  $\delta := \delta_{\log}(\mathcal{C}_{v'}, \mathcal{C}_{w'})$  between  $\mathcal{C}_{v'}$  and  $\mathcal{C}_{w'}$  such that  $\delta \leq D \leq \gamma\delta$ . For this we use Theorem 5.2 (c), see Corollary 7.6 below. When choosing  $\varepsilon = 2^{-\ell}$  sufficiently small, we show that  $\mathcal{C}_{v^*} = \mathcal{C}_{w^*}$  iff  $D < \gamma\varepsilon$ , which allows to decide orbit equality. Separation Hypothesis 1.7 is needed to make sure that it is sufficient to choose  $\ell$  as a polynomial in the input bit-length in order to obtain a polynomial time algorithm.

We now give the technical details. We recall that we restrict ourselves to torus actions where 0 lies in the interior of the convex hull  $P$  of the set of weights  $\Omega$ . This condition guarantees that *all* orbits of vectors in  $(\mathbb{C}^\times)^n$  are closed in  $\mathbb{C}^n$ ; on the other hand, when  $0 \notin \text{int}(P)$ , then no such orbit is closed, see [16].

We first state the following straightforward consequence of Theorem 5.2 (c).

► **Corollary 7.6.** *We assume the Separation Hypothesis 1.7. There is an exponentially bounded approximation factor  $\gamma = \gamma(d, n, B, b)$  such that, given a  $T$ -action by a weight matrix*

$M \in \mathbb{Z}^{d \times n}$ , and given vectors  $v, w \in (\mathbb{Q}(i)^\times)^n$  and  $x, y \in \mathbb{Q}^d$ , we can compute in polynomial time a number  $D \in \mathbb{Q}_{\geq 0}$  such that

$$\delta_{\log}(\mathcal{C}_{v'}, \mathcal{C}_{w'}) \leq D \leq \gamma \delta_{\log}(\mathcal{C}_{v'}, \mathcal{C}_{w'}),$$

where  $v' := e^x \cdot v$  and  $w' := e^y \cdot w$ .

Using the approximation factor  $\gamma$  of Corollary 7.6, we define the parameter

$$\varepsilon := \frac{1}{2\gamma} \text{sep}_T(d, n, B, b).$$

Assuming Conjecture 7.3, we can compute in polynomial time approximations

$$\|x - x^*\| < \frac{1}{2} \|M\|^{-1} \varepsilon, \quad \|y - y^*\| < \frac{1}{2} \|M\|^{-1} \varepsilon.$$

Recall that  $v^* = e^{x^*} \cdot v$ . We set  $v' := e^x \cdot v$ . Then  $\text{Log } v^* = \text{Log } v + M^T x^*$  and  $\text{Log } v' = \text{Log } v + M^T x$ , hence

$$\delta_{\log}(v', v^*) = \Delta(\text{Log } v + M^T x, \text{Log } v + M^T x^*) = \|M^T(x - x^*)\| \leq \frac{1}{2} \varepsilon,$$

which implies  $\delta_{\log}(\mathcal{C}_{v'}, \mathcal{C}_{v^*}) \leq \frac{1}{2} \varepsilon$ . Analogously, we get  $\delta_{\log}(\mathcal{C}_{w'}, \mathcal{C}_{w^*}) \leq \frac{1}{2} \varepsilon$  for  $w' := e^y \cdot w$ . Using the triangle inequality, we obtain

$$\delta_{\log}(\mathcal{C}_{v^*}, \mathcal{C}_{w^*}) \leq \delta_{\log}(\mathcal{C}_{v^*}, \mathcal{C}_{v'}) + \delta_{\log}(\mathcal{C}_{v'}, \mathcal{C}_{w'}) + \delta_{\log}(\mathcal{C}_{w'}, \mathcal{C}_{w^*}) \leq \delta_{\log}(\mathcal{C}_{v'}, \mathcal{C}_{w'}) + \varepsilon.$$

By symmetry, this implies

$$|\delta_{\log}(\mathcal{C}_{v^*}, \mathcal{C}_{w^*}) - \delta_{\log}(\mathcal{C}_{v'}, \mathcal{C}_{w'})| < \varepsilon. \quad (7.9)$$

