CONSTRAINED INHOMOGENEOUS SPHERICAL EQUATIONS: AVERAGE-CASE HARDNESS

ALEXANDER USHAKOV

Department of Mathematical Sciences, Stevens Institute of Technology, Hoboken NJ 07030, USA
e-mail address: aushakov@stevens.edu

ABSTRACT. In this paper we analyze computational properties of the Diophantine problem (and its search variant) for spherical equations $\prod_{i=1}^{m} z_i^{-1} c_i z_i = 1$ (and their variants) over the class of finite metabelian groups $G_{p,n} = \mathbb{Z}_p^n \rtimes \mathbb{Z}_p^*$, where $n \in \mathbb{N}$ and p is prime. We prove that the problem of finding solutions for certain constrained spherical equations is computationally hard on average (assuming that some lattice approximation problem is hard in the worst case).

1. INTRODUCTION

The main goal of this paper is to create bridges between problems of computational group theory (namely the problem of finding a solution for a spherical equation) and assumptions of lattice-based cryptography.

Modern uses of lattices in the design of cryptographic primitives started in 1996 with the paper [3], where M. Ajtai introduced a class of random problems hard for probabilistic polynomial-time (PPT) algorithms on any cryptographically non-negligible sets of instances unless certain lattice approximation problems (such as **SIVP**$_\gamma$, see Section 3.1) can be solved efficiently in the worst case (on every input). Since then Ajtai's construction was studied extensively; it was improved in different ways, various interesting applications were found, and it gained a lot of popularity, see [36] for a survey. Nowadays, lattice-based cryptography appears to be the most promising branch of post-quantum cryptography.

1.1. **Equations in groups.** Let $F = F(Z)$ denote the free group on countably many generators $Z = \{z_i\}_{i=1}^{\infty}$. For a group G, an *equation over G with variables in Z* is an equality of the form $W = 1$, where $W \in F * G$. If $W = z_{i_1}^{\varepsilon_1} g_1 \cdots z_{i_k}^{\varepsilon_k} g_k$, with $z_{i_j} \in Z$, $\varepsilon_j = \pm 1$, and $g_j \in G$, then we refer to $\{z_{i_1}, \ldots, z_{i_k}\}$ as the set of *variables* and to $\{g_1, \ldots, g_k\}$ as the set of *constants* (or *coefficients*) of W. We occasionally write $W(z_1, \ldots, z_k)$ or $W(z_1, \ldots, z_k; g_1, \ldots, g_k)$ to indicate that the variables in W are precisely $z_1, \ldots, z_k$ and

(in the latter case) the constants are precisely $g_1, \ldots, g_k$. A *solution* for an equation $W(z_1, \ldots, z_k) = 1$ over G is an assignment for variables $z_1, \ldots, z_k$ that makes $W = 1$ true.

In this paper we assume that G comes equipped with a fixed generating set X and elements of G are given as products of elements of X and their inverses. This naturally defines the length (or size) of the equation $W = 1$ as the length of its left-hand side W.

The Diophantine problem over a group G for a class of equations C is an algorithmic question to decide whether a given equation $W = 1$ in C has a solution.

By definition, the Diophantine problem is a *decision* problem (a yes/no question). Additionally, one can study the corresponding *search* problem that requires to find a solution for $W = 1$, provided one exists.

1.2. **Constrained equations.**
A *constrained equation* over a group G is an equation $W(z_1, \ldots, z_n) = 1$ equipped with a set $Z \subseteq G^n$. The Diophantine problem for a constrained equation requires to decide whether the following question has a solution or not:

$$\begin{cases} W(z_1, \ldots, z_n) = 1, \\ (z_1, \ldots, z_n) \in Z. \end{cases}$$

In this paper we work with a set of the form $Z = Z_1 \times \cdots \times Z_n$, in which case every variable z_i is constrained individually by Z_i.

Recently, constrained equations in groups have attracted significant attention, with most research focusing on free groups, hyperbolic groups, right-angled Artin groups, virtually abelian groups, and certain classes of finite groups. For a survey of recent results in this area see [7].

1.3. **Spherical equations.**
One class of equations over groups that has generated much interest is the class of *quadratic* equations: equations where each variable z appears exactly twice (as either z or z^{-1}). It was observed in the early 80's [9, 37] that such equations have an affinity with the theory of compact surfaces (for instance, via their associated van Kampen diagrams). This geometric point of view sparked the initial interest in their study and has led to many interesting results, particularly in the realm of quadratic equations over free groups: solution sets were studied in [16], **NP**-completeness was proved in [10, 20]. Systems of quadratic equations played an important role in the study of the first order theory of free groups (Tarski problem, [22]). quadratic equations in various classes of (infinite) groups such as hyperbolic groups (solution sets described in [17], **NP**-complete by [21]), the first Grigorchuk group (decidability proved in [25], commutator width computed in [5]), free metabelian groups (**NP**-hard by [26], in **NP** for orientable equations by [27]), metabelian Baumslag–Solitar groups (**NP**-complete by [28]), etc.

We say that equations $W = 1$ and $V = 1$ are *equivalent* if there is an automorphism $\phi \in \mathrm{Aut}(F * G)$ such that ϕ is the identity on G and $\phi(W) = V$. It is a well known consequence of the classification of compact surfaces that any quadratic equation over G is equivalent, via an automorphism ϕ, computable in time $O(|W|^2)$, to an equation in exactly

one of the following three *standard forms* (see [8, 16]):

$$\prod_{j=1}^{m} z_j^{-1} c_j z_j = 1 \qquad\qquad m \geq 1, \qquad\qquad (1.1)$$

$$\prod_{i=1}^{g} [x_i, y_i] \prod_{j=1}^{m} z_j^{-1} c_j z_j = 1 \qquad\qquad g \geq 1, m \geq 0, \qquad\qquad (1.2)$$

$$\prod_{i=1}^{g} x_i^2 \prod_{j=1}^{m} z_j^{-1} c_j z_j = 1 \qquad\qquad g \geq 1, m \geq 0. \qquad\qquad (1.3)$$

The number g is the *genus* of the equation, and both g and m (the number of constants) are invariants. The standard forms are called, respectively, *spherical, orientable of genus g,* and *non-orientable of genus g.*

In this paper we investigate spherical equations in finite groups. We say that the equation (1.1) is a *homogeneous* form of a spherical equation. For $m \in \mathbb{N}$ define the set of all homogeneous equations with m conjugates by

$$\mathrm{Sph}_m = \left\{ \prod_{j=1}^{m} z_j^{-1} c_j z_j = 1 \;\middle|\; c_1, \ldots, c_m \in G \right\} \quad \text{and} \quad \mathrm{Sph} = \bigcup_{m=1}^{\infty} \mathrm{Sph}_m.$$

It is easy to show that an equation of the form

$$\prod_{j=1}^{m} z_j^{-1} c_j z_j = c \qquad\qquad (1.4)$$

is equivalent to a spherical equation. We say that an equation (1.4) is an *inhomogeneous* form of a spherical equation.

Notice that spherical equations naturally generalize fundamental (Dehn) problems of group theory, as solving equations from Sph_1 is the same as solving the word problem and solving equations from Sph_2 is the same as solving the conjugacy problem. The complexity of solving equations in finite groups has been first studied by Goldmann and Russell [13] showing that the Diophantine problem in a fixed finite nilpotent group can be decided in polynomial time, while it is **NP**-complete in every finite non-solvable group. For some recent results, see [11, 18, 29].

