CHICAGO JOURNAL OF THEORETICAL COMPUTER SCIENCE 2013, Article 2, pages 1–22
http://cjtcs.cs.uchicago.edu/

Few Product Gates but Many Zeroes

Bernd Borchert Pierre McKenzie* Klaus Reinhardt

Received: June 6, 2010; revised: January 20, 2013, and March 9, 2013; published: March 10, 2013.

Abstract: A *d-gem* is a $\{+,-,\times\}$-circuit having very few $\times$-gates and computing from $\{x\}\cup\mathbb{Z}$ a univariate polynomial of degree d having d distinct integer roots. We introduce d-gems because they offer the remote possibility of being helpful for factoring integers and because their existence for infinitely many d would disprove a form of the Blum-Cucker-Shub-Smale conjecture (strengthened to allow arbitrary constants). A natural step towards a better understanding of the BCSS conjecture would thus be to construct d-gems or to rule out their existence. Ruling out d-gems for large d is currently totally out of reach. Here the best we can do towards that goal is to prove that skew 2^n-gems if they exist require n $\{+,-\}$-gates and that skew 2^n-gems for any $n\geq 5$ would provide new solutions to the Prouhet-Tarry-Escott problem in number theory (skew meaning the further restriction that each $\{+,-\}$-gate merely adds an integer to a polynomial). In the opposite direction, here we do manage to construct skew d-gems for several values of d up to 55.

1 Introduction

Consider the polynomials of degree 4, degree 8 and degree 16 computed by the $\{+,\times\}$-circuits depicted in Figure 1. Each polynomial factors completely over $\mathbb{Z}$ and has all distinct roots. Clearly the number of $\times$-gates used in each case is minimum. Our main question is: *Do $\{+,-,\times\}$-circuits having n $\times$-gates and computing a polynomial $f(x)\in\mathbb{Z}[x]$ having 2^n distinct integer roots exist for every n?*
Crandall [11, Prob. 3.1.13] found a *normalized* circuit for the case $n=3$ and asked whether such circuits exist for $n>3$, where a *normalized* circuit starts from x and alternates between a squaring operation and the addition of a constant. Dilcher [13] characterized the normalized circuits for $n=3$. Crandall and

*Supported by the Natural Sciences and Engineering Research Council of Canada

Key words and phrases: Polynomials, Arithmetic Circuits, Addition Chain, Prouhet-Tarry-Escott Problem, Integer Factoring, tau-conjecture

DOI: 10.4086/cjtcs.2013.002

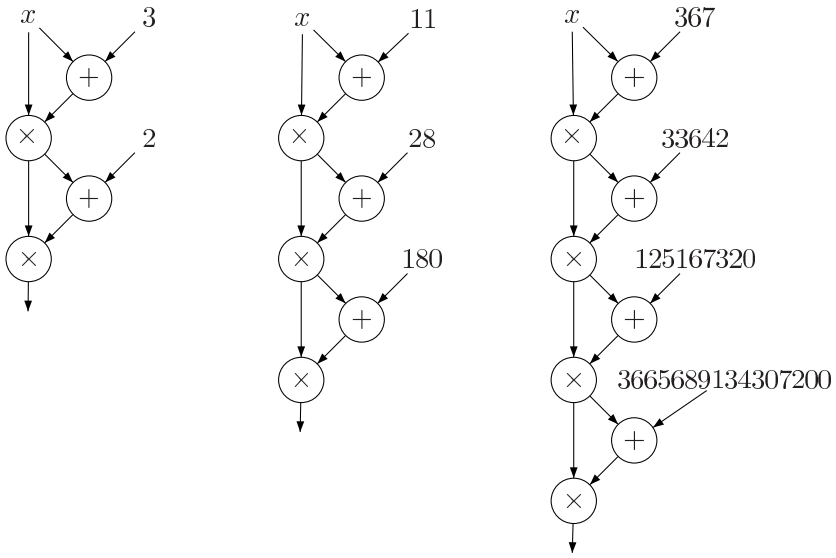

Figure 1: Circuits computing polynomials having the respective sets of roots $\{0,-1,-2,-3\}$, $\{0,-1,-2,-4,-7,-9,-10,-11\}$ and $\{0,-4,-7,-12,-118,-133,-145,-178,-189,-222,-234,-249,-355,-360,-363,-367\}$.

Pomerance [12] constructed an example for $n=4$ and Bremner [8] constructed two infinite families of examples for that case.

But why care about our main question?