We now observe the following two implications:

$$\begin{aligned} \delta_{\log}(\mathcal{C}_{v^*}, \mathcal{C}_{w^*}) = 0 &\implies \delta_{\log}(\mathcal{C}_{v'}, \mathcal{C}_{w'}) < \varepsilon \\ \delta_{\log}(\mathcal{C}_{v^*}, \mathcal{C}_{w^*}) > 0 &\implies \delta_{\log}(\mathcal{C}_{v'}, \mathcal{C}_{w'}) > \gamma \varepsilon. \end{aligned} \quad (7.10)$$

The first implication is an obvious consequence of Equation (7.9). For the second implication, note that  $\mathcal{C}_{v^*} \neq \mathcal{C}_{w^*}$  implies  $\delta_{\log}(\mathcal{C}_{v^*}, \mathcal{C}_{w^*}) \geq \delta_{\log}(\mathcal{O}_v, \mathcal{O}_w) \geq \text{sep}_T$ , since  $\mathcal{C}_{v^*} \subset \mathcal{O}_v$  and  $\mathcal{C}_{w^*} \subset \mathcal{O}_w$ . Combined with Equation (7.9), this gives

$$\delta_{\log}(\mathcal{C}_{v'}, \mathcal{C}_{w'}) > \text{sep}_T - \varepsilon = 2\gamma\varepsilon - \varepsilon \geq \gamma\varepsilon,$$

which proves Equation (7.10).

We use Corollary 7.6 to compute in polynomial time the number  $D$  such that  $\delta \leq D \leq \gamma\delta$ , where  $\delta := \delta_{\log}(\mathcal{C}_{v'}, \mathcal{C}_{w'})$ . Equation (7.10) implies now:

$$\begin{aligned} \mathcal{C}_{v^*} = \mathcal{C}_{w^*} &\implies D < \gamma\varepsilon \\ \mathcal{C}_{v^*} \neq \mathcal{C}_{w^*} &\implies D > \gamma\varepsilon. \end{aligned}$$

So the knowledge of  $D$  allows to decide whether  $\mathcal{C}_{v^*} = \mathcal{C}_{w^*}$  and hence whether  $\mathcal{O}_v = \mathcal{O}_w$ .

---

## References

- 1 Manindra Agrawal, Neeraj Kayal, and Nitin Saxena. PRIMES is in P. *Annals of Mathematics*, 160(2):781–793, 2004.



