Electronic Colloquium on Computational Complexity, Report No. 189 (2010)

On the Sum of Square Roots of Polynomials and related problems

Neeraj Kayal [*] Chandan Saha [†]

December 8, 2010

Abstract

The sum of square roots problem over integers is the task of deciding the sign of a nonzero sum, $S = \sum_{i=1}^{n} \delta_i \cdot \sqrt{a_i}$, where $\delta_i \in \{+1, -1\}$ and a_i's are positive integers that are upper bounded by N (say). A fundamental open question in numerical analysis and computational geometry is whether $|S| \geq 1/2^{(n \cdot \log N)^{O(1)}}$. We study a formulation of this problem over polynomials: Given an expression $S = \sum_{i=1}^{n} c_i \cdot \sqrt{f_i(x)}$, where c_i's belong to a field of characteristic 0 and f_i's are univariate polynomials with degree bounded by d and $f_i(0) \neq 0$ for all i, is it true that the minimum exponent of x which has a nonzero coefficient in the power series S is upper bounded by $(n \cdot d)^{O(1)}$, unless $S = 0$? We answer this question affirmatively. Further, we show that this result over polynomials can be used to settle (positively) the sum of square roots problem for a special class of integers: Suppose each integer a_i is of the form, $a_i = X^{d_i} + b_{i1}X^{d_i - 1} + \ldots + b_{id_i}$, $d_i > 0$, where X is a positive real number and b_{ij}'s are integers. Let $B = \max_{i,j}\{|b_{ij}|\}$ and $d = \max_i\{d_i\}$. If $X > (B+1)^{(n \cdot d)^{O(1)}}$ then a nonzero $S = \sum_{i=1}^{n} \delta_i \cdot \sqrt{a_i}$ is lower bounded as $|S| \geq 1/X^{(n \cdot d)^{O(1)}}$.

We then consider the following more general problem: given an arithmetic circuit computing a multivariate polynomial $f(\mathbf{X})$ and integer d, is the degree of $f(\mathbf{X})$ less than or equal to d? We give a $\mathsf{coRP}^{\mathsf{PP}}$-algorithm for this problem, improving previous results of [ABKPM09] and [KP07].

[*] Microsoft Research India, `neeraka@microsoft.com`
[†] Max Planck Institute for Informatics, `csaha@mpi-inf.mpg.de`

ISSN 1433-8092

1 Introduction

The sum of square roots is the following well-known problem: given a set $\{a_1, a_2, \ldots, a_n\}$ of positive integers and $\{\delta_1, \delta_2, \ldots, \delta_n\} \in \{-1, +1\}^n$, determine the sign of the sum

$$\sum_{i=1}^{n} \delta_i \cdot \sqrt{a_i} \tag{1}$$

It was posed as an open problem by Garey, Graham and Johnson [GGJ76] in connection with the Euclidean travelling salesman problem. Euclidean TSP is not known to be in NP but is easily seen to be in NP relative to the sum of square roots problem. More generally, the sum of square roots problem is of importance for many problems in computational geometry (cf. [MR08]), since the computation of the Euclidean distance between two points in general case involves the computation of a square root. The sum of square roots problem also arises in the algorithmic solution of semidefinite progamming which in turn is used for designing approximation algorithms (cf. the survey by Goemans [Goe98]). Although it has been conjectured [Mal96] that the problem lies in P, the best known result so far [ABKPM09] is containment in the counting hierarchy CH, which is a subclass of PSPACE that contains the polynomial hierarchy PH. For the related but easier problem of determining whether the sum (1) is zero or not, a deterministic polynomial-time algorithm is known [Blö91] [1]. A possible approach towards answering this question is to solve the following number-theoretic problem whose current status is still a conjecture.

Problem 1.1 (Lower bounding a nonzero 'signed' sum of square root of integers)**.** Given a sum $S = \sum_{i=1}^{n} \delta_i \cdot \sqrt{a_i}$, where $\delta_i \in \{+1, -1\}$ and a_i's are positive integers upper bounded by N, find a tight lower bound on $|S|$ in terms of n and N when $S \neq 0$. Is it true that for a nonzero S, $|S| \geq 1/2^{\mathsf{poly}(n, \log N)}$ for some fixed polynomial $\mathsf{poly}(\cdot)$?

If the answer to the above question is yes then computing the square roots up to $\mathsf{poly}(n, \log N)$ precision suffices to determine the sign of the sum of square roots. Reducing the immense gap between the known upper and lower bounds for Problem 1.1 is a challenging number-theoretic problem.

Now, there is a well known analogy between integers and polynomials (cf. [EHM05]). We refer the reader to a survey by Landau and Immerman [IL93] for some algorithmic aspects of this analogy. The investigation of the complexity of polynomial analogs of integer problems has occassionally given important insight into the integer problem itself. Indeed, Allender et al. [ABKPM09] proved a hardness result of a closely related problem, which they call BitSLP, by first observing that the corresponding problem for polynomials is $\#\mathsf{P}$-hard. This motivates us to examine the polynomial analogue of the sum of square roots problem as an interesting problem in its own right.

Here we study the natural analogue of Problem 1.1 in the world of polynomials, the precise statement of which is given below.

Problem 1.2 (Sum of square root of polynomials)**.** Given an expression $S = \sum_{i=1}^{n} c_i \sqrt{f_i(x)}$, where $c_i \in \mathbb{F}$ (a field of characteristic 0) and $f_i(x)$ are univariate polynomials with degree bounded by d and $f_i(0) \neq 0$ (for all i), [2] can we show that unless $S = 0$, the minimum exponent of x which has a nonzero coefficient in the power series S is bounded by a fixed polynomial in n and d.

[1] The conference version of Blömer's paper [Blö91] contains a randomized polynomial time algorithm, which was later derandomized in [Blö93]

[2] The condition $f_i(0) \neq 0$ is simply to ensure that $\sqrt{f(x)}$ has a well defined power series expansion around $x = 0$.

This problem being a close cousin of Problem 1.1, it seems reasonable to hope that solving it might shed some light on the latter problem. At the least, one might expect to solve Problem 1.1 for a nontrivial class of integers starting from a solution to Problem 1.2. We are not aware of any prior research work along this line. In this work, we answer the question posed in Problem 1.2 in the affirmative. Using this result we show that it is indeed sufficient to keep polynomial amount of precision in computing the sign of S in Problem 1.1 if the input integers belong to a special class that we call (by abusing terminology) the *polynomial integers*.