First, polynomials with distinct integer roots may hold the key to factoring integers. For example, knowing [3] that $\sim n^{1/4}$ operations modulo n suffice to evaluate $x(x-1)(x-2)\cdots(x-n^{1/4}+1)$ at the points $x=n^{1/4},2n^{1/4},3n^{1/4},\ldots,n^{1/2}$, Strassen [24] noted that $\log_2 n$-bit integers can be factored in time $\sim n^{1/4}$. No provably faster deterministic algorithm for factoring is known [26]. Lipton [16] later formulated a hypothesis, on circuits computing polynomials having many distinct integer roots, whose validity would imply that the integer factoring problem is "too easy" to support cryptography. A positive answer to our main question (with further size and constructivity assumptions) would validate Lipton's hypothesis. Crandall [11] discusses further connections with factoring.

Second, a positive answer would categorically *refute* a strong form of the τ-conjecture of Blum, Cucker, Shub and Smale [21, 1]. The τ-conjecture states that any polynomial $f(x)\in\mathbb{Z}[x]$ has at most $(\tau(f)+1)^\beta$ distinct roots in $\mathbb{Z}$, where $\tau(f)$ is the size of a smallest $\{+,-,\times\}$-circuit computing $f(x)$ from $x\cup\{1\}$. The stronger $\tau_\mathbb{Z}$-conjecture is obtained when $\tau(f)$ is replaced with $\tau_\mathbb{Z}(f)$, which we define here as the size of a $\{+,-,\times\}$-circuit computing $f(x)$ from $x\cup\mathbb{Z}$ rather than from $x\cup\{1\}$. Smale named the τ-conjecture the fourth most important millennium mathematical challenge [23]. The τ-conjecture is known to imply $\mathrm{P}_\mathbb{C}\neq\mathrm{NP}_\mathbb{C}$ in the Blum-Shub-Smale model of computation over the complex numbers [1]. Providing a negative answer to our main question could be seen as adding evidence towards the $\tau_\mathbb{Z}$-conjecture and a fortiori towards the τ-conjecture.

But against all odds, our main question remains open. Here we partly explain why, by relating the question to a classical number-theoretic problem. We also extend the question to encompass polynomials

of any degree $d > 0$ by defining *d-gems*, so named to reflect their "precious and seemingly rare" nature. Let ℓ_d be the length of a shortest addition chain for d, or equivalently, the size of a smallest $\{+\}$-circuit (Section 2 defines circuits formally) computing d from $\{1\}$.

Definition 1.1. A $\{+,-,\times\}$-circuit c with inputs from $\{x\} \cup \mathbb{Z}$ is a *d-gem* if c has at most ℓ_d $\times$-gates and if the polynomial $f_c(x) \in \mathbb{Z}[x]$ computed by c has degree d and has precisely d distinct integer roots.

Hence a 2^n-gem is a $\{+,-,\times\}$-circuit with n product gates that computes a polynomial of maximum degree 2^n and this polynomial factors completely with its 2^n roots integer and distinct. Our contributions are the following:

- we observe, following [18], that any polynomial of degree d can be computed from $\{x\} \cup \mathbb{Z}$ by a $\{+,-,\times\}$-circuit using $2(\sqrt{d}+1)$ product gates;
- if d-gems exist for infinitely many d, then the $\tau_{\mathbb{Z}}$-conjecture fails;
- by theoretical considerations and computer search, we construct *skew* d-gems for $d \leq 22$ and d-gems for $d \leq 31$ and $d = 36, 37, 42, 54, 55$ (see Figure 2), where a circuit is *skew* if each of its $\{+,-\}$-gates merely adds an integer to a polynomial;
- it is not known whether there exists *any* polynomial $q \in \mathbb{Z}[x]$ computable by a $\{+,-,\times\}$-circuit having *fewer* than $\ell_{\deg(q)}$ product gates (see discussion leading to Lemma 2.1); here we show that if such a polynomial of degree d exists, then $8d \leq 2^{\ell_d}$; since $8d > 2^{\ell_d}$ holds for every $d < 71$, we conclude that all the d-gems we are able to construct so far have a minimal number of product gates;
- we prove that a skew 2^n-gems for any $n \geq 5$ would provide new solutions of size 2^{n-1} to the Prouhet-Tarry-Escott problem which arguably has a 200-year history in number theory (see for instance [6]);
- we spell out sufficient conditions implying a 2^n-gem;
- we construct skew *d-gems over the reals*, i.e. with inputs from $\mathbb{R} \cup \{x\}$ and with the requirement of distinct roots in $\mathbb{R}$, for every d;
- we prove that any skew 2^n*-gem over the reals* requires at least n $\{+,-\}$-gates; we conclude that the skew 2^n-gems (over $\mathbb{Z}$) depicted in Figure 1 have a minimal number of $\{+,-\}$-gates among all skew 2^n-gems.