- 2 Dorit Aharonov and Oded Regev. Lattice problems in  $\text{NP} \cap \text{coNP}$ . *J. ACM*, 52(5):749–765, sep 2005. doi:10.1145/1089023.1089025.
- 3 Zeyuan Allen-Zhu, Ankit Garg, Yuanzhi Li, Rafael Oliveira, and Avi Wigderson. Operator scaling via geodesically convex optimization, invariant theory and polynomial identity testing. In *STOC'18—Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 172–181. ACM, New York, 2018. doi:10.1145/3188745.3188942.
- 4 Carlos Améndola, Kathlén Kohn, Philipp Reichenbach, and Anna Seigal. Invariant theory and scaling algorithms for maximum likelihood estimation. *SIAM J. Appl. Algebra Geom.*, 5(2):304–337, 2021. doi:10.1137/20M1328932.
- 5 Carlos Améndola, Kathlén Kohn, Philipp Reichenbach, and Anna Seigal. Toric invariant theory for maximum likelihood estimation in log-linear models. *Algebraic Statistics*, 12(2):187–211, 2021.
- 6 Michele Audin. *Torus actions on symplectic manifolds*, volume 93 of *Progress in Mathematics*. Birkhäuser Basel, 2012.
- 7 Alan Baker. Experiments on the abc-conjecture. *Publicationes Mathematicae*, 65, 11 2004.
- 8 Alan Baker. Logarithmic forms and the abc-conjecture. In *Number Theory: Diophantine, Computational and Algebraic Aspects. Proceedings of the International Conference held in Eger, Hungary, July 29-August 2, 1996*, pages 37–44. De Gruyter, 2011. URL: <https://doi.org/10.1515/9783110809794.37>, doi:doi:10.1515/9783110809794.37.
- 9 Alan Baker and Gisbert Wüstholz. *Logarithmic Forms and Diophantine Geometry*. New Mathematical Monographs. Cambridge University Press, 2008. doi:10.1017/CB09780511542862.
- 10 Markus Bläser, Christian Ikenmeyer, Vladimir Lysikov, Anurag Pandey, and Frank-Olaf Schreyer. Variety membership testing, algebraic natural proofs, and geometric complexity theory. arXiv:1911.02534, 2020.
- 11 Enrico Bombieri and Walter Gubler. *Heights in Diophantine Geometry*. New Mathematical Monographs. Cambridge University Press, 2006. doi:10.1017/CB09780511542879.
- 12 Jonathan M. Borwein and Peter B. Borwein. *Pi and the AGM : a study in analytic number theory and computational complexity*. Canadian Mathematical Society series of monographs and advanced texts. Wiley, New York, 1987.
- 13 Jonathan M. Borwein and Peter B. Borwein. On the complexity of familiar functions and numbers. *SIAM Rev.*, 30(4):589–601, 1988. doi:10.1137/1030134.
- 14 Stephen Boyd, Seung-Jean Kim, Lieven Vandenbergh, and Arash Hassibi. A tutorial on geometric programming. *Optimization and Engineering*, 8(1):67–127, March 2007. doi:10.1007/s11081-007-9001-7.
- 15 Peter Bürgisser, Matthias Christandl, Ketan Mulmuley, and Michael Walter. Membership in moment polytopes is in NP and coNP. *SIAM J. Comput.*, 46(3):972–991, 2017. doi:10.1137/15M1048859.
- 16 Peter Bürgisser, Levent Doğan, Visu Makam, Michael Walter, and Avi Wigderson. Polynomial Time Algorithms in Invariant Theory for Torus Actions. In Valentine Kabanets, editor, *36th Computational Complexity Conference (CCC 2021)*, volume 200 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 32:1–32:30, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. URL: <https://drops.dagstuhl.de/opus/volltexte/2021/14306>, doi:10.4230/LIPIcs.CCC.2021.32.
- 17 Peter Bürgisser, Cole Franks, Ankit Garg, Rafael Mendes de Oliveira, Michael Walter, and Avi Wigderson. Towards a theory of non-commutative optimization: geodesic first and second order methods for moment maps and polytopes. In *60th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2019*, pages 845–861. IEEE Computer Soc., Los Alamitos, CA, 2019. URL: <http://arxiv.org/abs/1910.12375>.
- 18 Peter Bürgisser, Cole Franks, Ankit Garg, Rafael Oliveira, Michael Walter, and Avi Wigderson. Efficient algorithms for tensor scaling, quantum marginals, and moment polytopes. In *59th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2018*, pages 883–897. IEEE Computer Soc., Los Alamitos, CA, 2018. doi:10.1109/FOCS.2018.00088.