We have mentioned earlier that the sum of square roots problem lies in the counting hierarchy CH. This result is due to Allender, Bürgisser, Kjeldgaard-Pedersen and Miltersen [ABKPM09]. In fact, they showed that the more general problem PosSLP, which is the task of checking if the integer produced by a given division-free straight-line program is greater than zero, belongs to the complexity class $\mathsf{P}^{\mathsf{PP}^{\mathsf{PP}^{\mathsf{PP}}}}$ that is contained in the fourth level of CH. The polynomial analog of the PosSLP problem is the task of comparing the degree of the polynomial computed by a given arithmetic circuit with a given integer. More precisely:

Problem 1.3 (Degree Computation)**.** Let $\mathbb{F}$ be a field (say the rational numbers $\mathbb{Q}$). Given an arithmetic circuit computing a multivariate polynomial $f(\mathbf{X})$ over $\mathbb{F}$ and an integer d, is the degree of $f(\mathbf{X})$ at most d?

This problem, which Allender et al. [ABKPM09] refer to as DegSLP, was also studied by Koiran and Perifel [KP07] and they put it in the second level of the counting hierarchy. Here we give a (slight) improvement to the complexity theoretic upper bound for DegSLP. We show its containment in the class $\mathsf{coRP}^{\mathsf{PP}}$.

1.1 Previous work

The work of Burnikel, Fleischer, Mehlhorn and Schirra [BFMS00] considered the problem of finding the sign of an arithmetic expression E involving the operations additions, subtractions, multiplications and square root (in fact, division as well), and with integer operands. They showed that if u is the bound on the value of E when all the subtraction operations in E are replaced by additions and k is the number of distinct square root operations in E then $|E| \geq 1/u^{2^k-1}$ unless $E = 0$. This result immediately gives an exponential bound on the bit size of S in Problem 1.1: if $S \neq 0$ then $|S| \geq 1/2^{2^n \cdot \log(nN)}$. (In this regard, the work of Mehlhorn and Schirra [MS00] is also relevant). It is also noted in [BFMS00] that the bound obtained for E is nearly optimal in general. For instance, if $E = (2^{2^k}+1)^{1/2^k} - 2$ then $u = (2^{2^k}+1)^{1/2^k} + 2 \leq 5$ and hence by the result in [BFMS00], $|E| \geq 1/5^{2^k-1}$. On the other hand, it was also shown that $|E| \leq 1/2^{2^k}$. However, Problem 1.1 is just a special case of the problem studied in [BFMS00] where there is no occurrence of nested square roots. So, it remains a conceivable possibility that there is a better lower bound for $|S|$. Indeed, for a certain choice of parameters a better result is known due to the work of Cheng, Meng, Sun and Chen [CMSC10] (see also [Che06]). By connecting Problem 1.1 to the shortest vector problem over a certain integer lattice, they showed that $|S| \geq 1/N^{2^{O(N/\log N)}}$, which is an improvement over earlier results when $N \leq c \cdot n \cdot \log n$ for some constant c. However, a less desirable aspect of this result is the doubly exponential dependency of the bit complexity of S on $\log N$.

On the other side, one seeks to construct/prove the existence of sets of integers $\{a_1, a_2, \ldots, a_n\}$ for which the sum S of Problem 1.1 is as small as possible in absolute value. In other words, what could be a good upper bound on $|S|$. Qian and Wang [QW06] gave an explicit construction that gives $|S| = O(N^{-2n+3/2})$. The integers they construct are closely related to the special class of integers that we look at (see Section 1.2). For Qian and Wang, every integer a_i is essentially of the form $(X+i)$ (suitably scaled).

1.2 Our contribution

Our first contribution is an affirmative answer to Problem 1.2. Using this, we prove the following theorem.

Theorem 1.4 (Sum of square root of 'polynomial integers'). *Suppose $S = \sum_{i=1}^{n} \delta_i \sqrt{a_i}$ ($\delta_i \in \{+1, -1\}$) such that every positive integer a_i is of the form $a_i = X^{d_i} + b_{i1} \cdot X^{d_i - 1} + \ldots + b_{id_i}$ ($d_i > 0$), where X is a positive real number and b_{ij} are integers. Let $B = \max_{i,j}\{|b_{ij}|\}$ and $d = \max_i\{d_i\}$. If $X > (B+1)^{p_1(n,d)}$, where $p_1(n,d)$ is a fixed polynomial in n and d, then a nonzero S is lower bounded as, $|S| \geq 1/X^{p_2(n,d)}$, where $p_2(n,d)$ is another fixed polynomial in n and d.*

The polynomials $p_1(n,d)$ and $p_2(n,d)$ can be taken to be $12 \cdot dn^2 \log 2d$ and $8 \cdot dn^2$, respectively. Note that the integers b_{ij} need not be positive.

Expressing each a_i as $X^{d_i} + b_{i1} \cdot X^{d_i-1} + \ldots + b_{id_i}$ is nothing very unusual - it is like a base-X representation of a_i when X is a positive integer. What makes the 'polynomial integers' special is the condition that X is exponentially large compared to the b_{ij}'s; or in other words, all the digits are small in X-ary representation. Indeed, if one can prove Theorem 1.4 without this condition then Problem 1.1 would stand solved in its full generality by taking $X = 2$.

Finally, we would like to note that we have not made an attempt to find the best possible expressions for $p_1(\cdot)$ and $p_2(\cdot)$, our primary intention being to just show that the functions p_1, p_2 are some fixed polynomials in n and d.

For the more general DegSLP problem, we show containment in the first level of the counting hierarchy (modulo the use of randomization), thereby improving the previous best result [KP07] for this problem which was the second level of the counting hierarchy. More precisely, we show

Theorem 1.5. DegSLP *is in* $\mathsf{coRP}^{\mathsf{PP}}$.

Organization - The rest of this paper is organized as follows. In Section 2 we give a solution to Problem 1.2 and in Section 3 we prove Theorem 1.4. The result on the complexity upper bound of DegSLP is presented in Section 4.