Section 2 defines circuits and proves basic facts. Section 3 relates the existence of d-gems to the Prouhet-Tarry-Escott problem. Section 4 deals with gems over the real numbers. Section 5 describes our d-gem constructions. Section 6 concludes.

d	$c_\times$	c_+	$f_c(x)$	$\texttt{izeros}_c$
1	0	0	x	$\{0\}$
2	1	1	x^2-1	$\{-1,1\}$
3	2	1	$(x^2-1)x$	$\{-1,0,1\}$
4	2	2	$((x^2-5)^2-16$	$\{-1,1,-3,3\}$
5	3	2	$(((x^2-5)^2-16)x$	$\{0,-2,2,-3,3\}$
5	3	2	$(x^2-1)((x^2-4)x)$	$\{0,-1,1,-2,2\}$
6	3	2	$((x^2-25)^2-24^2)(x^2-25)$	$\{-1,1,-7,7,-5,5\}$
6	3	2	$((x^2-7)x)^2-36$	$\{1,2,-3,-1,-2,3\}$
7	4	2	$(((x^2-7)x)^2-36)x$	$\{0,-1,1,-2,2,-3,3\}$
7	4	2	$((x^2-25)^2-24^2)(x^2-25)x$	$\{0,-1,1,-7,7,-5,5\}$
8	3	3	$(((x^2-65)^2-1696)^2-2073600$	$\{-3,3,-11,11,-7,7,-9,9\}$
9	4	2	$(((x^2-49)x)^2-120^2)((x^2-49)x)$	$\{0,-3,3,-5,5,-8,8,-7,7\}$
10	4	3	$((y^2-236448)^2-123552^2)y$ with $y=(x^2-625)$	$\{\pm5,35,17,31,25\}$
10	4	3	$(((x^2-250)^2-14436)x)^2-1612802$	$\{\pm4,8,14,18,20\}$
12	4	3	$(((x^2-91)x)^2-58500)^2-50400^2$	$\{\pm1,5,6,9,10,11\}$
14	5	3	$((((x^2-7^4)x)^2-...)^2-...)(x^2-7^4)$	$\{\pm49,16,39,55,21,35,56\}$
15	5	3	$y\times(y^2-34320^2)\times(y^2-41160^2)$ with $y=(x^2-7^4)x$	$\{\pm0,49,16,39,55,21,35,56\}$
16	4	4	$((((x^2-67405)^2-3525798096)^2-...)^2-...$	$\{\pm11,367,131,343,77,359,101,353\}$
18	5	5	$f_{c16}\cdot(x^2-1)$	$\text{Set}_{16}\cup\{-1,1\}$
18	5	4	$(y^2-2484^2)\times(y^2-4116^2)\times(y^2-5916^2)$ with $y=(x^2-7^2\cdot13)x$	$\{\pm4,23,27,7,21,27,12,17,29\}$
20	5	5	$f_{c_{16}}\cdot((x^2-67405)^2-3958423056)$	$\{\pm67,361,11,367,131,343,77,359,101,353\}$
21	6	4	$y\times(y^2-89760^2)\times(y^2-150480^2)\times(y^2-263640^2)$ with $y=(x^2-7^2\cdot13^2)x$	$\{\pm0,91,11,85,96,19,80,99,39,65,104\}$
22	6	6	$f_{c20}\cdot(x^2-1)$	$\text{Set}_{20}\cup\{-1,1\}$
23	6		$y\times(y-4838400x^3+208051200x)\times(x^3-16x)$ with $y=z\times(z+45x^3-700x^2-2835x+630)$ and $z=(x^3+x^2-197x+195)\times(x^2-x-42)$	$\{\pm0,1,2,3,4,6,7,9,10,13,14,15\}$
24	5		$f_{c_4}(y^2)$ with $y=(x^2-7\cdot13\cdot19)x$	$\{\pm3,40,43,8,37,45,15,32,47,23,25,48\}$
24	5	442	$z(z+c_{Prop.3.5})$ with $y=(x^2-11763)^2$ and $z=(y+241x^2+..)(y+195x^2+..)(y+x^2+..)$	$\{\pm22,61,86,127,140,151,35,47,94,121,146,148\}$
26	6	443	$f_{c24}\cdot(x^2-1)$	$\text{Set}_{24}\cup\{-1,1\}$
27	6		$y\times f_{c_4}(y^2)$ with $y=(x^2-7^2\cdot13^2)x$	$\text{Set}_{21}\cup\{\pm49,56,105\}$
28	6	560	$f_{c24}\cdot(y+117x^2+...)$	$\text{Set}_{24}\cup\{-1,1,-153,153\}$
30	6		$f_{c_5}(y^2)$ with $y=(x^2-7^2\cdot13^2\cdot19)x$	$\{\pm13,390,403,35,378,413,70,357,427,103,335,438,117,325,442\}$
36	6		$f_{c_6}(y^2)$ with $y=(x^2-7^2\cdot13^2\cdot19)x$	$\text{Set}_{30}\cup\{\pm137,310,447\}$
42	7		$f_{c_7}(y^2)$ with $y=(x^2-7^2\cdot13^2\cdot19)x$	$\text{Set}_{36}\cup\{\pm182,273,455\}$
54	7		$f_{c_9}(y^2)$ with $y=(x^2-7^2\cdot13^2\cdot19)x$	$\text{Set}_{42}\cup\{\pm202,255,457,225,233,458\}$
55	8		The above $\times x$	The above $\cup\{0\}$