- 19 Peter Bürgisser, Ankit Garg, Rafael Oliveira, Michael Walter, and Avi Wigderson. Alternating minimization, scaling algorithms, and the null-cone problem from invariant theory. In *9th Innovations in Theoretical Computer Science*, volume 94 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 24, 20. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2018.
- 20 Peter Bürgisser, Yinan Li, Harold Nieuwboer, and Michael Walter. Interior-point methods for unconstrained geometric programming and scaling problems. *arXiv:2008.12110*, 2020.
- 21 Giuseppe C. Calafiore, Stephane Gaubert, and Corrado Possieri. Log-sum-exp neural networks and posynomial models for convex and log-log-convex data. *IEEE Transactions on Neural Networks and Learning Systems*, 31(3):827–838, 2020. doi:10.1109/TNNLS.2019.2910417.
- 22 Giuseppe C. Calafiore, Stephane Gaubert, and Corrado Possieri. A universal approximation result for difference of log-sum-exp neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 31(12):5603–5612, 2020. doi:10.1109/TNNLS.2020.2975051.
- 23 John H. Conway and Neil J. A. Sloane. Low-dimensional lattices v. integral coordinates for integral lattices. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 426(1871):211–232, 1989. URL: <http://www.jstor.org/stable/2398341>.
- 24 Harm Derksen and Gregor Kemper. *Computational Invariant Theory*. BV035421342 Encyclopaedia of Mathematical Sciences volume 130. Springer, Heidelberg ; New York ; Dordrecht ; London, second enlarged edition with two appendices by vladimir l. popov, and an addendum by norbert a. campo and vladimir l. popov edition, 2015.
- 25 Harm Derksen and Visu Makam. Polynomial degree bounds for matrix semi-invariants. *Adv. Math.*, 310:44–63, 2017. doi:10.1016/j.aim.2017.01.018.
- 26 Harm Derksen and Visu Makam. Algorithms for orbit closure separation for invariants and semi-invariants of matrices. *Algebra Number Theory*, 14(10):2791–2813, 2020. doi:10.2140/ant.2020.14.2791.
- 27 Harm Derksen and Visu Makam. Maximum likelihood estimation for matrix normal models via quiver representations. *SIAM Journal on Applied Algebra and Geometry*, 5(2):338–365, 2021.
- 28 Harm Derksen, Visu Makam, and Michael Walter. Maximum likelihood estimation for tensor normal models via castling transforms. In *Forum of Mathematics, Sigma*, volume 10, page e50. Cambridge University Press, 2022.
- 29 Irit Dinur, Guy Kindler, Ran Raz, and Shmuel Safra. Approximating CVP to within-almost polynomial factors is NP-hard. *Combinatorica*, 23(2):205–243, April 2003. doi:10.1007/s00493-003-0019-y.
- 30 Kousha Etessami, Alistair Stewart, and Mihalis Yannakakis. A note on the complexity of comparing succinctly represented integers, with an application to maximum probability parsing. *ACM Trans. Comput. Theory*, 6(2), may 2014. doi:10.1145/2601327.
- 31 Michael A. Forbes and Amir Shpilka. Explicit Noether normalization for simultaneous conjugation via polynomial identity testing. In *Approximation, randomization, and combinatorial optimization*, volume 8096 of *Lecture Notes in Comput. Sci.*, pages 527–542. Springer, Heidelberg, 2013. doi:10.1007/978-3-642-40328-6\_37.
- 32 Cole Franks, Rafael Oliveira, Akshay Ramachandran, and Michael Walter. Near optimal sample complexity for matrix and tensor normal models via geodesic convexity. *arXiv preprint arXiv:2110.07583*, 2021.
- 33 Ankit Garg, Leonid Gurvits, Rafael Oliveira, and Avi Wigderson. A deterministic polynomial time algorithm for non-commutative rational identity testing. In *57th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2016*, pages 109–117. IEEE Computer Soc., Los Alamitos, CA, 2016. doi:10.1109/FOCS.2016.95.
- 34 Ankit Garg, Leonid Gurvits, Rafael Oliveira, and Avi Wigderson. Operator scaling: theory and applications. *Found. Comput. Math.*, 20(2):223–290, 2020. doi:10.1007/s10208-019-09417-z.