2 Sum of square roots of polynomials

In this section, we prove the following theorem.

Theorem 2.1. *Given a sum $S = \sum_{i=1}^{n} c_i \cdot g_i(x) \cdot \sqrt{f_i(x)}$ where $c_i \in \mathbb{F}$ (a field of characteristic 0), f_i and g_i are univariate polynomials of degree at most d and $f_i(0) \neq 0$ for all $1 \leq i \leq n$, either $S = 0$ or the minimum exponent of x which has a nonzero coefficient in the power series S is bounded by $dn^2 + n$.*

The solution to Problem 1.2 follows immediately if we take $g_i(x)$ to be 1 in the above theorem. But, we will need the slightly general form, that is, when the g_i's are not assumed to be 1, to prove

Theorem 1.4. We will see that the proof of Theorem 2.1 is not difficult in hindsight - it uses a mathematical object called the Wronskian, which is used to study linear dependence of functions. Let us spend some time to briefly discuss this concept.

2.1 The Wronskian and linear independence

A set of functions n functions $\{h_1, h_2, \ldots, h_n\}$ over a field $\mathbb{F}$ is said to be linearly dependent if there exist elements $c_1, c_2, \ldots, c_n \in \mathbb{F}$ such that the function $(c_1 h_1 + c_2 h_2 + \ldots + c_n h_n)$ is identically zero. If each of the h_i's is n times differentiable, then the *Wronskian* of this set, denoted $\mathsf{W}(h_1, \ldots, h_n)$ (or, $\mathsf{W}(\mathbf{h})$ for short) is defined as the following determinant.

$$\mathsf{W}(h_1, \ldots, h_n) \stackrel{\text{def}}{=} \det \begin{pmatrix} h_1 & h_2 & \ldots & h_n \\ h_1^{(1)} & h_2^{(1)} & \ldots & h_n^{(1)} \\ \vdots & \vdots & \vdots & \vdots \\ h_1^{(n-1)} & h_2^{(n-1)} & \ldots & h_n^{(n-1)} \end{pmatrix},$$

where $h_i^{(j)}$ is the j^{th} derivative of h_i. It is a well known function used in the study of differential equations. It is easy to observe that if the functions $h_1, \ldots, h_n$ are $\mathbb{F}$-linearly dependent then their Wronskian is identically zero. But, the converse need not be true in general. Bôcher [BÔ0] showed that there are families of infinitely differentiable functions which are linearly independent and yet their Wronskian vanishes identically. However, for analytic functions this is not the case: a finite family of linearly independent (real or complex valued) analytic functions has a nonzero Wronskian. More generally, this property is true for any family of formal power series over any characteristic zero field.

Theorem 2.2 (Wronskian of a family of power series). *Let $\mathbb{F}$ be a field of characteristic zero. A finite family of power series in $\mathbb{F}[[x]]$ has a zero Wronskian if and only if it is $\mathbb{F}$-linearly dependent.*

A short and simple proof of the above fact appears in [BD10]. Let us now see how to use this result to prove Theorem 2.1.

2.2 Proof of Theorem 2.1

Let $S = \sum_{i=1}^{n} c_i \cdot g_i(x) \cdot \sqrt{f_i(x)}$ be a given nonzero sum. Assume without loss of generality that $f_i(0) = 1$, for all i. If this is not the case then take out $\sqrt{f_i(0)}$ common from the term $\sqrt{f_i(x)}$ and work with an appropriate extension of $\mathbb{F}$ that contains $\sqrt{f_i(0)}$. This is simply to ensure that $\sqrt{f_i}$ can be expressed as a formal power series in x over $\mathbb{F}$. Denote $g_i\sqrt{f_i}$ by h_i. We can also assume that $h_1, \ldots, h_n$ are $\mathbb{F}$-linearly independent - if not, simply work with an $\mathbb{F}$-basis of $h_1, \ldots, h_n$. (Note that, we are not finding a basis, we are only using it for the sake of argument.)

Suppose, x^t divides the power series S, where t is the maximum possible. Pretend that,

$$\sum_{i=1}^{n} c_i h_i = S \tag{2}$$

is a linear equation in the 'variables' $c_1, \ldots, c_n$. By taking derivatives of both sides of Equation 2 with respect to x, we have the following system of linear equations in $c_1, \ldots, c_n$, for $0 \le j \le n-1$,

$$\sum_{i=1}^{n} c_i h_i^{(j)} = S^{(j)}, \tag{3}$$

where $h_i^{(j)}$ and $S^{(j)}$ are the j^{th} derivatives of h_i and S, respectively. Let $\mathcal{C}$ be the coefficient matrix of the above system of linear equations. That is,

$$\mathcal{C} = \begin{pmatrix} h_1 & h_2 & \dots & h_n \\ h_1^{(1)} & h_2^{(1)} & \dots & h_n^{(1)} \\ \vdots & \vdots & \vdots & \vdots \\ h_1^{(n-1)} & h_2^{(n-1)} & \dots & h_n^{(n-1)} \end{pmatrix}.$$

Observe that $\det(\mathcal{C})$ is the Wronskian $\mathsf{W}(\mathbf{h})$. The following simple claim about $\mathsf{W}(\mathbf{h})$ is crucial to the proof.

Claim 2.3. *The Wronskian* $\mathsf{W}(\mathbf{h}) = \prod_{i=1}^{n} f_i^{-\frac{2n-3}{2}} \cdot \det(\mathsf{M})$, *where* M *is an* $n \times n$ *matrix whose every entry is a polynomial in* x *of degree at most* $n \cdot d$.