1.4. **Groups under consideration.**

Consider the group $\mathbb{Z}_p^n = \{ (x_1, \ldots, x_n) \mid x_i \in \mathbb{Z}_p \}$. The group of units $\mathbb{Z}_p^*$ acts on $\mathbb{Z}_p^n$ by (scalar) multiplication

$$(x_1, \ldots, x_n) \overset{\alpha}{\mapsto} (\alpha x_1, \ldots, \alpha x_n),$$

where $\alpha \in \mathbb{Z}_p^*$ and $(x_1, \ldots, x_n) \in \mathbb{Z}_p^n$. The semidirect product $G = \mathbb{Z}_p^n \rtimes \mathbb{Z}_p^*$ is a set of pairs $(\overline{x}, \alpha)$ equipped with the binary operation

$$(\overline{x}, \alpha)(\overline{y}, \beta) = (\overline{x} + \alpha \overline{y}, \alpha \cdot \beta),$$

with the identity $(\overline{0}, 1)$. The following useful formulae are used throughout the paper without referencing:

$$(\overline{x}, \alpha)^{-1} = (-\alpha^{-1} \overline{x}, \alpha^{-1}),$$

$$(\overline{x}, \alpha)^{-1}(\overline{y}, \beta)(\overline{x}, \alpha) = (\alpha^{-1}((\beta - 1)\overline{x} + \overline{y}), \beta).$$

1.5. **Model of computation.** We assume that all computations are performed on a random access machine. Elements of $\mathbb{Z}_p$ are given in binary as bit-strings of length $\lceil \log_2(p) \rceil$. Elements of $\mathbb{Z}_p^n \rtimes \mathbb{Z}_p^*$ are given as $n+1$-tuples of elements from $\mathbb{Z}_p$. If $f(n)$ is $O(T(n) \cdot n^\varepsilon)$ for every $\varepsilon > 0$, then we say that a function $f(n)$ is "nearly $T(n)$" and write $f(n)$ is $\tilde{O}(T(n))$. Operations in $\mathbb{Z}_p$ have the following complexity:

- Addition and subtraction can be done in $O(\lceil \log_2(p) \rceil)$ time in a straightforward way.
- Multiplication can be done in $\tilde{O}(\lceil \log_2(p) \rceil)$ time using fast Fourier transform.
- Computing the multiplicative inverse of a unit modulo p can be done in $O(\lceil \log_2(p) \rceil^2)$ time using the extended Euclidean algorithm.

2. PRELIMINARIES: SUBSET SUM PROBLEM IN GROUPS

Here we review several useful definitions of discrete optimization in groups. Let G be a group generated by a finite set $X = \{x_1, \ldots, x_n\} \subseteq G$. Elements in G can be expressed as products of the generators in X and their inverses. Hence, we can state the following combinatorial problem.

The subset sum problem SSP(G, X): Given $g_1, \ldots, g_k, g \in G$ decide if

$$g = g_1^{\varepsilon_1} \cdots g_k^{\varepsilon_k} \tag{2.1}$$

for some $\varepsilon_1, \ldots, \varepsilon_k \in \{0, 1\}$.

By [30, Proposition 2.5], computational properties of **SSP** do not depend on the choice of a finite generating set X and, hence, the problem can be abbreviated as **SSP**(G). The same paper, [30], provides a variety of examples of groups with **NP**-complete (or polynomial time) subset sum problems.

Consider the infinitely generated group $\mathbb{Z}_3^\omega$ whose elements can be formally viewed as functions $f \colon \mathbb{N} \to \mathbb{Z}_3$ with finite support. For algorithmic purposes, we assume that elements of $\mathbb{Z}_3^\omega$ are encoded by finite ternary strings (as in [34, Section 4]). [35, Proposition 2.1] proves that **SSP**$(\mathbb{Z}_3^\omega)$ is **NP**-complete, which can be reformulated as follows.

Proposition 2.1 ([35, Proposition 2.1])**.** *For every $m \geq 3$, **SSP** is **NP**-complete for the class of finite groups $\{\mathbb{Z}_m^n\}_{n=1}^\infty$.*

3. PRELIMINARIES: LATTICE PROBLEMS

3.1. **Lattices.** Here we review several definitions of the theory of lattices. Recall that a set of points $S \subseteq \mathbb{R}^n$ is *discrete* if every point $x \in \mathbb{R}^n$ has an ε-neighborhood that contains x only. A *lattice* is a discrete subgroup of $\mathbb{R}^n$. An *integer lattice* is a subgroup of $\mathbb{Z}^n$. We discuss integer lattices only. We say that $\mathcal{L} \leq \mathbb{Z}^n$ is a *full-rank* lattice if the dimension of the corresponding vector space $\mathrm{Span}(\mathcal{L})$ is n (we can simply write $\dim(\mathcal{L}) = n$). We typically assume that $\mathcal{L}$ is a full-rank lattice.

The *minimum distance* of a lattice $\mathcal{L} \leq \mathbb{R}^n$ is the length of a shortest nonzero lattice vector

$$\lambda_1(\mathcal{L}) = \min_{\overline{v} \in \mathcal{L} \setminus \{\overline{0}\}} \|\overline{v}\|.$$

More generally, for $i = 1, \ldots, n$, the ith *successive minimum* of $\mathcal{L}$ is

$$\lambda_i(\mathcal{L}) = \min \left\{ r \mid \dim(\mathrm{Span}(B(r) \cap \mathcal{L})) \geq i \right\},$$

which is the minimum radius of a ball that contains at least i linearly independent points. Below we recall several computational problems for lattices. Some of them are of direct importance to lattice-based cryptography (e.g., $\mathbf{SIVP}_\gamma$ and $\mathbf{GapSVP}_\gamma$) and others have more historical importance.

Shortest vector problem, SVP. Given a basis of a lattice $\mathcal{L} \leq \mathbb{Z}^n$, find a shortest nonzero vector $\overline{v} \in \mathcal{L}$, i.e., a vector satisfying $\|\overline{v}\| = \lambda_1(\mathcal{L})$.

Approximate shortest vector problem, $\mathbf{SVP}_\gamma$. Given a basis of a lattice $\mathcal{L} \leq \mathbb{Z}^n$, find a nonzero vector $\overline{v} \in \mathcal{L}$ satisfying $\|\overline{v}\| \leq \gamma \cdot \lambda_1(\mathcal{L})$.

Decisional approximate SVP, $\mathbf{GapSVP}_\gamma$. Given a basis of a lattice $\mathcal{L}$, where either $\lambda_1(\mathcal{L}) \leq 1$ or $\lambda_1(\mathcal{L}) > \gamma$, decide which is the case.

Shortest independent vector problem, SIVP. Given a basis of a full-rank lattice $\mathcal{L} \leq \mathbb{Z}^n$, find n linearly independent vectors $\overline{v}_1, \ldots, \overline{v}_n \in \mathcal{L}$ satisfying $\max \|\overline{v}_i\| \leq \lambda_n(\mathcal{L})$.

Approximate shortest independent vectors problem, $\mathbf{SIVP}_\gamma$. Given a basis of a full-rank lattice $\mathcal{L} \leq \mathbb{Z}^n$, find n linearly independent vectors $\overline{v}_1, \ldots, \overline{v}_n \in \mathcal{L}$ satisfying

$$\max \|\overline{v}_i\| \leq \gamma(n) \cdot \lambda_n(\mathcal{L}).$$

Formally, we say that $\mathbf{SIVP}_\gamma$ is hard in the worst case for probabilistic polynomial-time (PPT) algorithms, if for every PPT algorithm $\mathcal{A}$ and for every $n \in \mathbb{N}$ there is an n-dimensional basis $B_{n,\mathcal{A}} = \{\overline{b}_1, \ldots, \overline{b}_n\} \subseteq \mathbb{Z}^n$ satisfying

$$\Pr[\mathcal{A} \text{ solves } \mathbf{SIVP}_\gamma \text{ for } B_{n,\mathcal{A}}] \quad \text{is} \quad o(n^{-d}), \tag{3.1}$$

for every $d > 0$, where the probability is taken over the coin-tosses of $\mathcal{A}$.