- 35 Craig Gentry. Fully homomorphic encryption using ideal lattices. In *Proceedings of the Forty-First Annual ACM Symposium on Theory of Computing*, STOC '09, pages 169–178, New York, NY, USA, 2009. Association for Computing Machinery. doi:10.1145/1536414.1536440.
- 36 Dorian Goldfeld. Beyond the last theorem. *Math Horizons*, 4(1):26–34, 1996. arXiv:<https://doi.org/10.1080/10724117.1996.11974985>, doi:10.1080/10724117.1996.11974985.
- 37 Oded Goldreich, Shafi Goldwasser, and Shai Halevi. Public-key cryptosystems from lattice reduction problems. In Burton S. Kaliski, editor, *Advances in Cryptology — CRYPTO '97*, pages 112–131, Berlin, Heidelberg, 1997. Springer Berlin Heidelberg.
- 38 Andrew Granville and Thomas Tucker. It's as easy as abc. *Notices of the American Mathematical Society*, 49, 01 2002.
- 39 Victor Guillemin and Shlomo Sternberg. *Symplectic techniques in physics*. Cambridge Univ. Press, Cambridge u.a., 1. publ., reprint. edition, 1986.
- 40 Leonid Gurvits. Classical complexity and quantum entanglement. *J. Comput. Syst. Sci.*, 69(3):448–484, 2004. doi:10.1016/j.jcss.2004.06.003.
- 41 Marshall Hall. Integral matrices  $a$  for which  $aa^t = mi$ . *Number Theory and Algebra*, pages 119 – 134, 1977. URL: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84887022329&partnerID=40&md5=5f91330ec7e4ddf22586cd2717a71b3a>.
- 42 Marshall Hall. Combinatorial completions. In *Advances in Graph Theory*, volume 3 of *Annals of Discrete Mathematics*, pages 111–123. Elsevier, 1978. URL: <https://www.sciencedirect.com/science/article/pii/S0167506008705016>, doi:[https://doi.org/10.1016/S0167-5060\(08\)70501-6](https://doi.org/10.1016/S0167-5060(08)70501-6).
- 43 Masaki Hamada and Hiroshi Hirai. Computing the nc-rank via discrete convex optimization on  $\text{cat}(0)$  spaces. *SIAM Journal on Applied Algebra and Geometry*, 5(3):455–478, 2021. arXiv:<https://doi.org/10.1137/20M138836X>, doi:10.1137/20M138836X.
- 44 Godfrey H. Hardy and Edward M. Wright. *An introduction to the theory of numbers*. Oxford University Press, Oxford, sixth edition, 2008. Revised by D. R. Heath-Brown and J. H. Silverman, With a foreword by Andrew Wiles.
- 45 Charles J. Himmelberg. Pseudo-metrizability of quotient spaces. *Fund. Math.*, pages 1–6, 1968.
- 46 Warren Hoburg, Philippe Kirschen, and Pieter Abbeel. Data fitting with geometric-programming-compatible softmax functions. *Optimization and Engineering*, 17(4):897–918, December 2016. doi:10.1007/s11081-016-9332-3.
- 47 Jeffrey Hoffstein, Jill Pipher, and Joseph H. Silverman. Ntru: A ring-based public key cryptosystem. In Joe P. Buhler, editor, *Algorithmic Number Theory*, pages 267–288, Berlin, Heidelberg, 1998. Springer Berlin Heidelberg.
- 48 Jeffrey Hoffstein and Joseph Silverman. Optimizations for ntru. In Kazimierz Alster, Jerzy Urbanowicz, and Hugh C. Williams, editors, *Proceedings of the International Conference organized by the Stefan Banach International Mathematical Center Warsaw, Poland, September 11-15, 2000*, pages 77–88, Berlin, New York, 2001. De Gruyter. URL: <https://doi.org/10.1515/9783110881035.77> [cited 2023-06-20], doi:doi:10.1515/9783110881035.77.
- 49 Gábor Ivanyos and Youming Qiao. On the orbit closure intersection problems for matrix tuples under conjugation and left-right actions, pages 4115–4126. Society for Industrial and Applied Mathematics, 2023. URL: <https://epubs.siam.org/doi/abs/10.1137/1.9781611977554.ch158>, arXiv:<https://epubs.siam.org/doi/pdf/10.1137/1.9781611977554.ch158>, doi:10.1137/1.9781611977554.ch158.
- 50 Gábor Ivanyos, Youming Qiao, and K. V. Subrahmanyam. Non-commutative Edmonds' problem and matrix semi-invariants. *Comput. Complexity*, 26(3):717–763, 2017. doi:10.1007/s00037-016-0143-x.
- 51 Gábor Ivanyos, Youming Qiao, and K. V. Subrahmanyam. Constructive non-commutative rank computation is in deterministic polynomial time. *Comput. Complexity*, 27(4):561–593, 2018. doi:10.1007/s00037-018-0165-7.