Proof. Expanding $h^{(j)}$ we get the following. (Superscripts indicate the order of the derivatives.)

$$h_i^{(j)} = \sum_{k=0}^{j} g_i^{(j-k)} (\sqrt{f_i})^{(k)} \Rightarrow f_i^{\frac{2j-1}{2}} \cdot h_i^{(j)} = \sum_{k=0}^{j} g_i^{(j-k)} \cdot f_i^{\frac{2j-1}{2}} \cdot (\sqrt{f_i})^{(k)},$$

multiplying both sides by $f_i^{2j-1/2}$. Now notice that, $f_i^{2j-1/2} \cdot (\sqrt{f_i})^{(k)}$ is a polynomial of degree at most $j \cdot d$. Hence, $f_i^{2j-1/2} \cdot h_i^{(j)}$ is also a polynomial of degree at most $(j+1) \cdot d$, although individually they are power series in x. Since j is at max $n-1$, the statement of the claim follows. $\square$

Since $S \neq 0$, there must be one c_i which is nonzero. Let it be c_1. Then, by applying Cramer's rule,

$$c_1 = \frac{\det(\mathsf{M}_1)}{\mathsf{W}(\mathbf{h})} = \prod_{i=1}^{n} f_i^{\frac{2n-3}{2}} \cdot \frac{\det(\mathsf{M}_1)}{\det(\mathsf{M})} \quad \text{(by Claim 2.3)},$$

where M_1 is the following matrix,

$$\mathsf{M}_1 = \begin{pmatrix} S & h_2 & \dots & h_n \\ S^{(1)} & h_2^{(1)} & \dots & h_n^{(1)} \\ \vdots & \vdots & \vdots & \vdots \\ S^{(n-1)} & h_2^{(n-1)} & \dots & h_n^{(n-1)} \end{pmatrix}.$$

Note that, Cramer's rule applies here because $\mathsf{W}(\mathbf{h}) \neq 0$, as $h_1, \dots, h_n$ are assumed to be linearly independent, which in turn implies that $\det(\mathsf{M}) \neq 0$ (by Claim 2.3). Since x^t divides S, x^{t-n+1} must divide $S^{(j)}$ for every $0 \leq j \leq n-1$ and hence x^{t-n+1} divides $\det(\mathsf{M}_1)$.

Claim 2.4. *The maximum power of x dividing* $\det(\mathsf{M}_1)$ *and* $\det(\mathsf{M})$ *must be the same.*

Proof. This is because c_1 is an element of the field $\mathbb{F}$ and $\prod_{i=1}^{n} f_i(0)^{\frac{2n-3}{2}} \neq 0$ by assumption. $\square$

Therefore, $t - n$ must be less than the degree of $\det(\mathsf{M})$, which is at most $d \cdot n^2$ (again by Claim 2.3), and hence $t \leq d \cdot n^2 + n$. This proves Theorem 2.1.

With the polynomial version of the sum of square roots problem at hand, one wonders as to what can be inferred about the corresponding problem over integers. In turns out that indeed something nontrivial can be shown about a special class of integers that we have called before as the 'polynomial integers' (see Theorem 1.4). This constitutes the content of the following section.

3 Sum of square roots of 'polynomial integers'

This section is devoted to the proof of Theorem 1.4. Let $S = \sum_{i=1}^{n} \delta_i \sqrt{a_i}$; $\delta_i \in \{+1, -1\}$, be a given nonzero sum, where each positive integer a_i is of the following form.

$$a_i \quad = \quad X^{d_i} + b_{i1}X^{d_i - 1} + \ldots + b_{id_i} \tag{4}$$

where X is a positive real number and b_{ij}'s are integers (not necessarily positive).

Overview. The overall idea of the proof is to do a Taylor series expansion for each $\sqrt{a_i}$ so that we get a Taylor expansion for the sum S overall. Using Theorem 2.1 and the nonzeroness of S, we deduce that we must get a nonzero term 'very early' in the Taylor expansion. That is, there must be some nonzero S_ℓ for ℓ 'relatively small'. We use the fact that ℓ is small to deduce that such an S_ℓ is 'fairly large' in absolute value. We then use the fact that each b_{ij} is much smaller than X to upper bound each of the remaining terms and thereby deduce that the sum of the remaining terms *cannot* almost cancel out S_ℓ. More specifically, the sum of the remaining terms is at most $\frac{1}{2}|S_\ell|$ in absolute value. This helps us deduce that S itself is fairly large in absolute value.

Doing the Taylor expansion. From (4) we have

$$\sqrt{a_i} \quad = \quad (\sqrt{X})^{d_i} \cdot \sqrt{1 + \frac{b_{i1}}{X} + \ldots + \frac{b_{id_i}}{X^{d_i}}}.$$

Adding these expressions together with the appropriate sign, we get an expression for S.

$$S \quad = \quad \sum_{i=1}^{n} \delta_i \cdot X^{d_i/2} \cdot \sqrt{1 + \frac{b_{i1}}{X} + \ldots + \frac{b_{id_i}}{X^{d_i}}}.$$

Let $y = 1/X$ and $d = \max_i\{d_i\}$. Then,

$$S \cdot y^{d/2} \quad = \quad \sum_{i=1}^{n} \delta_i \cdot y^{(d-d_i)/2} \cdot \sqrt{1 + (b_{i1} \cdot y) + \ldots + (b_{id_i} \cdot y^{d_i})}.$$

Now notice that, by pretending that $f_i(y) = 1 + b_{i1}y + \ldots + b_{id_i}y^{d_i}$ is a polynomial in the 'formal variable' y, the sum $S \cdot y^{d/2}$ is of the form $\sum_{i=1}^{n} \delta_i \cdot g_i(y) \cdot \sqrt{f_i(y)}$, where $g_i(y) = y^{(d-d_i)/2}$. Therefore,

$$S \cdot y^{d/2} = \sum_{i=1}^{n} \delta_i \cdot \sum_{j \geq 0} c_{ij} \cdot y^j = \sum_{j \geq 0} y^j \cdot \sum_{i=1}^{n} \delta_i \cdot c_{ij} \quad \text{(by exchanging the summations)},$$

where c_{ij} is the coefficient of y^j coming from the i^{th} power series $g_i(y) \cdot \sqrt{f_i(y)}$. This is the Taylor series expansion of S that we will work with. We give the name S_j to each summand in the expression above, namely $S_j = y^j \cdot \sum_{i=1}^{n} \delta_i \cdot c_{ij}$. Thus

$$S \cdot y^{d/2} = \sum_{j \geq 0} S_j. \tag{5}$$

The Proof Strategy ahead. Applying Theorem 2.1, the minimum exponent ℓ of y with $\sum_{i=1}^{n} \delta_i \cdot c_{i\ell} \neq 0$ is such that $\ell \leq dn^2 + n$. Suppose we could show that

$$\frac{|S_{\ell+t}|}{|S_\ell|} \leq \frac{1}{2^{t+1}}, \tag{6}$$

for every $t \geq 1$, then from (5) it would follow that

$$\begin{aligned} |S| \cdot \left|y^{d/2}\right| &\geq ||S_\ell| - |S_{\ell+1}| - \ldots| \geq \left||S_\ell| - \frac{1}{2}|S_\ell|\right| = \frac{1}{2} \cdot |S_\ell| \\ \Rightarrow |S| &\geq \left|X^{d/2}\right| \cdot \frac{1}{2} \cdot |S_\ell| . \end{aligned} \tag{7}$$

This (potentially) gives us a lower bound on $|S|$ via a lower bound of $|S_\ell|$. But, to satisfy the condition given by equation (6), we also need an upper bound on $|S_{\ell+t}|$ for every t.