The above lattice problems have been intensively studied and appear to be intractable, except for very large approximation factors $\gamma(n)$. Known polynomial-time algorithms like the Lenstra–Lenstra–Lovász [24] and its descendants obtain only slightly subexponential approximation factors for all the above problems. Known algorithms that obtain polynomial approximation factors, such as [19, 4, 33, 1], either require superexponential time, or exponential time and space.

Many lattice problems are **NP**-hard, even to approximate to within various sub-polynomial $n^{o(1)}$ approximation factors. However, such hardness is not of any direct consequence to cryptography, since lattice-based cryptographic constructions so far rely on polynomial approximation problems factors $\gamma(n) \geq n$. Indeed, there is evidence that for factors $\gamma(n) \geq \sqrt{n}$, the lattice problems relevant to cryptography are not **NP**-hard, because they lie in $\mathbf{NP} \cap \mathbf{coNP}$ [14, 2]. For a survey on lattice-based cryptography see [36].

3.2. **Short integer solution problem.** In general, the *short integer solution* problem can be formulated as follows.

Short integer solution (SIS) problem. For a given matrix $A \in \mathbb{Z}_p^{n \times m}$, find $\overline{x} \in \mathbb{Z}_p^m$ satisfying

$$\begin{cases} A\overline{x} \equiv_p \overline{0}, \\ \overline{x} \text{ is short and nontrivial}, \end{cases} \tag{3.2}$$

assuming that a solution exists.

In other words, **SIS** requires to find a short and nontrivial solution for a homogeneous system of linear congruences $A\overline{x} \equiv_p \overline{0}$. In the original paper [3], $\overline{x}$ was called short if $\|\overline{x}\|_2 \leq n$ (can be relaxed to $\|\overline{x}\|_2 \leq poly(n)$). In [15], instead of $\|\overline{x}\|_2 \leq n$, the following two constraints on $\overline{x}$ were discussed:

- $\overline{x} \neq 0$ and $x_i \in \{0, 1\}$;
- $\overline{x} \neq 0$ and $x_i \in \{-1, 0, 1\}$.

We denote the corresponding versions of **SIS** by $\mathbf{SIS}_{\{0,1\}}$ and $\mathbf{SIS}_{\{-1,0,1\}}$. To discuss the average case complexity of **SIS** we use the uniform distribution on its instances.

Randomized SIS problem. For a matrix $A \in \mathbb{Z}_p^{n \times m}$ sampled uniformly randomly, find $\overline{x} \in \mathbb{Z}_p^m$ such that $A\overline{x} = \overline{0}$, and $\overline{x}$ is short and nontrivial.

Theorem 3.1 (Goldreich, Goldwasser, Halevi, [15, Theorem 1], cf. Ajtai, [3, Theorem 1])**.** *Suppose that a PPT algorithm $\mathcal{A}$ solves the randomized $\mathbf{SIS}_{\{0,1\}}$ problem (or $\mathbf{SIS}_{\{-1,0,1\}}$ problem) with parameters n, m, p satisfying*

$$n \log(p) < m < \frac{p}{2n^4} \quad and \quad p = O(n^c) \text{ for some } c > 0, \tag{3.3}$$

with probability at least n^{-c_0} for some fixed constant $c_0 > 0$, where the probability is taken over the choice of the instance as well as the coin-tosses of $\mathcal{A}$. Then there is a PPT algorithm that solves $\mathbf{GapSVP}_{\gamma=pn^6}$ and $\mathbf{SIVP}_{\gamma=pn^6}$ (among others) on every n-dimensional lattice with probability at least $1 - 2^{-n}$. [1]

In a similar way we can define deterministic and randomized *inhomogeneous short integer solution* (ISIS) problem.

Randomized ISIS problem. For a matrix $A \in \mathbb{Z}_p^{n \times m}$ and a vector $\overline{y} \in \mathbb{Z}_p^n$ sampled uniformly randomly, find $\overline{x} \in \mathbb{Z}_p^m$ such that $A\overline{x} = \overline{y}$, and $\overline{x}$ is short.

Corollary 3.2. *The statement of Theorem 3.1 holds for the randomized $\mathbf{ISIS}_{\{0,1\}}$ problem.*

Proof. Indeed, suppose that there is a PPT algorithm $\mathcal{A}$ that solves $\mathbf{ISIS}_{\{0,1\}}$ problem with probability at least n^{-c_0}. Then for an instance A of $\mathbf{SIS}_{\{0,1\}}$ with columns $\overline{v}_1, \ldots, \overline{v}_m$

(1) "guess" a nonzero index i in a solution $\overline{x} \in \{0, 1\}^m$,
(2) form an instance $(A', -\overline{v}_i)$, where A' is obtained by deleting the ith column from A,
(3) solve $(A', -\overline{v}_i)$ using $\mathcal{A}$.

This gives a PPT algorithm that solves $\mathbf{SIS}_{\{0,1\}}$ with probability at least n^{-c_0}. $\qquad\square$

4. Spherical equations over $G_{p,n}$

In this section we analyse complexity of solving a spherical equation (1.1) over $G_{p,n}$, with constants $c_i = (\overline{c}_i, \beta_i) \in G_{p,n}$ and unknowns $z_1, \ldots, z_m$.

Lemma 4.1. $z_i = (\overline{z}_i, \alpha_i)$ *satisfy* (1.1) $\Leftrightarrow$ *the following two conditions are satisfied:*

(S1) $\beta_1 \cdots \beta_m = 1$;
(S2) $\sum_{i=1}^m B_i \alpha_i^{-1}((\beta_i - 1)\overline{z}_i + \overline{c}_i) \equiv_p \overline{0}$, *where* $B_i = \beta_1 \cdots \beta_{i-1}$.

Proof. Straightforward verification. $\qquad\square$

---

[1] For better bounds and approximation factors of γ see [32].

4.1. Spherical equations over $G_{p,n}$: generic-case.

Here, we investigate generic-case hardness for spherical equations, i.e., hardness of a "typical" equation, see [31, Chapter 10] for basic definitions of generic-case complexity.

Proposition 4.2. *If $\beta_i \neq 1$ for some i, then (1.1) has a solution and a solution can be found in polynomial time.*

Proof. It takes nearly linear time to compute $B_1, \ldots, B_m$ and check the condition (S1). If $\beta_i \neq 1$, then the following assignment:

- $\alpha_1 = \cdots = \alpha_m = 1$,
- $\bar{z}_j = \bar{0}$ for $j \neq i$,
- $\bar{z}_i \equiv_p \frac{1}{B_i(\beta_i - 1)} \sum_{j=1}^m B_j \bar{c}_j$

is a solution that can be computed in polynomial time. $\qquad\square$

Next, we claim that the property $\exists i \; \beta_i \neq 1$ is strongly generic, i.e., a typical equation satisfies this property. Since the problem involves three parameters, n, m, and p, we use the following stratification for the set of instances of the uniform problem. For $s \in \mathbb{N}$ define a set of pairs

$$\mathcal{I}_s = \{ (E_m, G_{p,n}) \mid E_m \in \mathrm{Sph}_m(G_{p,n}), \; m, n, p \leq s \}$$

equipped with the uniform distribution. Then the following holds:

$$|G_{p,n}| = p^n \cdot (p-1),$$
$$|\mathrm{Sph}_m| = (p^n \cdot (p-1))^m,$$
$$|\mathcal{I}_s| = \sum_{p,n,m \leq s} (p^n \cdot (p-1))^m.$$

Lemma 4.3. $\Pr\{ (E_m, G_{p,n}) \in \mathcal{I}_s \mid m \geq s/2 \} \to 1$ *exponentially fast as $s \to \infty$.*