Figure 2: Some d-gems and their numbers $c_\times$ of $\times$-gates and c_+ of $\{+,-\}$-gates. When two examples are given for some d, these arise from different minimal addition chains for d. The functions f_{c_i} are from Lemma 5.1 if $i\le 9$ and from the i-gem in this table otherwise. We omitted the cases $d=11,13,17,19,25,29,31,37$ which, like the case $d=55$ here, are obtained by extending a $(d-1)$-gem; we note that such an extension fails for 43 which has a shorter addition chain bypassing 42.

2 Preliminaries and basic facts

By an (*arithmetic*) $\{+,-,\times\}$-*circuit* c we mean a rooted directed acyclic graph with in-degree-2 nodes called *product gates* labeled with $\times$, in-degree-2 nodes called *additive gates* labeled with $+$ or $-$ and in-degree-0 nodes called *input gates* labeled with an integer or the variable x. We write $c_\times$ and c_+ for the numbers of product and additive gates in c respectively. The *size* of c is $c_\times + c_+$. A circuit c *represents* or *computes* a polynomial $f_c(x) \in \mathbb{Z}[x]$. A *zero* or *root* of c is an integer a such that $f_c(a) = 0$. We write $\texttt{izeros}_c$ for the set of zeros of c. For example, if c is the leftmost circuit in Figure 1, then $c_\times = c_+ = 2$ and c represents $f_c(x) = (x(x+3))(x(x+3)+2) = x^4+6x^3+11x^2+6x$ having $\texttt{izeros}_c = \{0,-1,-2,-3\}$. An *addition chain* for a natural number d is an increasing sequence $d_0 = 1, d_1, \ldots, d_k = d$ of natural numbers such that each d_i for $i > 0$ is the sum of two earlier numbers in the sequence. The polynomial x^d is computable by an optimal $\{\times\}$-circuit having ℓ_d product gates, where ℓ_d is the minimum k for which there is an addition chain such that $d_k = d$ (see [15] for extensive related facts on addition chains, such as $\lceil \log_2 d \rceil \leq \ell_d \leq 2\lfloor \log_2 d \rfloor$ for all d).

Does a $\{+,-,\times\}$-circuit c having fewer $\times$-gates than $\ell_{\deg(f_c(x))}$ exist? Strangely, this question has never been resolved[1]. In particular, a polynomial $q(x)$ could well equal $p_1(x) - p_2(x)$ where $\deg(p_1) > \deg(q)$ but $\ell_{\deg(p_1)} < \ell_{\deg(q)}$. It is not inconceivable in such a case that p_1 and p_2 could be computed using in total fewer $\times$-gates than are required to compute q via an addition chain for $\deg(q)$. In view of this, and because we did not want declaring the gem status to require a lower bound proof, we imposed "$c_\times \leq \ell_d$" rather than "no circuit with fewer $\times$-gates than $c_\times$ computes $f_c(x)$" in the Definition 1.1 of a d-gem c. When c is a 2^n-gem for some $n \geq 0$ however, $2^{\ell_{2^n}} = 2^n = \deg(f_c(x)) \leq 2^{c_\times} \leq 2^{\ell_{2^n}}$ so $c_\times$ is then minimal. Since $d < 71$ implies $d > 2^{l_d - 3}$ (see [15, p. 446]), by Lemma 2.1 the minimality of $c_\times$ also extends to all d-gems c constructed in this paper. Our proof of Lemma 2.1 is technical and appears in an Appendix.