Upper bound on $|S_j|$:

$$|S_j| = y^j \cdot \left|\sum_{i=1}^{n} \delta_i c_{ij}\right| \leq n \cdot y^j \cdot \max_i\{|c_{ij}|\}$$

Let us upper bound the quantity $|c_{ij}|$. Fix any index i. For the ease of presentation, we will avoid writing the index i whenever it is clear from the context that we have a specific i in mind. For example, we write c_j as the coefficient of y^j coming from the power series,

$$y^{\frac{d-d_i}{2}} \cdot \sqrt{1 + b_1 y + \ldots + b_{d_i} y^{d_i}} = y^{\frac{d-d_i}{2}} \cdot \sum_{k=0}^{\infty} u_k \cdot (b_1 y + \ldots + b_{d_i} y^{d_i})^k,$$

where $u_k = \frac{\frac{1}{2}\cdot(\frac{1}{2}-1)\ldots(\frac{1}{2}-(k-1))}{k!} = (-1)^k \cdot \frac{1\cdot 3\cdot 5\ldots(2(k-1)-1)}{2^k\cdot k!}$. Expressed differently,

$$u_k = (-1)^{k+1} \cdot \frac{(2k)!}{(2k-1)\cdot(k!)^2 \cdot 2^{2k}} \Rightarrow |u_k| = \frac{\binom{2k}{k}}{(2k-1)\cdot 2^{2k}}$$

Now, notice that $\binom{2k}{k}/(2k-1) = 2 \cdot C_{k-1}$, where C_k is the k^{th} Catalan number $1/(k+1) \cdot \binom{2k}{k}$. Hence,

$$|u_k| = \frac{2 \cdot C_{k-1}}{2^{2k}}, \text{ and also } |u_k| \leq 1. \tag{8}$$

For any j, the coefficient c_j is contributed to by those terms of $\sum_{k=0}^{\infty} u_k \cdot (b_1 y + \ldots + b_{d_i} y^{d_i})^k$ for which k is in the range $[(j-(d-d_i)/2)/d_i, j-(d-d_i)/2]$. For any fixed $k \in [j/d_i, j]$, the coefficient of y^j in $(b_1 y + \ldots + b_{d_i} y^{d_i})^k$ is exactly,

$$v_{kj} = \sum_{\substack{k_1+2k_2+\ldots+d_i k_{d_i}=j,\\ k_1+\ldots+k_{d_i}=k}} \binom{k}{k_1,\ldots,k_{d_i}} \cdot b_1^{k_1} \cdot b_2^{k_2} \ldots b_{d_i}^{k_{d_i}}.$$

where $k_1, \ldots, k_d$ are positive integers. Then, assuming $B = \max_{ij}\{|b_{ij}|\}$,

$$|v_{kj}| \leq \sum_{k_1+\ldots+k_{d_i}=k} \binom{k}{k_1,\ldots,k_{d_i}} \cdot |b_1|^{k_1} \cdot |b_2|^{k_2} \ldots |b_{d_i}|^{k_{d_i}} \leq (B \cdot d_i)^k.$$

Since $c_{j+(d-d_i)/2} = \sum_{k\in[j/d_i,j]} u_k \cdot v_{kj}$,

$$\begin{aligned} |c_j| &\leq \sum_{k\in[0,j]} (B\cdot d_i)^k \leq (B\cdot d)^{j+1} \quad \text{(using Equation 8)} \\ \Rightarrow |S_j| &\leq n\cdot y^j \cdot (B\cdot d)^{j+1} \end{aligned} \tag{9}$$

Lower bound on $|S_j|$:

$$|S_j| = y^j \cdot \left|\sum_{i=1}^{n} \delta_i c_{ij}\right|$$

Let us lower bound the sum $|\sum_{i=1}^{n} \delta_i c_{ij}|$. Notice that, in the previous discussion on upper bounding $|S_j|$, the integer v_{kj} depends on the index i whereas u_k solely depends on k. So, to make the following discussion more precise, we switch to the notation v_{ikj}. Moreover, for simplicity the range of k is not specified in the following equations - the appropriate range should be clear from the context.

$$\sum_{i=1}^{n} \delta_i c_{ij} = \sum_{i=1}^{n} \delta_i \sum_k u_k \cdot v_{ikj} = \sum_k u_k \cdot \sum_{i=1}^{n} \delta_i v_{ikj}.$$

Now notice that, the sum $\sum_{i=1}^{n} \delta_i v_{ikj}$ is an integer. Hence, if $\sum_{i=1}^{n} \delta_i c_{ij} \neq 0$ then $|\sum_{i=1}^{n} \delta_i c_{ij}| \geq 1/2^{2j+1}$ (by Equation 8). Therefore,

$$|S_j| \geq \frac{y^j}{2^{2j+1}} \text{ if } S_j \neq 0. \tag{10}$$

Putting everything together. With the upper and the lower bounds on $|S_j|$ at hand, we are now ready to pinpoint the requirement that ensures that condition (6) is satisfied. We want $|S_{\ell+t}| \,/\, |S_\ell| \leq 1/2^{t+1}$, for all $t \geq 1$. Hence, by combining equations (9) and (10), it is sufficient if,

$$\begin{aligned} \frac{n\cdot y^{\ell+t}\cdot (B\cdot d)^{\ell+t+1}}{y^\ell/2^{2\ell+1}} &\leq \frac{1}{2^{t+1}} \\ \Rightarrow X^t &\geq n\cdot 2^{2\ell+t+2}\cdot (B\cdot d)^{\ell+t+1} \end{aligned}$$

Therefore, it suffices if $X \geq (B+1)^{12\cdot dn^2 \log 2d}$ (taking into consideration that $\ell \leq dn^2 + n$). And if this happens then condition (6) is satisfied and by equation (7),

$$|S| \geq \left|X^{d/2}\right| \cdot \frac{1}{2} \cdot |S_\ell| \geq \frac{1}{2^{2\ell+2}\cdot X^{\ell-d/2}} \quad \text{(by Equation 10)}$$

This implies that $|S| \geq 1/X^{8\cdot dn^2}$, which proves Theorem 1.4.