Proof. For any fixed $p \geq 2, n \geq 1$ we have

$$V_s = \sum_{m=0}^s (p^n \cdot (p-1))^m = \frac{(p^n \cdot (p-1))^{s+1} - 1}{p^n \cdot (p-1) - 1},$$

which implies that

$$\frac{V_{s/2}}{V_s} = \frac{(p^n \cdot (p-1))^{s/2+1} - 1}{(p^n \cdot (p-1))^{s+1} - 1} \leq \frac{1}{(p^n \cdot (p-1))^{s/2}} \leq \frac{1}{2^{s/2}}$$

that converges to 0 exponentially fast. The obtained bound $\frac{1}{2^{s/2}}$ does not depend on p or n. Therefore, the bound works on the whole $\mathcal{I}_s$. $\qquad\square$

Lemma 4.4. $\Pr\{ (E_m, G_{p,n}) \in \mathcal{I}_s \mid \exists i \; \beta_i \neq 1 \} \to 1$ *exponentially fast as $s \to \infty$.*

Proof. By Lemma 4.3, we may assume that $m \geq s/2$, in which case

$$\Pr\{ (E_m, G_{p,n}) \in \mathcal{I}_s \mid \exists i \; \beta_i \neq 1 \} \geq 1 - \frac{1}{(p-1)^{s/2}},$$

which converges to 1 exponentially fast as $s \to \infty$. $\qquad\square$

Corollary 4.5. *The Diophantine (decision) problem for spherical equations over $\{G_{p,n}\}$ is decidable in strongly generically linear time. The corresponding search problem can be solved in strongly generically polynomial time.*

Proof. Follows from Proposition 4.2 and Lemma 4.4   □

Since spherical equations in which at least one coefficient $c_i = (\bar{c}_i, \beta_i)$ satisfies $\beta_i \neq 1$ are computationally easy, we put a restriction on the coefficients that we use in spherical equations. Define the set

$$C_{p,n} = \left\{ (\bar{c}, 1) \mid \bar{c} \in \mathbb{Z}_p^n \right\}. \tag{4.1}$$

From now on we assume that all $c_i \in C_{p,n}$. In that case condition (S1) of Lemma 4.1 is trivially satisfied and condition (S2) translates into the following:

$$\exists \alpha_i \in \mathbb{Z}_p^* \quad \text{s.t.} \quad \sum_{i=1}^{m} \alpha_i^{-1} \bar{y}_i \equiv_p \bar{0}. \tag{4.2}$$

The obtained condition defines a homogeneous system of linear congruences that does not allow zero values for unknowns α_i, which makes the problem nontrivial (we conjecture it is **NP**-hard).

4.2. Spherical equations over $G_{p,n}$: worst case.

Here we investigate the worst case complexity for spherical equations. For $\bar{w} \in \mathbb{Z}_3^n$ define

$$c_{\bar{w}} = (\bar{w}, 1) \in \mathbb{Z}_3^n \rtimes \mathbb{Z}_3^*.$$

Consider an instance I of $\bar{v}_1, \ldots, \bar{v}_m, \bar{v} \in \mathbb{Z}_3^n$ of **SSP**. Recall that I is a positive instance if $\varepsilon_1 \bar{v}_1 + \cdots + \varepsilon_m \bar{v}_m = \bar{v}$ for some $\varepsilon_1, \ldots, \varepsilon_m \in \{0, 1\}$. For I define $\bar{v}' = \bar{v} + \sum \bar{v}_i$ and the spherical equation E_I as

$$\prod_i^m z_i^{-1} c_{\bar{v}_i} z_i = c_{\bar{v}'}. \tag{4.3}$$

Proposition 4.6. *I is a positive instance of* **SSP** $\quad \Leftrightarrow \quad$ *E_I has a solution.*

Proof. By design we have

$$I \text{ is a positive instance of } \textbf{SSP} \;\Leftrightarrow\; \bar{v} = \sum_{i=1}^{m} \varepsilon_i \bar{v}_i \text{ for } \varepsilon_1, \ldots, \varepsilon_m \in \{0, 1\}$$

$$\Leftrightarrow\; \bar{v} + \sum_{i=1}^{m} \bar{v}_i = \sum_{i=1}^{m} \beta_i \bar{v}_i \text{ for } \beta_1, \ldots, \beta_m \in \{1, 2\} = \mathbb{Z}_3^*$$

$$\Leftrightarrow\; \bar{v} + \sum_{i=1}^{m} \bar{v}_i = \sum_{i=1}^{m} \alpha_i^{-1} \bar{v}_i \text{ for } \alpha_1, \ldots, \alpha_m \in \{1, 2\} = \mathbb{Z}_3^*$$

$$\overset{4.1}{\Leftrightarrow}\; E_I \text{ has a solution.}$$

        □

Corollary 4.7. *The Diophantine problem for spherical equations over $\mathbb{Z}_{3,n} \rtimes \mathbb{Z}_3^*$ is* **NP**-*complete.*

Proof. By Proposition 4.6, $I \to E_I$ is a many-one polynomial-time reduction of **SSP**$(\mathbb{Z}_3^n)$ to spherical equations over **SSP**$(\mathbb{Z}_{3,n} \rtimes \mathbb{Z}_3^*)$. By Proposition 2.1, **SSP** is **NP**-complete for groups $\{\mathbb{Z}_3^n\}_{n=1}^{\infty}$. Hence the result.   □

5. CONSTRAINED SPHERICAL EQUATIONS: AVERAGE-CASE HARDNESS

In this section we discuss spherical equations over the class of groups $\mathcal{G} = \{G_{p,n}\}$ with constraints on values of z_i's. In full generality the problem can be formulated as follows. Given a group $G \in \mathcal{G}$, a spherical equation $\prod_{i=1}^{m} z_i^{-1} c_i z_i = 1$ over G, and subsets $Z_1, \ldots, Z_m \subseteq G$, find $z_1, \ldots, z_m \in G$ satisfying

$$\begin{cases} \prod_{i=1}^{m} z_i^{-1} c_i z_i = 1 \\ z_i \in Z_i \end{cases}$$

or a similar *inhomogeneous* form

$$\begin{cases} \prod_{i=1}^{m} z_i^{-1} c_i z_i = c \\ z_i \in Z_i. \end{cases}$$

The problem to decide if a given constrained inhomogeneous spherical equation has a solution is abbreviated **CISE**.

It is not hard to prove that **CISE** for groups $G_{p,n}$ is hard in the worst case. In fact, for $p = 3$, the proof of Proposition 4.6 works as is, with constraints $Z_i = \{(\overline{0}, 1), (\overline{0}, 2)\}$ (so chosen sets Z_i effectively define no constraints). Using these constraints we can easily achieve the same result for any odd value of p.

Theorem 5.1. CISE *is* **NP***-hard for groups* $\{G_{p,n}\}_n$ *for any fixed odd prime* p.

Proof. For a given instance $\overline{v}_1, \ldots, \overline{v}_m, \overline{v} \in \mathbb{Z}_p^n$, call it I, of **SSP**$(\mathbb{Z}_p^n)$, as in the proof of Proposition 4.6, construct a constrained inhomogeneous spherical equation E_I over $\mathbb{Z}_p^n \rtimes \mathbb{Z}_p^*$

$$\begin{cases} \prod_{i=1}^{m} z_i^{-1} c_{\overline{v}_i} z_i = c_{\overline{v}'} \\ z_i = (\overline{0}, 1) \text{ or } (\overline{0}, 2^{-1}) \end{cases}$$

and observe that I is a positive instance of **SSP**$(\mathbb{Z}_p^n)$ if and only if E_I has a solution. This gives a polynomial time reduction from **SSP** over groups $\{\mathbb{Z}_p^n\}_{n=1}^{\infty}$ to decidability of constrained inhomogeneous spherical equations over $\mathbb{Z}_p^n \rtimes \mathbb{Z}_p^*$. $\square$

Our next goal is to demonstrate the average-case hardness of constrained spherical equations. To achieve this, we randomize equations as follows. First, we stratify parameters:

- the parameter $n \in \mathbb{N}$ is the main independent parameter;
- parameters m and p are functions of n satisfying conditions (3.3).