- 52 Valentine Kabanets and Russell Impagliazzo. Derandomizing polynomial identity tests means proving circuit lower bounds. *Computational Complexity*, 13(1-2):1–46, 2004. doi:10.1007/s00037-004-0182-6.
- 53 Ravi Kannan. Minkowski’s convex body theorem and integer programming. *Mathematics of Operations Research*, 12(3):415–440, 1987. URL: <http://www.jstor.org/stable/3689974>.
- 54 Ravindran Kannan and Achim Bachem. Polynomial algorithms for computing the Smith and Hermite normal forms of an integer matrix. *SIAM J. Comput.*, 8(4):499–507, 1979. doi:10.1137/0208040.
- 55 George Kempf and Linda Ness. The length of vectors in representation spaces. In Knud Lønsted, editor, *Algebraic Geometry*, pages 233–243, Berlin, Heidelberg, 1979. Springer Berlin Heidelberg.
- 56 Pascal Koiran. Hilbert’s Nullstellensatz is in the Polynomial Hierarchy. *Journal of Complexity*, 12(4):273–286, 1996. doi:<https://doi.org/10.1006/jcom.1996.0019>.
- 57 Joseph P. S. Kung and Gian-Carlo Rota. The invariant theory of binary forms. *Bull. Amer. Math. Soc. (N.S.)*, 10(1):27–85, 1984. doi:10.1090/S0273-0979-1984-15188-7.
- 58 Greg Kuperberg. Knottedness is in NP, modulo GRH. *Advances in Mathematics*, 256:493–506, 2014. URL: <https://www.sciencedirect.com/science/article/pii/S0001870814000188>, doi: <https://doi.org/10.1016/j.aim.2014.01.007>.
- 59 Serge Lang. *Elliptic Curves: Diophantine Analysis*, volume 231 of *Grundlehren der mathematischen Wissenschaften*. Springer, 2013.
- 60 Jonathan Leake and Nisheeth K. Vishnoi. On the computability of continuous maximum entropy distributions with applications. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2020, pages 930–943, New York, NY, USA, 2020. Association for Computing Machinery. doi:10.1145/3357713.3384302.
- 61 Arjen K. Lenstra, Hendrik. W. Lenstra, and László Lovász. Factoring polynomials with rational coefficients. *Mathematische Annalen*, 261(4):515–534, December 1982. doi:10.1007/BF01457454.
- 62 Seymour Lipschutz. *Schaum’s outline of theory and problems of linear algebra : [including 600 solved problems ; completely solved in detail]*. Schaum’s outline series. McGraw-Hill, New York u.a., 6th ed. edition, 1974.
- 63 Visu Makam and Avi Wigderson. Singular tuples of matrices is not a null cone (and the symmetries of algebraic varieties). *J. Reine Angew. Math.*, 780:79–131, 2021. doi:10.1515/crelle-2021-0044.
- 64 Petros Maragos, Vasileios Charisopoulos, and Emmanouil Theodosis. Tropical geometry and machine learning. *Proceedings of the IEEE*, 109(5):728–755, 2021. doi:10.1109/JPROC.2021.3065238.
- 65 Jerrold Marsden and Alan Weinstein. Reduction of symplectic manifolds with symmetry. *Reports on Mathematical Physics*, 5(1):121–130, 1974. URL: <https://www.sciencedirect.com/science/article/pii/0034487774900214>, doi:[https://doi.org/10.1016/0034-4877\(74\)90021-4](https://doi.org/10.1016/0034-4877(74)90021-4).
- 66 E. M. Matveev. An explicit lower bound for a homogeneous rational linear form in logarithms of algebraic numbers. *Izvestiya: Mathematics*, 62(4):723, aug 1998. URL: <https://dx.doi.org/10.1070/IM1998v062n04ABEH000190>, doi:10.1070/IM1998v062n04ABEH000190.
- 67 Daniele Micciancio and Shafi Goldwasser. *Complexity of lattice problems : a cryptographic perspective*. The Kluwer international series in engineering and computer science BV000632170 671. Kluwer Academic, Boston, 2002.
- 68 Gary L. Miller. Riemann’s hypothesis and tests for primality. *Journal of Computer and System Sciences*, 13(3):300–317, 1976. URL: <https://www.sciencedirect.com/science/article/pii/S0022000076800438>, doi:[https://doi.org/10.1016/S0022-0000\(76\)80043-8](https://doi.org/10.1016/S0022-0000(76)80043-8).
- 69 Louis J. Mordell. On the representation of a binary quadratic form as a sum of squares of linear forms. *Mathematische Zeitschrift*, 35(1-15):1432–1823, 1932. URL: <https://doi.org/10.1007/BF01186544>.