4 The complexity of DegSLP

In this section we consider the algorithmic complexity of the following problem: given a polynomial $f(\mathbf{X})$ as an arithmetic circuit, and an integer d in binary determine if $\deg(f) \leq d$. Towards this end, we need to define another natural computational problem.

CoeffSLP : Given an arithmetic circuit computing a polynomial $f(\mathbf{X})$ over integers, a monomial $\mathbf{X}^{\alpha}$ and a prime p, determine the coefficient of $\mathbf{X}^{\alpha}$ in $f(\mathbf{X})$ modulo p. (The additional input p is to ensure that the output is not too large). [3]

We will need the following theorem from [KP07]. A simpler, self-contained proof is given in the appendix.

Theorem 4.1. *[KP07]* CoeffSLP *is* #P*-complete.*

Theorem 4.2.

$$\mathsf{DegSLP} \leq_T^{\mathsf{coRP}} \mathsf{CoeffSLP}.$$

Here we prove the theorem assuming over the field of rational numbers. A similar proof will go through over other fields of zero characteristic. The theorem is valid even if the underlying field has small characteristic. We relegate the proof of the general case to the appendix.

Proof. We first reduce the multivariate to the univariate problem by making a random substitution of the form $g(z) = f(a_1 \cdot z, a_2 \cdot z, \ldots, a_n \cdot z)$.

Claim 4.3. *With high probability over a random choice of the vector* $\mathbf{a} = (a_1, \ldots, a_n)$ *we have:* $\deg(g(z)) = \deg(f(\mathbf{X}))$.

Proof of Claim 4.3: Let f have degree d. We can write the polynomial $f(\mathbf{X})$ as $\sum_{i=0}^{d} f_i(\mathbf{X})$, where each $f_i(\mathbf{X})$ is a homogeneous polynomial of degree i. Applying the substitution $x_i := a_i \cdot z$, we get that

$$g(z) = f_0 + z \cdot f_1(\mathbf{a}) + z^2 \cdot f_2(\mathbf{a}) + \ldots + z^d \cdot f_d(\mathbf{a}).$$

By the Schwartz-Zippel lemma, $f_d(\mathbf{a})$ is nonzero with high probability so that $\deg(g(z)) = \deg(f(\mathbf{X}))$ also with high probability. □

Our problem thus is the following: given an arithmetic circuit computing a univariate polynomial $g(z)$ and an integer d in binary, we want to determine if $\deg(g(z)) \geq d$. The most natural thing to do is to use the CoeffSLP oracle to determine whether the coefficient of z^d in $g(z)$ is nonzero. [4] If this happens to be nonzero we have a certificate that the degree of g is at least d. The converse is easily seen to be false: the coefficient of z^d in g can be zero and yet the degree of g can be larger than d. To fix this, we take a 'random shift' of g and compute its degree instead. Specifically, we look at the polynomial $g(z+\beta)$, where β is chosen uniformly at random from a large enough subset of $\mathbb{F}$. Clearly, $\deg(g(z)) = \deg(g(z+\beta))$. It suffices then to prove the following claim:

Claim 4.4. *With high probability over a random choice of $\beta \in \mathbb{F}$, we have: coefficient of z^d in $g(z+\beta)$ is nonzero if and only if* $\deg(g(z)) \geq d$.

Proof of Claim 4.4: We first observe that the coefficient of z^d in $g(z+\beta)$ is $\frac{1}{d!}g^{(d)}(\beta)$, where $g^{(d)}(z)$ denotes as usual the d-th order derivative of g. (To see this, first use linearity of derivatives to reduce the problem to the case where $g(z)$ is a single monomial, say $g(z) = a \cdot z^e$ and then use binomial expansion to compute the coefficient of z^d in $a \cdot (z+\beta)^e$). Since the characteristic of the

[3] Our definition is slightly different from that of [ABKPM09, KP07] and more tailored towards our needs.

[4] For the oracle call to CoeffSLP, we choose the prime p at random. This ensures that with high probability, the coefficient α of z^d in $g(z)$ is zero as a rational number if and only if α is zero modulo p.

field is zero, $g^{(d)}(z)$ has degree precisely $\deg(g) - d$. In particular, $g^{(d)}(z)$ is identically zero if and only if $\deg(g) < d$. Now the claim follows from an application of the Schwarz-Zippel lemma. □

This completes the proof of the theorem. □

Combining Theorems 4.1 and 4.2, we immediately get:

Theorem 4.5. DegSLP *is in* $\mathsf{coRP}^{\mathsf{PP}}$.

5 Discussion

We have seen that for the class of 'polynomial integers' it is possible to compare two sums of square roots by keeping precision of up to polynomially many bits (during square root computations). Although, 'polynomial integers' form a nontrivial class, the condition that X is sufficiently large also makes them very restrictive at the same time. The hope is that it may be possible to exploit results similar to that of Theorem 2.1 to show something stronger for the case of integers. As a next step, we would be interested in a similar result where X is constrained as $X \geq \mathsf{poly}(n, d) \cdot B^c$ (c is a constant), instead of $X \geq (B + 1)^{\mathsf{poly}(n,d)}$ as is the case in our analysis. Could encoding the integers as multivariate polynomials be useful in this regard?