Then, for an arbitrary but fixed $n \in \mathbb{N}$, consider the set of coefficients $C_{p,n}$ defined in (4.1) and a (finite) set of constrained inhomogeneous spherical equations

$$\begin{cases} \prod_{i=1}^{m} z_i^{-1} c_i z_i = c & \text{with } c_i, c \in C_{p,n} \\ z_i = (\overline{0}, 1) \text{ or } (\overline{0}, 2^{-1}). \end{cases} \tag{5.1}$$

The problem to find a solution of (5.1) is called **CISE**$_{\{1,2\}}$. Also, we want to study slightly relaxed equations

$$\begin{cases} \prod_{i=1}^{m} z_i^{-1} c_i z_i = c & \text{with } c_i, c \in C_{p,n} \\ z_i = (\overline{0}, 1), (\overline{0}, 2^{-1}), \text{ or } (\overline{0}, 3^{-1}). \end{cases} \tag{5.2}$$

The problem to find a solution of (5.2) is called **CISE**$_{\{1,2,3\}}$.

Randomized constrained inhomogeneous spherical equation problem. Find a solution for a uniformly distributed system (5.1) (or (5.2)).

Theorem 5.2. *Suppose that a PPT algorithm $\mathcal{A}$ solves the randomized $\mathbf{CISE}_{\{1,2\}}$ problem (or $\mathbf{CISE}_{\{1,2,3\}}$ problem) with parameters n, m, p satisfying (3.3) with probability at least n^{-c_0} for some fixed constant $c_0 > 0$, where the probability is taken over the choice of the instance as well as the coin-tosses of $\mathcal{A}$. Then there is a PPT algorithm that solves $\mathbf{SIVP}_{\gamma=pn^6}$ for any lattice with overwhelming probability (e.g. $1 - 2^{-n}$).*

Proof. Consider an arbitrary instance of $(A, \overline{y})$ of $\mathbf{ISIS}_{\{0,1\}}$. Let $\overline{v}_1, \ldots, \overline{v}_m$ be the system of columns of A. Consider the equation (4.3), where, as in Section 4.2, $c_{\overline{v}} = (\overline{v}, 1)$ and $\overline{v}' = \overline{y} + \sum_{i=1}^{m} \overline{v}_i$. Then for $\overline{x} = (x_1, \ldots, x_m) \in \{0,1\}^m$ we have

$$\overline{x} \text{ is a solution for } (A, \overline{y}) \quad \Leftrightarrow \quad \sum_{i=1}^{m} x_i \overline{v}_i = \overline{y}$$

$$\Leftrightarrow \quad \sum_{i=1}^{m} (x_i + 1)\overline{v}_i = \overline{y} + \sum_{i=1}^{m} \overline{v}_i$$

$$\Leftrightarrow \quad \prod_{i=1}^{m} (\overline{0}, (x_i + 1)^{-1})^{-1} c_{\overline{v}_i} (\overline{0}, (x_i + 1)^{-1}) = c_{\overline{v}'}$$

$$\Leftrightarrow \quad \{z_i = (\overline{0}, (x_i + 1)^{-1})\}_{i=1}^{m} \text{ is a solution for (5.1).}$$

This establishes a one-to-one correspondence between $\mathbf{ISIS}_{\{0,1\}}$ and $\mathbf{CISE}_{\{1,2\}}$ that translates the uniform distribution for the instances of $\mathbf{ISIS}_{\{0,1\}}$ to the uniform distribution for the instances of $\mathbf{CISE}_{\{1,2\}}$ (the element $\overline{v}' \in C_{p,n}$ is uniformly distributed because $\overline{y}$ uniformly distributed). A similar reduction can be used to establish a one-to-one correspondence between $\mathbf{SIS}_{\{-1,0,1\}}$ and $\mathbf{CISE}_{\{1,2,3\}}$. $\qquad\square$

Remark 1. *The one-to-one correspondence established in the proof of Theorem 5.2 also implies that uniformly chosen instances of $\mathbf{CISE}_{\{1,2\}}$ and $\mathbf{CISE}_{\{1,2,3\}}$ have solutions with probability approaching 1 as $n \to \infty$.*

Remark 2. *Similar results can be proven for the class of symmetric groups $\{S_n\}_{n=1}^{\infty}$, or any other class of groups that contain the groups $G_{p,n}$ in a "compact" way.*

6. Spherical functions

Let G be a group. For $\overline{c} = (c_1, \ldots, c_m) \in G^m$ define a function $f_{\overline{c}} \colon G^m \to G$ by

$$(z_1, \ldots, z_m) \quad \overset{f_{\overline{c}}}{\mapsto} \quad (z_1^{-1} c_1 z_1) \cdots (z_m^{-1} c_m z_m).$$

We call $f_{\overline{c}}$ a *spherical function* because the problem of finding an $f_{\overline{c}}$-preimage of $g \in G$ is equivalent to finding a solution for a spherical equation $(z_1^{-1} c_1 z_1) \cdots (z_m^{-1} c_m z_m) = g$. Fix distinct $g_0, g_1 \in G$ and define a function $H_{\overline{c}} \colon \{0,1\}^m \to G$ by

$$(b_1, \ldots, b_m) \quad \overset{H_{\overline{c}}}{\mapsto} \quad (g_{b_1}^{-1} c_1 g_{b_1}) \cdots (g_{b_m}^{-1} c_m g_{b_m}).$$

called a *0/1-spherical function* of length m. Similarly, we can define $-1/0/1$-spherical functions of length m.

    Consider the family of groups $\{G_{p,n}\}$ and $g_0 = (\overline{0}, 1), g_1 = (\overline{0}, 2^{-1}) \in G_{p,n}$, where $2^{-1} = \frac{1}{2}(p + 1)$ is the multiplicative inverse of 2 modulo an odd prime p. As in Section 5, n

is considered to be the main parameter; the parameters m and p are functions of n satisfying conditions (3.3). Define a system of functions

$$\mathcal{H}_n = \{\, H_{\bar{c}} \mid c_1, \ldots, c_m \in C_{p,n} \,\} \text{ and } \mathcal{H} = \bigcup_{n=1}^{\infty} \mathcal{H}_n.$$

Proposition 6.1. *If* $\mathbf{SIVP}_{\gamma=pn^6}$ *is hard in the worst case, then* $\mathcal{H}$ *is a one-way function family.*

Proof. To prove that $\mathcal{H}$ is a one-way function family it is sufficient to show that for every PPT algorithm $\mathcal{A}$ the sequence of values

$$P_{\mathcal{A}}(n) \;=\; \Pr[\mathcal{A}(1^n, \bar{c}, H_{\bar{c}}(\bar{x})) \in H_{\bar{c}}^{-1}(H_{\bar{c}}(\bar{x}))]$$

converges to 0 faster than every $1/poly(n)$, where the probability is taken over
- uniform choices of $\bar{c} \in (C_{p,n})^m$,
- uniform choices of $\bar{x} \in \{0,1\}^n$, and
- the coin-tosses of $\mathcal{A}$.