Lemma 2.1. *If $d > 2^{l_d-3}$, then any $\{\times,+,-\}$-circuit computing a polynomial $f(x) \in \mathbb{Z}[x]$ of degree d requires at least ℓ_d product gates.*

Because any fixed integer can be constructed from 1 by a fixed $\{+,-\}$-circuit, the $\tau_\mathbb{Z}$-conjecture clearly fails if skew d-gems exist for infinitely many d, but this also holds for d-gems:

Proposition 2.2. *If d-gems exist for infinitely many values of d, then the $\tau_\mathbb{Z}$-conjecture fails.*

Proof. Since $c_\times \leq \ell_d \leq 2\log_2 d$, a d-gem c computes, even without division, a polynomial having $d = \Omega(2^{c_\times/2})$ distinct integer roots. To disprove the $\tau_\mathbb{Z}$-conjecture, which implies that the maximum number of distinct integer roots of a polynomial computed by a circuit of size t is roughly t^β for some fixed β, we need a lower bound for d in terms of the *total* number of gates in c. It thus suffices to argue that a d-gem c can be simulated by a $\{\times,+,-\}$-circuit of total size polynomial in $c_\times$. To do this, we note as in [18] that for $1 \leq i \leq c_\times$, the ith product gate g in c computes

$$\left(\sum_{j=-1}^{i-1} a_{i,j}\mu_j\right) \times \left(\sum_{j=-1}^{i-1} b_{i,j}\mu_j\right) \tag{2.1}$$

[1]The issue was credited to Volker Strassen and brought to our attention by Allan Borodin. A cryptic mention of this question can be found in [4, p. 26]. Volker Strassen in 2008 kindly acknowledged having considered the question thirty years earlier, but did not have an answer. The existence of such circuits seems unlikely, which might be the weak consensus among the few experts we consulted.

where $a_{i,j}, b_{i,j} \in \mathbb{Z}$, $\mu_{-1} = 1$, $\mu_0 = x$ and $\mu_1, \ldots, \mu_{i-1}$ are polynomials computed by earlier product gates. Hence a subcircuit of size linear in $c_\times$ can compute (2.1) from $\mu_{-1}, \mu_0, \ldots, \mu_{i-1}$ and the integer constants. It follows that a circuit c' with $c'_\times + c'_+ = O(c_\times^2)$ computes $f_c(x)$. □

To qualify as a d-gem, a circuit c must be extremely efficient in terms of its number of product gates, that is, $c_\times \leq \ell_d \leq 2\log_2 d$. At the opposite end of the spectrum, we have:

Proposition 2.3. *[18] Any degree-d polynomial $f(x) \in \mathbb{Z}[x]$ can be computed by a $\{+,-,\times\}$-circuit having at most $2\sqrt{d}+1$ product gates.*

Proof. Write $f(x) = (\cdots((g_k x^k + g_{k-1})x^k + g_{k-2})x^k + \cdots + g_1)x^k + g_0$ where each $g_i \in \mathbb{Z}[x]$ has degree $k = \lceil \sqrt{d} \rceil$. Once $x^2, x^3, \cdots, x^k$ are available, each g_i is computable using additive gates alone. Another k products by x^k suffice. □

3 Gems and the Prouhet-Tarry-Escott problem

Gems by definition need not be skew. But for any $n > 4$, we are unable to rule out the existence of a 2^n-gem that even fulfills skewness. In Subsection 3.1, we relate skew 2^n-gems to the number-theoretic Prouhet-Tarry-Escott problem by showing as Corollary 3.6 that for any $n > 4$, the existence of a skew 2^n-gem would yield new PTE solutions. In Subsection 3.2, we explore the possibility of going the other way, i.e., of extracting gems or skew gems from solutions to the Prouhet-Tarry-Escott problem.

3.1 Skew 2^n-gems yield PTE solutions

Recall from Section 1 that a 2^n-gem is *normalized* if its computation starts from x and iterates the sequence "squaring then adding a constant" at least once. A normalized 2^n-gem thus has n $\times$-gates and n $+$-gates and is entirely described by a sequence $\gamma_1, \ldots, \gamma_n$ of integers with $n \geq 1$. Normalized 8-gems and normalized 16-gems are investigated in [11, 13, 12, 8].