Nonetheless, 'polynomial integers' are perhaps interesting from one perspective. A plausible way to make the sum S very small is to assume that the number of $+$ and $-$ signs in $S = \sum_{i=1}^{n} \delta_i \sqrt{a_i}$ are equal and all the integers a_i's are somewhat very close to each other. Notice that, because of the assumption that X is large, all integers of the form $X^d + b_1 X^{d-1} + \ldots + b_d$ are reasonably close to X^d. Our analysis shows that at least for this case any nonzero sum S is still sufficiently large. As a final remark on the sum of square roots problem, we would like to note that the proofs and the results presented here generalize in a straightforward manner to general sums of radicals - like sums of cube roots or fourth roots of integers.

We feel that the complexity of the problem DegSLP is not understood well enough. In particular no hardness results are known for it. We conclude by posing the following problem:

Problem 5.1. (Hardness of DegSLP): Does there exist an efficient randomized algorithm for DegSLP ? ... or, will the existence of such an algorithm for DegSLP lead to a collapse of the polynomial hierarchy?

Acknowledgement

We thank Kurt Mehlhorn and Piyush Kurur for some insightful discussions on this work.

References

[ABKPM09] Eric Allender, Peter Bürgisser, Johan Kjeldgaard-Pedersen, and Peter Bro Miltersen. On the Complexity of Numerical Analysis. *SIAM J. Comput.*, 38(5):1987–2006, 2009.

[Bô00] M. Bôcher. On linear dependence of functions of one variables. *Bull. Amer. Math. Soc.*, 7:120–121, 1900.

[BD10] Alin Bostan and Philippe Dumas. Wronskians and linear independence. *American Mathematical Monthly*, 117(8):722–727, 2010.

[BFMS00] Christoph Burnikel, Rudolf Fleischer, Kurt Mehlhorn, and Stefan Schirra. A Strong and Easily Computable Separation Bound for Arithmetic Expressions Involving Radicals. *Algorithmica*, 27(1):87–99, 2000.

[Blö91] Johannes Blömer. Computing Sums of Radicals in Polynomial Time. In *Proceedings of the 32nd IEEE Symposium on Foundations of Computer Science, FOCS'91*, pages 670–677, 1991.

[Blö93] Johannes Blömer. Computing Sums of Radicals in Polynomial Time. 1993. Available from http://www.cs.uni-paderborn.de/uploads/tx_sibibtex/ComputingSumsOfRadicals.pdf.

[Che06] Qi Cheng. On Comparing Sums of Square Roots of Small Integers. In *MFCS*, pages 250–255, 2006.

[CMSC10] Qi Cheng, Xianmeng Meng, Celi Sun, and Jiazhe Chen. Bounding the Sum of Square Roots via Lattice Reduction. *Mathematics of Computation*, 79:1109–1122, 2010.

[EHM05] G. Effinger, K. H. Hicks, and G. L. Mullen. Integers and Polynomials: comparing the close cousins Z and F[x]. *Mathematical Intelligencer*, (27):26–34, 2005.

[GGJ76] M. R. Garey, Ronald L. Graham, and David S. Johnson. Some NP-Complete Geometric Problems. In *STOC*, pages 10–22, 1976.

[Goe98] Michele Goemans. Semidefinite Programming and Combinatorial Optimization. *Documenta Mathematica, Extra Volume ICM 1998*, III:657–666, 1998.

[IL93] Neil Immerman and Susan Landau. The Similarities (and Differences) between Polynomials and Integers. In *International Conference on Number Theoretic and Algebraic Methods in Computer Science*, pages 57–59, 1993.

[KP07] Pascal Koiran and Sylvain Perifel. The complexity of two problems on arithmetic circuits. *Theor. Comput. Sci.*, 389(1-2):172–181, 2007.

[Mal96] G Malojovich. An effective version of Kronecker's theorem on simultaneous Diophantine equation. Technical report, City University of Hong Kong, 1996.

[MR08] Wolfgang Mulzer and Günter Rote. Minimum-weight triangulation is NP-hard. *Journal of the ACM*, 55(2), 2008.

[MS00] Kurt Mehlhorn and Stefan Schirra. Generalized and improved constructive separation bound for real algebraic expressions. Research Report MPI-I-2000-1-004, Max-Planck-Institut für Informatik, Stuhlsatzenhausweg 85, 66123 Saarbrücken, Germany, November 2000.

[QW06] Jianbo Qian and Cao An Wang. How much precision is needed to compare two sums of square roots of integers? *Inf. Process. Lett.*, 100(5):194–198, 2006.

[vL99] Jacobus Hendricus van Lint. *Introduction to coding theory.* Springer, 1999.

Appendix

A The complexity of CoeffSLP

The aim of this section of the appendix is to give a simpler, self-contained proof of the following theorem.

Theorem 4.1. CoeffSLP is #P-complete.

We first give some warm-up lemmas.

Lemma A.1. *For any $m \geq 7$, the lcm of the first m numbers is at least 2^m.*

Lemma A.2. *For any integer $t \in \mathbb{Z}_{\geq 1}$ and prime p, there is a prime $r = O(t^2 \cdot \log p)$ such that the ring*

$$R \stackrel{\text{def}}{=} \mathbb{F}_p[z]/\langle \frac{z^r - 1}{z - 1} \rangle$$

is the direct sum of finite fields of size $q > p^t$.

Proof. Consider the integer

$$M := (p-1) \cdot (p^2 - 1) \cdot \ldots \cdot (p^t - 1).$$

Then $M < p^{t^2}$. By lemma A.1, there exists a prime $r < \log M$ such that r does not divide M. This is the prime r that we seek. Let m denote the order of p modulo r , i.e. m is the smallest positive integer such that $p^m = 1 \pmod r$. Since r does not divide $M = \prod_{i \in [t]} (p^i - 1)$, therefore r does not divide any $(p^i - 1)$ for $1 \leq i \leq t$ and therefore $m > t$. Let $\phi_r(z)$ denote the r-th cyclotomic polynomial, that is

$$\phi_r(z) \stackrel{\text{def}}{=} \frac{z^r - 1}{z - 1}.$$

It is known that over $\mathbb{F}_p$, $\phi_r(z)$ factors into $\frac{r-1}{m}$ irreducible polynomials each of degree m. Thus, $R \stackrel{\text{def}}{=} \mathbb{F}_p[z]/\langle \phi_r(z) \rangle$ is the direct sum of finite fields of size $p^m > p^t$. $\square$