It is not difficult to check that for uniformly distributed $\bar{c}$ and $\bar{x}$, the values of $H_{\bar{c}}(\bar{x}) \in C_{p,n}$ are distributed nearly uniformly (the sequence of distributions on $C_{p,n}$ converges to the uniform distribution exponentially fast in terms of n) and we may assume that $\mathcal{A}$ deals with uniformly randomized $\mathbf{CISE}_{\{1,2\}}$.

Therefore, if it is not true that $P_{\mathcal{A}}(n)$ is $o(n^{-d})$ for some PPT algorithm $\mathcal{A}$ and some $d > 0$, then $P_{\mathcal{A}}(n) \geq c \cdot n^{-d}$ (for some $c > 0$) satisfied for infinitely many indices n. By Theorem 5.2 that means that there is a PPT algorithm solving $\mathbf{SIVP}_{\gamma=pn^6}$ with overwhelming probability on every lattice for infinitely many dimensions n, which contradicts the assumption (3.1). $\qquad\square$

Proposition 6.2. *If* $\mathbf{SIVP}_{\gamma=pn^6}$ *is hard in the worst case, then* $\mathcal{H}$ *is a collision-free function family.*

Proof. To prove that $\mathcal{H}$ is a collision-free hash function family it is sufficient to show that for every PPT algorithm $\mathcal{A}$ the sequence of values

$$P_{\mathcal{A}}(n) \;=\; \Pr[(\bar{x}, \bar{y}) = \mathcal{A}(1^n, \bar{c}) \ \& \ H_{\bar{c}}(\bar{x}) = H_{\bar{c}}(\bar{y}))]$$

converges to 0 faster than every $1/poly(n)$, where the probability is taken over
- uniform choices of $\bar{c} \in (C_{p,n})^m$,
- the coin tosses of $\mathcal{A}$.

If this condition is not satisfied, then there is a PPT algorithm $\mathcal{A}$ and $c, d > 0$ satisfying $P_{\mathcal{A}}(n) \geq c \cdot n^{-d}$ for infinitely many indices n. Observe that

$$
\begin{aligned}
H_{\bar{c}}(\bar{x}) = H_{\bar{c}}(\bar{y}) \quad &\Leftrightarrow \quad \sum_{j=1}^{m} x_j \bar{c}_j = \sum_{j=1}^{m} y_j \bar{c}_j \\
&\Leftrightarrow \quad \sum_{j=1}^{m} (2 + x_j - y_j)\bar{c}_j = \bar{0}, \quad \text{where } 2 + x_j - y_j \in \{1, 2, 3\} \\
&\Rightarrow \quad \{z_j = (\bar{0}, (2 + x_j - y_j)^{-1})\}_{j=1}^{m} \text{ satisfies (5.2).}
\end{aligned}
$$

Thus, if we can efficiently find collisions for uniformly distributed 0/1-spherical equations $H_{\bar{c}} \in \mathcal{H}_n$ for infinitely many indices n, then we can efficiently find solutions for uniformly

distributed $\mathbf{CISE}_{\{1,2,3\}}$. Then, by Theorem 5.2, there is a PPT algorithm solving $\mathbf{SIVP}_{\gamma=pn^6}$ with overwhelming probability on every lattice for infinitely many dimensions n, which contradicts the assumption (3.1). Contradiction. $\square$

7. The acyclic graph word problem: average-case hardness

CISE can be naturally reduced to different knapsack-type problems in groups. One such reduction is discussed in this section. The *acyclic graph word problem* was introduced in [12] as a convenient generalization of **SSP** and as a tool for studying **SSP**. Let X be a generating set for G.

The acyclic graph word problem, AGWP(G, X): Given an acyclic directed graph Γ with edges labeled by letters in $X \cup X^{-1} \cup \{\varepsilon\}$ with two marked vertices, α and ω, decide whether there is an oriented path in Γ from α to ω labeled by a word w such that $w = 1$ in G.

It is easy to show that complexity of **AGWP** does not depend on a choice of a generating set for G and the problem can be abbreviated **AGWP**(G). Also, we can work with edges labeled with words over the alphabet X. There is a natural polynomial-time reduction from (decision/search)-**CISE** to (decision/search)-**AGWP**. Indeed, for an instance of **CISE**

$$\begin{cases} \prod_{i=1}^{m} z_i^{-1} c_i z_i = c \\ z_i \in Z_i, \end{cases}$$

we can construct a labelled digraph $\Gamma = (V, E)$, where

- $V = \{0, 1, \ldots, m, m+1\}$
- $E = \left\{ j-1 \overset{z^{-1}c_j z}{\longrightarrow} j \ \middle|\ z \in Z_j, \ j = 1, \ldots, m \right\} \cup \{m \overset{c^{-1}}{\longrightarrow} m+1\}$
- the starting point α is 0,
- the terminal point ω is $m+1$.

The construction is visualized in Figure 1. By construction of Γ, the following statement holds.
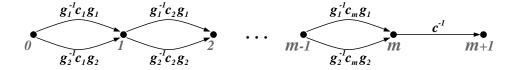

Figure 1: An instance of **AGWP** corresponding to an instance of **CISE** (assuming that $Z_j = \{g_1, g_2\}$).

Proposition 7.1. *The instance of* **CISE** *is positive if and only if the instance of* **AGWP** *is positive.*

The reduction described above induces a system of probability measures $\{\mu_n\}_{n\in\mathbb{N}}$ on instances of **AGWP** over the class of groups $\mathbb{Z}_p^n \rtimes \mathbb{Z}_p^*$. In fact, each μ_n gives uniform distribution on a subset of instances of **AGWP** of type shown in Figure 7.1. That establishes a one-to one correspondence between uniformly randomized **CISE**$_{\{1,2\}}$ described in Section 5 and **AGWP** endowed with $\{\mu_n\}_{n\in\mathbb{N}}$, which implies average-case hardness of **AGWP** (again, assuming (3.1)).

The converse reduction appears to be implausible. In fact, for some groups the converse is clearly not true, e.g., for $\mathbb{Z}_3^\omega$ (and for the class of groups $\{\mathbb{Z}_3^n\}_{n\in\mathbb{N}}$) we have

- **CISE** is polynomial-time decidable, but
- **SSP** (and hence **AGWP**) is **NP**-complete.

8. Open problems and questions for further study

In this section we outline several problems related to spherical equations, constrained spherical equations, relations between constrained/unconstrained problems, average-case complexity, and design of hash functions.

8.1. **Spherical equations over finite groups: constrained versus unconstrained.**

Problem 8.1. Does there exist a class of finite groups $\mathcal{G}$ satisfying the following conditions:
- the Diophantine problem for spherical equations over $\mathcal{G}$ is efficiently decidable,
- the Diophantine problem for constrained spherical equations is computationally hard?

Denote by $\mathbb{P}$ the set of all prime numbers. It was proved in [29] that the Diophantine problem for spherical equations over $\mathcal{G} = \{\mathrm{GL}(2,p)\}_{p\in\mathbb{P}}$ can be solved efficiently.

Problem 8.2. Does $\{\mathrm{GL}(2,p)\}_{p\in\mathbb{P}}$ satisfy the second condition of Problem 8.1?

Problem 8.2 can be generalized as follows.

Problem 8.3. Does any family of classical finite groups (such as $\mathrm{SL}(n,p)$, $\mathrm{PSL}(n,p)$, etc., see [23]) satisfy conditions of Problem 8.1?

8.2. **Constrained spherical equations: foundation of average-case hardness.** In Section 5 we showed that for groups $\{G_{p,n}\}$ constrained spherical equations (randomized in a certain way) are hard on average by establishing a one-to-one correspondence between **CISE**$_{\{1,2\}}$ and **ISIS**$_{\{0,1\}}$. Average-case hardness of **ISIS**$_{\{0,1\}}$ is linked (by M. Ajtai) to the worst-case hardness of lattice approximation problems, such as **SIVP**$_\gamma$, which makes hardness of **SIVP**$_\gamma$ the foundation of average hardness for **CISE**$_{\{1,2\}}$ and **ISIS**$_{\{0,1\}}$. Informally, is it possible to untie **CISE**$_{\{1,2\}}$ and **ISIS**$_{\{0,1\}}$ from **SIVP**$_\gamma$?