Consider an arbitrary skew 2^n-gem c. Consecutive $+$-gates in c can be merged, i.e. $(g+a)+b$ for a gate g and $a, b \in \mathbb{Z}$ can be rewritten $g+(a+b)$. And for c to reach degree 2^n, each $\times$-gate g must reach twice the degree of the nearest $\times$-gate g' having a path to g, i.e. g must perform $(g'+a) \times (g'+b)$ with $a, b \in \mathbb{Z}$. Hence c can be taken to be the circuit $Skew(\alpha_0, \beta_0, \ldots, \alpha_{n-1}, \beta_{n-1}, \alpha_n)$ depicted in Figure 3.

Proposition 3.1. *Let $\alpha_0, \ldots, \alpha_n, \beta_0, \ldots, \beta_{n-1}, a_1, \ldots, a_{2^n} \in \mathbb{Z}$ be such that the circuit*

$$c = Skew(\alpha_0, \beta_0, \ldots, \alpha_{n-1}, \beta_{n-1}, \alpha_n)$$

computes the polynomial $p(x) = \prod_{i=1}^{2^n}(x - a_i)$. Then for any $t \in \mathbb{Z} \setminus \{0\}$, the circuit

$$c' = Skew(\alpha_0 t^{2^0}, \beta_0 t^{2^0}, \ldots, \alpha_{n-1} t^{2^{n-1}}, \beta_{n-1} t^{2^{n-1}}, \alpha_n t^{2^n})$$

computes the polynomial $q(x) = \prod_{i=1}^{2^n}(x - t a_i)$.

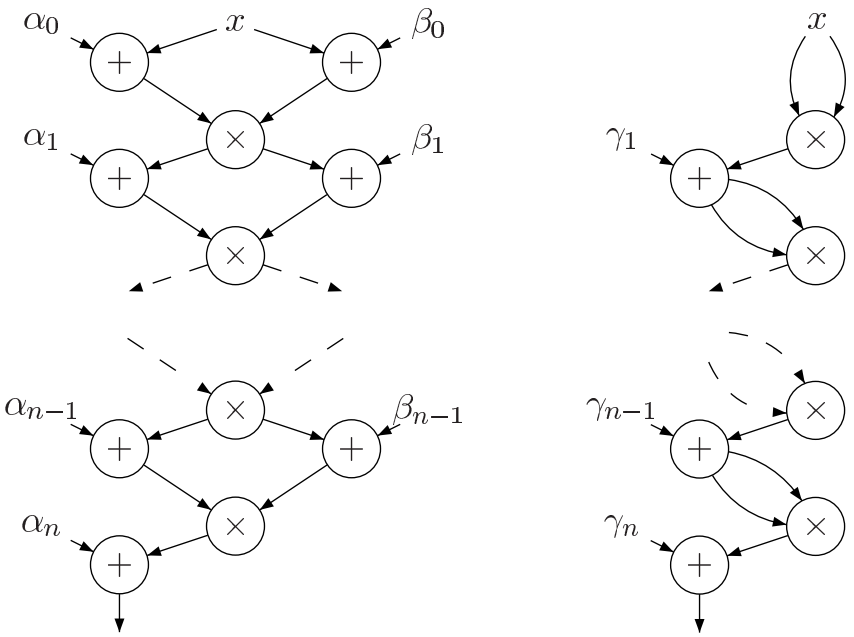

Figure 3: The skew 2^n-gem $Skew(\alpha_0, \beta_0, \ldots, \alpha_{n-1}, \beta_{n-1}, \alpha_n)$ and a skew circuit in the normal form produced by Lemma 3.2.