Proof of Theorem 4.1: The #P-hardness of this problem is well-known and a proof can be found for example in [ABKPM09]. It is sufficient to show this for univariate polynomials (by replacing each indeterminate x_i by an exponentially increasing sequence of monomials, if necessary). That is, our problem now becomes the following: given a circuit of size s computing a univariate polynomial

$f(x)$ and an $\alpha \in \mathbb{Z}_{\geq 0}$ given in binary, compute the coefficient of x^α in $f(x)$. Notice that $D \stackrel{\text{def}}{=} 2^s$ is an upper bound on $\deg(f(x))$. Using lemma A.2, we obtain an extension ring R of the form $R = \mathbb{F}_p[z]/\langle \frac{z^r-1}{z-1} \rangle$ such that $r \leq (\log D)^2 \cdot (\log p)$ and

$$R \cong \mathbb{F}_q \oplus \ldots \oplus \mathbb{F}_q,$$

with $q - 1 > D$. We now observe that the coefficient of x^α in $f(x)$ is given by

$$\mathrm{Coeff}(x^\alpha, f(x)) = -\sum_{\beta \in R^*} \beta^\alpha \cdot f(\beta^{-1}).$$

The number of terms in the above summation is exponentially large but notice that each summand in the above expression, $(\beta \cdot f(\beta^{-1}))$, is polynomial-time computable so that overall this sum is computable in $\mathsf{P}^{\#\mathsf{P}}$.

□

B DegSLP over fields of small characteristic

In this section of the appendix, we give the proof of Theorem 4.2 in the general case, i.e even when the underlying field has small characteristic. We first record a lemma that was implcit stated and used in Section 4 earlier.

Lemma B.1. *Over a field of characteristic larger than d, the coefficient of x^d in $f(x+\beta)$ is precisely $\frac{1}{d!} f_d(\beta)$.*

Proof. By the linearity of derivatives, it is sufficient to show this for monomials. So let $f(x) = a \cdot x^e$. If $e < d$ then $f_d(x)$ is the zero polynomial and we are done. So let $e \leq d$. Expanding $(x+\beta)^e$ using binomial theorem we get that coefficient of x^d is $a \cdot \binom{e}{d} \cdot \beta^{e-d}$. On the other hand $f_d(x) = a \cdot e \cdot (e-1) \cdot \ldots \cdot (e-d+1) x^{e-d}$. It is now easily verified that the coefficient of x^d in $f(x+\beta)$ is $\frac{1}{d!} f_d(\beta)$. □

We will also need a lemma originally due to Edouard Lucas.

Lemma B.2. *[vL99, p.55] Let n, m be positive inte gers whose p-ary representation is the following:*

$$m = m_0 + m_1 p + \ldots + m_d p^d, \quad \forall i : 0 \leq m_i \leq p-1$$

$$n = n_0 + n_1 p + \ldots + n_d p^d \quad \forall i : 0 \leq n_i \leq p-1.$$

Then

$$\binom{n}{m} = \binom{n_0}{m_0} \cdot \binom{n_1}{m_1} \cdot \ldots \cdot \binom{n_d}{m_d} \pmod{p}$$

In particular, for any intger $n \geq 1$, the binomial coefficient $\binom{n}{p^i}$ is divisible by p if and only if the n_i, the i-th digit in the p-ary representation of n is zero.

Proof of Theorem 4.2: Proceeding as before, we can assume without loss of generality that the given circuit computes a univariate polynomial $g(z)$. Our problem then is the following: given an arithmetic circuit computing a univariate polynomial $g(z)$ and an integer d in binary, we want to determine if $\deg(g(z)) \geq d$. Recall that in the large characteristic situation, our strategy was to choose a random $\beta \in \mathbb{F}$ and look at the coefficient of z^d in $g(z+\beta)$. To get the reduction over fields of small characteristic, we need to examine the polynomial $g(z+y)$. Specifically, we need to determine as to when does it happen that the coefficient of z^d as a polynomial in y is the identically zero polynomial. We sketch the proof below. Let the size of the circuit computing f be s. Then the formal degree of f is bounded by 2^s. First observe that multiplying $g(z)$ with a suitable power of z, we may assume without loss of generality that d is a power of p, say $d = p^t$. Notice that $\deg(g(z)) \geq p^t$ if and only if $g(z)$ contains a monomial z^m where the p-ary (base-p) representation of the positive integer m contains a non-zero digit at the i-th position, for some $t \leq i \leq s$. Thus, to achieve our objective, it is sufficient to devise a randomized procedure that given an integer $i \in [s]$, tests whether $g(z)$ contains any non-zero monomial z^m such that in the p-ary representation of the integer m, the i-th digit is non-zero. This procedure works as before: choose a random β (in a suitably large field extension of $\mathbb{F}_p$) and accept if and only if the the coefficient of z^{p^i} (computed via an oracle call to $\mathsf{CoeffSLP}$) is nonzero. We next describe why the test gives the correct answer with high probability.

Suppose that

$$g(z) = \sum_{0 \leq m \leq 2^s} a_m \cdot z^m.$$

Then the coefficient of z^{p^i} in $g(z+\beta)$ is given by

$$h(\beta) = \sum_{p^i \leq m \leq 2^s} a_m \cdot \binom{m}{p^i} \cdot \beta^{m-p^i}.$$

Our test accepts with high probability if and only if $h(\beta)$ is not the identically zero polynomial with respect to β. Use Lucas's lemma B.2 to observe that $\binom{m}{p^i}$ is zero modulo p if and only if the i-th digit in the p-ary representation of m is zero. Thus $h(\beta)$ is a nonzero polynomial if and only if there exists an $p^i \leq m \leq 2^s$, such that a_m is nonzero and in the p-ary representation of the integer m, the i-th digit is nonzero. Thus $h(\beta)$ is nonzero if and only if $g(z)$ contains a non-zero monomial z^m such that in the p-ary of the integer m, the i-th digit is non-zero. This completes the proof of the theorem.

□

Combining Theorems 4.1 and 4.2, we immediately get:

Theorem B.3. DegSLP *is in* $\mathsf{coRP}^{\mathsf{PP}}$.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de