One way to approach this question is to utilize a discrete logarithm type self-reduction idea. For **ISIS**$_{\{0,1\}}$ that means to design a class of randomized self-reductions Φ between instances of **ISIS**$_{\{0,1\}}$ and use Φ to enhance **ISIS**-solvers as follows. If A is an **ISIS**-solver, then an enhanced **ISIS**-solver B for a given instance I performs the following:

(a) Apply A to I and if it succeeds, then output the result.
(b) Choose a random $\varphi \in \Phi$ and apply A to $\varphi(I)$. If it succeeds and produces a solution for $\varphi(I)$, then use it to "reconstruct" and output a solution for I.
(c) Repeat (b) until it succeeds.

If Φ is sufficiently large, then this approach might work. Notice that existence of uniform self-reductions for **ISIS** (like for the discrete logarithm problem) appears to be implausible.

Problem 8.4. Design a class of randomized self-reductions Φ for $\mathbf{ISIS}_{\{0,1\}}$ and use them to prove the following. If $\mathbf{ISIS}_{\{0,1\}}$ can be solved by a PPT algorithm A on a non-negligible set of instances, then it can be solved by a PPT algorithm B on every instance with overwhelming probability.

Similarly, it would be interesting to investigate self-reductions for **CISE** over some groups $\mathcal{G}$ and use them to demonstrate that the worst-case hardness for **CISE** implies its average-case hardness.

Problem 8.5. For a class of spherical equations over a class of groups $\mathcal{G}$, design a class of randomized self-reductions for **CISE** and use them to prove the following. If **CISE** can be solved by a PPT algorithm A on a non-negligible set of instances, then it can be solved by a PPT algorithm B on every instance with overwhelming probability.

8.3. **Constrained spherical equations: generalization.** As we have shown in Section 7, instances of **CISE** can be transformed into instances of **AGWP** of certain type (shown in Figure 7.1). Let us generalize the digraph shown in Figure 7.1 and modify it into a finite-state transducer. For a table of elements $C = \{c_{ij}\}$ from G (where $i = 1, \ldots, m$ and $j = 0, 1$) define a transducer Γ_C shown in Figure 2. Each Γ_C defines a function $J_C \colon \{0,1\}^m \to G$ as

Figure 2: The transducer Γ_C.

follows:

$$(b_1, \ldots, b_m) \quad \overset{J_C}{\mapsto} \quad c_{1b_1} \cdots c_{mb_m}.$$

By design, the problem of finding a J_C-preimage $(b_1, \ldots, b_m) \in \{0,1\}^m$ of a given element c generalizes **CISE**. Also, notice that the functions J_C generalize hash function construction introduced in [39] and its numerous variations (e.g., [38, 6]), including those utilizing monoids.

Problem 8.6. Investigate average-case hardness of computing J_C^{-1} for C sampled uniformly randomly.

The most promising approach to Problem 8.6 appears to be via self-reductions similar to those discussed in Section 8.2. Furthermore, it might be easier to design self-reductions for functions J_C because J_C does not require pairs c_{i0} and c_{i1} to be conjugate.

Problem 8.7. For a class of groups $\mathcal{G} = \{G_n\}$ design a randomized self-reduction for the problem of computing J_C^{-1} satisfying the following: if J_C^{-1} can be computed by a PPT algorithm on a non-negligible set of instances, then it can be computed by a PPT algorithm on every instance with overwhelming probability.

Problem 8.8. For a class of groups $\mathcal{G}$ investigate cryptographic properties for the family of functions $\{J_C\}_{C \in G^{2 \times m}}$.

8.4. Constrained spherical equations: learning without errors.

Consider a black-box device D that computes values of a "hidden" function $H_{\bar{c}} \colon \{0,1\}^m \to G$ for some $\bar{c} \in G^m$ hardwired inside of D. The device has a single button which, when pressed, produces a pair $(\bar{b}, g)$, where

- $\bar{b} = (b_1, \ldots, b_m) \in \{0,1\}^m$ is chosen uniformly randomly,
- $g = H_{\bar{c}}(\bar{b})$.

To learn the hidden function means to find the secret element $\bar{c}$ using a number of sampled pairs $(\bar{b}, g)$.

Problem 8.9. Investigate computational complexity of learning $\bar{c}$ from a hidden function $H_{\bar{c}}$ for different (classes of) finite/infinite groups.

In a similar way, we can consider a device D that computes values of a "hidden" function $J_C \colon \{0,1\}^m \to G$ and the problem of learning C.

Problem 8.10. Investigate computational complexity of learning C from a hidden function J_C for different (classes of) finite/infinite groups.

8.5. Spherical equations over infinite groups: decidability.

For what (infinite) groups G is the Diophantine problem for spherical equations decidable, but the problem for constrained spherical equations is not? Obviously any instance of **CISE** with a finite search space (e.g. when $|G| < \infty$ or when each variable can attain finitely many values) can be solved by enumerating all possible solutions. If $|G| = \infty$ and Z_j are allowed to be infinite, then **CISE** can become undecidable. This question depends on how we allow to constrain variables, e.g.

- sets Z_j can be finitely generated subgroups of G,
- or (more generally) sets Z_j can be defined by rational subsets of G.

Problem 8.11. For what (infinite) groups G the Diophantine problem for spherical equations is decidable, but spherical equations with constraints of the form $z_j \in Z_j$, where Z_j *is a finitely generated subgroup* of G is not decidable.

Obvious candidates are groups containing a subgroup H with undecidable membership problem, such as partially commutative groups. We believe that for these groups the Diophantine problem for spherical equations belongs to **NP**, but the problem for the constrained conjugacy equations

$$\begin{cases} z^{-1}cz = c \\ z \in H, \end{cases}$$

is not decidable. We also formulate a similar problem for rationally constrained spherical equations.

Problem 8.12. For what (infinite) groups G the Diophantine problem for spherical equations is decidable, but spherical equations with constraints of the form $z_j \in Z_j$, where Z_j *is a rational subset* of G, is not decidable.

Another class of groups with undecidable membership problem is the braid groups. For the braid groups it is not known how to approach even unconstrained spherical equations.

Problem 8.13. Is the Diophantine problem for spherical (quadratic) equations over braid groups decidable?

8.6. Spherical equations over infinite groups: universality.
Recall that a set of functions $H = \{h\colon D \to R\}$ is *universal* if for any distinct $x, y \in D$

$$\Pr[h(x) = h(y)] \leq \tfrac{1}{|R|},$$

where the probability is taken over a uniformly chosen $h \in H$. This condition can be relaxed by allowing $\Pr[h(x) = h(y)]$ to be $O(\tfrac{1}{|R|})$. Notice that in this paper for R we have used certain classes of finite groups only (namely, $\{\mathbb{Z}_p^n \rtimes \mathbb{Z}_p^*\}$ and $\{S_n\}$). Is it possible to construct a family of efficient universal functions using a single infinite group?

Problem 8.14. Design a family of universal $0/1$-spherical functions or J_C functions with the ranges R being a part of the same infinite group.

8.7. Spherical functions: public-key encryption.

Problem 8.15. Is it possible to design a public-key encryption scheme which security is based on computational hardness of solving constrained spherical equations (or another related group-theoretic problem)?

References

[1] D. Aggarwal, D. Dadush, O. Regev, and N. Stephens-Davidowitz. Solving the shortest vector problem in 2^n time using discrete gaussian sampling: Extended abstract. STOC '15, page 733–742, New York, NY, USA, 2015. Association for Computing Machinery.