- 70 Ketan Mulmuley. Geometric complexity theory V: Efficient algorithms for Noether normalization. *J. Amer. Math. Soc.*, 30(1):225–309, 2017. doi:10.1090/jams/864.
- 71 Ketan Mulmuley and Milind Sohoni. Geometric complexity theory I: An approach to the P vs. NP and related problems. *SIAM Journal on Computing*, 31(2):496–526, 2001.
- 72 David Mumford. *The red book of varieties and schemes*, volume 1358 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1999. doi:10.1007/b62130.
- 73 David Mumford, John Fogarty, and Frances Kirwan. *Geometric invariant theory*, volume 34 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (2) [Results in Mathematics and Related Areas (2)]*. Springer-Verlag, Berlin, third edition, 1994. doi:10.1007/978-3-642-57916-5.
- 74 Arkadi Nemirovski and Uriel Rothblum. On complexity of matrix scaling. *Linear Algebra and its Applications*, 302-303:435–460, 1999. URL: <https://www.sciencedirect.com/science/article/pii/S0024379599002128>, doi:[https://doi.org/10.1016/S0024-3795\(99\)00212-8](https://doi.org/10.1016/S0024-3795(99)00212-8).
- 75 Joseph Oesterlé. Nouvelles approches du théorème de Fermat. In *Séminaire Bourbaki : volume 1987/88, exposés 686-699*, number 161-162 in *Astérisque*. Société mathématique de France, 1988. talk:694. URL: [http://www.numdam.org/item/SB\\_1987-1988\\_\\_30\\_\\_165\\_0/](http://www.numdam.org/item/SB_1987-1988__30__165_0/).
- 76 Victor Y. Pan and Zhao Q. Chen. The complexity of the matrix eigenproblem. In *Annual ACM Symposium on Theory of Computing (Atlanta, GA, 1999)*, pages 507–516. ACM, New York, 1999. doi:10.1145/301250.301389.
- 77 Michael O. Rabin and Jeffery O. Shallit. Randomized algorithms in number theory. *Communications on Pure and Applied Mathematics*, 39(S1):S239–S256, 1986. doi:<https://doi.org/10.1002/cpa.3160390713>.
- 78 Alexander Schrijver. *Theory of linear and integer programming*. Wiley-Interscience series in discrete mathematics and optimization. Wiley, Chichester u.a., reprinted edition, 1999.
- 79 Mohit Singh and Nisheeth K. Vishnoi. Entropy, optimization and counting. In *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*, STOC '14, pages 50–59, New York, NY, USA, 2014. Association for Computing Machinery. doi:10.1145/2591796.2591803.
- 80 Cameron L. Stewart and R. Tijdeman. On the Oesterlé-Masser conjecture. *Monatshefte für Mathematik*, 102(3):251–257, September 1986. doi:10.1007/BF01294603.
- 81 Cameron L. Stewart and Kunrui Yu. On the abc conjecture. *Mathematische Annalen*, 291(2):225–230, 1991. URL: <http://eudml.org/doc/164860>.
- 82 Damian Straszak and Nisheeth K. Vishnoi. Maximum entropy distributions: Bit complexity and stability. In Alina Beygelzimer and Daniel Hsu 0001, editors, *Conference on Learning Theory, COLT 2019, 25-28 June 2019, Phoenix, AZ, USA*, volume 99 of *Proceedings of Machine Learning Research*, pages 2861–2891. PMLR, 2019. URL: <http://proceedings.mlr.press/v99/straszak19a.html>.
- 83 Bernd Sturmfels. *Algorithms in Invariant Theory*. Texts & Monographs in Symbolic Computation. Springer, 2008. doi:10.1007/978-3-211-77417-5.
- 84 Alfred J. van der Poorten. On Baker’s inequality for linear forms in logarithms. *Mathematical Proceedings of the Cambridge Philosophical Society*, 80(2):233–248, 1976. doi:10.1017/S0305004100052877.
- 85 Michel Waldschmidt. *Diophantine approximation on linear algebraic groups : transcendence properties of the exponential function in several variables*. Grundlehren der mathematischen Wissenschaften BV000000395 326. Springer, Berlin u.a., 2000.
- 86 Michel Waldschmidt. Open diophantine problems. *Mosc. Math. J.*, 4:245–305, 2004.
- 87 Michel Waldschmidt. Lecture on the abc conjecture and some of its consequences. In Pierre Cartier, A.D.R. Choudary, and Michel Waldschmidt, editors, *Mathematics in the 21st Century*, pages 211–230, Basel, 2015. Springer Basel.
- 88 Gisbert Wüstholz and Alan Baker. Logarithmic forms and group varieties. *Journal für die reine und angewandte Mathematik*, 442:19–62, 1993. URL: <http://eudml.org/doc/153550>.