Proof. We claim that c' computes $q(x) = t^{2^n} p(x/t)$. This concludes the proof since then $q(ta_1) = \cdots = q(ta_{2^n}) = 0$. So let $n = 0$. Then c' computes $x + \alpha_0 t^{2^0} = t \cdot (x/t + \alpha_0) = t \cdot p(x/t)$. Now consider $n > 0$ and let q_1, q_2 and p_1, p_2 be computed by the gates input to the lowest $\times$-gate in c' and c respectively. Then $q(x) = q_1(x) \times q_2(x) + \alpha_n t^{2^n} = t^{2^{n-1}} p_1(x/t) \times t^{2^{n-1}} p_2(x/t) + \alpha_n t^{2^n}$ by induction, and the latter equals $t^{2^n}[p_1(x/t) \times p_2(x/t) + \alpha_n] = t^{2^n} p(x/t)$. □

Lemma 3.2. *(Normal form) Let $n \geq 1$. Given $a_1, \ldots, a_{2^n} \in \mathbb{Z}$ and a skew 2^n-gem c such that $f_c(x) = \prod_{i=1}^{2^n}(x - a_i)$, there exist $s \in \mathbb{Z}$ and $t \in \{1, 2\}$ and a normalized 2^n-gem computing the polynomial $\prod_{i=1}^{2^n}(x - ta_i - s)$.*

Proof. Let $c = Skew(\alpha_0, \beta_0, \ldots, \alpha_{n-1}, \beta_{n-1}, \alpha_n)$ (see Figure 3). If $\alpha_i + \beta_i$ is odd for some $0 \leq i < n$ then let $t = 2$, else let $t = 1$. The skew 2^n-gem $c' = Skew(\alpha_0 t^{2^0}, \beta_0 t^{2^0}, \ldots, \alpha_{n-1} t^{2^{n-1}}, \beta_{n-1} t^{2^{n-1}}, \alpha_n t^{2^n})$ computes $q(x) = \prod_{i=1}^{2^n}(x - ta_i)$ by Proposition 3.1. We will now normalize c'. First we rewrite each $\times$-gate $(g + a) \times (g + b) = g^2 + (a + b)g + ab$ as $[g + (a + b)/2]^2 + [ab - ((a + b)/2)^2]$, noting that any such $a + b$ occurring in c' is even. Then we merge consecutive $+$-gates since these are skew. The result would be normalized, if we did not have an extraneous $x + s$ gate at the input level. We replace $x + s$ with x. This yields a normalized 2^n-gem computing $q(x - s) = \prod_{i=1}^{2^n}(x - ta_i - s)$. □

Definition 3.3. (see [6, 7]) Two sets $\{a_1, ..., a_m\}, \{b_1, ..., b_m\}$ solve the *PTE (Prouhet-Tarry-Escott) problem* of degree k if $a_1^i + ... + a_m^i = b_1^i + ... + b_m^i$ for all $i \leq k$. A solution is called *ideal* if $k = m - 1$ and then m is called the *size* of the solution. A solution of the form $\{a_1, -a_1, \ldots, a_{m/2}, -a_{m/2}\}$, $\{b_1, -b_1, \ldots,$

$b_{m/2}, -b_{m/2}\}$ is called *symmetric* and we abbreviate it by $\{a_1,\ldots,a_{m/2}\},\{b_1,\ldots,b_{m/2}\}$; a solution of the form $\{a_1,\ldots,a_m\},\{-a_1,\ldots,-a_m\}$ is also called symmetric.

Remark 3.4. The largest ideal PTE solution known today is due to Shuwen [22] and only has degree 11, i.e., size 12 [27]. Borwein and Ingalls [6, p. 8] point out that "it has been conjectured for a long time that ideal PTE solutions exist for every n". Borwein [5, p. 87] recalls that Wright in 1934 "specifically conjectured that it is always possible to find ideal PTE solutions". Borwein then raises the fact that "heuristic arguments suggest that Wright's conjecture should be false" yet "finds intriguing that ideal solutions exist for as many n as they do".

Let $p(x)=(x-a_1)(x-a_2)\cdots(x-a_m)$ and $q(x)=(x-b_1)(x-b_2)\cdots(x-b_m)$. Define, for $k=1,2,\ldots,m$, $s_k=\sum_{i=1}^{m} a_i^k$ and $t_k=\sum_{i=1}^{m} b_i^k$.

Proposition 3.5. *([14], see [6].) The following are equivalent:*

- $s_1=t_1$ *and* $s_2=t_2$ *and* $s_3=t_3$ *and* $\cdots$ *and* $s_k=t_k$
- $degree[p(x)-q(x)]\leq m-(k+1)$.