[2] D. Aharonov and O. Regev. Lattice Problems in **NP** ∩ **coNP**. *J. ACM*, 52(5):749–765, 2005.

[3] M. Ajtai. Generating hard instances of lattice problems (extended abstract). In *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '96, pages 99–108, New York, NY, USA, 1996. Association for Computing Machinery.

[4] M. Ajtai, R. Kumar, and D. Sivakumar. A sieve algorithm for the shortest lattice vector problem. In *Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing*, STOC '01, page 601–610, New York, NY, USA, 2001. Association for Computing Machinery.

[5] L. Bartholdi, T. Groth, and I. Lysenok. Commutator width in the first Grigorchuk group. *Groups Geom. Dyn.*, 16:493–522, 2022.

[6] L. Bromberg, V. Shpilrain, and A. Vdovina. Navigating in the Cayley graph of $\mathrm{SL}_2(\mathbb{F}_p)$ and applications to hashing. *Semigroup Forum*, 94, 2015.

[7] L. Ciobanu. Word equations, constraints, and formal languages. preprint. Available at `http://arxiv.org/abs/2406.02160`.

[8] L. Comerford and C. Edmunds. Quadratic equations over free groups and free products. *J. Algebra*, 68:276–297, 1981.

[9] M. Culler. Using surfaces to solve equations in free groups. *Topology*, 20(2):133–145, 1981.

[10] V. Diekert and J. Robson. *Quadratic word equations*, pages 314–326. Springer Berlin Heidelberg, Berlin, Heidelberg, 1999.

[11] A. Földvári and G. Horváth. The complexity of the equation solvability and equivalence problems over finite groups. *Int. J. Algebra Comput.*, 30(03):607–623, 2020.

[12] L. Frenkel, A. Nikolaev, and A. Ushakov. Knapsack problems in products of groups. *J. Symbolic Comput.*, 74:96–108, 2016.

[13] M. Goldmann and A. Russell. The complexity of solving equations over finite groups. *Inf. Comput.*, 178(1):253–262, 2002.

[14] O. Goldreich and S. Goldwasser. On the limits of nonapproximability of lattice problems. *J. Comput. Syst. Sci.*, 60(3):540–563, jun 2000.

[15] O. Goldreich, S. Goldwasser, and S. Halevi. Public-key cryptosystems from lattice reduction problems. In *Advances in Cryptology – CRYPTO 1997*, volume 1294 of *Lecture Notes Comp. Sc.*, pages 112–131, London, UK, 1997. Springer-Verlag.

[16] R. Grigorchuk and P. Kurchanov. On quadratic equations in free groups. In *Proc. Int. Conf. on Algebra Dedicated to the Memory of A. I. Malcev*, volume 131 of *Contemporary Mathematics*, pages 159–171. American Mathematical Society, 1992.

[17] R. Grigorchuk and I. Lysenok. A description of solutions of quadratic equations in hyperbolic groups. *Int. J. Algebra Comput.*, 2(3):237–274, 1992.

[18] P. Idziak, P. Kawalek, J. Krzaczkowski, and A. Weiß. Satisfiability problems for finite groups. In *ICALP 2022, Proc.*, volume 229 of *LIPIcs*, pages 127:1–127:20, 2022.

[19] R. Kannan. Improved algorithms for integer programming and related lattice problems. In *Proceedings of the Fifteenth Annual ACM Symposium on Theory of Computing*, STOC '83, page 193–206, New York, NY, USA, 1983. Association for Computing Machinery.

[20] O. Kharlampovich, I. Lysenok, A. Miasnikov, and N. Touikan. The solvability problem for quadratic equations over free groups is NP-complete. *Theor. Comput. Syst.*, 47:250–258, 2010.

[21] O. Kharlampovich, A. Mohajeri, A. Taam, and A. Vdovina. Quadratic equations in hyperbolic groups are NP-complete. *Transactions of the American Mathematical Society*, 369(9):6207–6238, 2017.

[22] O. Kharlampovich and A. Myasnikov. Irreducible affine varieties over a free group. I: Irreducibility of quadratic equations and Nullstellensatz. *J. Algebra*, 200(2):472–516, 1998.

[23] P. Kleidman and M. Liebeck. *The Subgroup Structure of the Finite Classical Groups*. London Mathematical Society Lecture Note Series. Cambridge University Press, 1990.

[24] A. Lenstra, H. Lenstra, and L. Lovász. Factoring polynomials with rational coefficients. *Math. Ann.*, 261:1432–1807, 1982.

[25] I. Lysenok, A. Miasnikov, and A. Ushakov. Quadratic equations in the Grigorchuk group. *Groups, Geometry, and Dynamics*, 10:201–239, 2016.

[26] I. Lysenok and A. Ushakov. Spherical quadratic equations in free metabelian groups. *Proc. Amer. Math. Soc.*, 144:1383–1390, 2016.

[27] I. Lysenok and A. Ushakov. Orientable quadratic equations in free metabelian groups. *Journal of Algebra*, 581:303–326, 2021.

[28] R. Mandel and A. Ushakov. Quadratic equations in metabelian Baumslag-Solitar groups. Accepted to International Journal of Algebra and Computation. Available at `https://arxiv.org/abs/2302.06974`, 2023.

[29] C. Mattes, A. Ushakov, and A. Weiss. Complexity of spherical equations in finite groups. In *SOFSEM 2024: Theory and Practice of Computer Science*, volume 14519 of *Lecture Notes Comp. Sc.*, pages 383–397. Springer, 2024.

[30] A. G. Miasnikov, A. Nikolaev, and A. Ushakov. Knapsack problems in groups. *Math. Comput.*, 84:987–1016, 2015.

[31] A. G. Miasnikov, V. Shpilrain, and A. Ushakov. *Non-Commutative Cryptography and Complexity of Group-Theoretic Problems*. Mathematical Surveys and Monographs. AMS, 2011.

[32] D. Micciancio and O. Regev. Worst-Case to Average-Case Reductions Based on Gaussian Measures. *SIAM Journal on Computing*, 37(1):267–302, 2007.

[33] D. Micciancio and P. Voulgaris. A deterministic single exponential time algorithm for most lattice problems based on voronoi cell computations. In *Proceedings of the Forty-Second ACM Symposium on Theory of Computing*, STOC '10, pages 351–358, New York, NY, USA, 2010. Association for Computing Machinery.

[34] A. G. Myasnikov, A. Nikolaev, and A. Ushakov. The Post correspondence problem in groups. *J. Group Theory*, 17:991–1008, 2014.

[35] A. Nikolaev and A. Ushakov. On subset sum problem in branch groups. *Groups, Complexity, Cryptology*, 12, 2020.

[36] C. Peikert. A Decade of Lattice Cryptography. *Found. Trends Theor. Comput. Sci.*, 10:283–424, 2016.

[37] P. E. Schupp. Quadratic equations in groups, cancellation diagrams on compact surfaces, and automorphisms of surface groups. In *Word problems, II (Conf. on Decision Problems in Algebra, Oxford, 1976)*, volume 95 of *Stud. Logic Foundations Math.*, pages 347–371. North-Holland, Amsterdam, 1980.

[38] J. Tillich and G. Zémor. Hashing with sl_2. In *Advances in Cryptology – CRYPTO 1994*, volume 839 of *Lecture Notes Comp. Sc.*, pages 40–49. Springer, 1994.

[39] G. Zémor. Hash Functions And Graphs With Large Girths. In *Advances in Cryptology – EUROCRYPT 1991*, volume 547 of *Lecture Notes Comp. Sc.*, pages 508–511, Berlin, 1991. Springer.