Proof. Apply Newton's formulas [25] to $p(x)$ and $q(x)$. These formulas relate the roots $a_1,\ldots,a_m$ of the polynomial $p(x)=\sum_{i=0}^{m} c_i x^{m-i}$ to its coefficients by means of $s_1,s_2,\ldots,s_k$:

$$\begin{aligned} s_1+c_1&=0\\ s_2+c_1s_1+2c_2&=0\\ s_3+c_1s_2+c_2s_1+3c_3&=0\\ &\vdots\\ s_m+c_1s_{m-1}+c_2s_{m-2}+\cdots+mc_m&=0. \end{aligned}$$

Setting $q(x)=\sum_{i=0}^{m} d_i x^{m-i}$, one proves by induction on $k\geq 1$ that $s_i=t_i$ for $1\leq i\leq k$ if and only if $c_i=d_i$ for $0\leq i\leq k$. □

Corollary 3.6. *For any $n\geq 1$ and any skew 2^n-gem c such that $\texttt{izeros}_c=\{a_1,\ldots,a_{2^n}\}$, there is a partition $S\uplus T=\{a_1,\ldots,a_{2^n}\}$ into two equal size sets such that the pair S,T is an ideal PTE solution of size 2^{n-1}.*

Proof. We apply Lemma 3.2 to c and obtain a circuit c' computing a polynomial $(r(x))^2+\gamma_n=\prod_{i=1}^{2^n}(x-ta_i-s)$ for some $\gamma_n\in\mathbb{Z}$ (Figure 3), $s\in\mathbb{Z},t\in\{1,2\}$. Hence $-\gamma_n=e$ for some $e\in\mathbb{N}$. Now $(r(x))^2-e=p(x)q(x)$, where $p(x)=r(x)+\sqrt{e}$ and $q(x)=r(x)-\sqrt{e}$. Since $\mathbb{Z}[x]$ is a Euclidean ring, $p(x)$ and $q(x)$ must each have 2^{n-1} distinct roots. (So $\sqrt{e}\in\mathbb{N}$.) Now $\deg(p(x)-q(x))=0$, so applying Proposition 3.5 with $k=2^{n-1}-1$ shows that $\{a\in\mathbb{Z}:p(a)=0\}$ and $\{a\in\mathbb{Z}:q(a)=0\}$ form an ideal PTE solution of size 2^{n-1}. It is well known [6] that shifting from ta_i+s to ta_i preserves PTE solutions. Further dividing out the ta_i by t to get back to the a_i also preserves PTE solutions. □

A skew 32-gem, if it exists, would therefore yield an ideal PTE solution of degree 15, i.e., of size 16, while the largest ideal PTE known today has size 12 (Remark 3.4).

3.2 On the possibility of extracting gems from PTE solutions

The 16-gem in Figure 2 leads to the symmetric ideal solution $\{\pm 11, 367, 131, 343\}, \{\pm 77, 359, 101, 353\}$ as an example. This solution has the additional property that $11^2 + 367^2 = 134810 = 131^2 + 343^2$.
For the converse direction, consider the ideal symmetric PTE solution $\{2, 16, 21, 25\}$, $\{5, 14, 23, 24\}$ of degree 7 [6, p. 8]. We could try to unravel the Corollary 3.6 construction, by expressing $p(x) = (x^2-2^2)(x^2-16^2)(x^2-21^2)(x^2-25^2)$ using 3 products. Here this happens to be possible, by calculating $(x^2-2^2)(x^2-25^2)$ using 2 products, then forming $(x^2-16^2)(x^2-21^2)$ from $(x^2-2^2)(x^2-25^2)$ by repeatedly subtracting x^2 (unavoidable since $2^2+25^2 \neq 16^2+21^2$), and finally obtaining $p(x)$ using one last product. In this special case, we could therefore construct a (non-skew) 16-gem from an ideal symmetric PTE solution.
Figure 2 contains further gems (i.e. 24 and 23) constructed with the help of PTE solutions where we do not care about the number of additive gates:
We use the symmetric ideal degree-11 solution $\{22, 61, 86, 127, 140, 151\}$, $\{35, 47, 94, 121, 146, 148\}$ found by Kuosa, Meyrignac and Shuwen [22] to construct a circuit with 5 multiplications and 24 zeros as follows. Note that the polynomial

$$p(x) = (x^2-22^2)(x^2-61^2)(x^2-86^2)(x^2-127^2)(x^2-140^2)(x^2-151^2)$$

can be computed using 4 products, for example by calculating x^2 and $(x^2-11763)^2 = x^4 - 23526x^2 + 11763^2$ using 2 multiplications, then calculating $(x^4 - 23285x^2 + 22^2 \cdot 151^2)$, $(x^4 - 23321x^2 + 61^2 \cdot 140^2)$ and $(x^4 - 23525x^2 + 86^2 \cdot 127^2)$ using (only a few hundred) additions, and finally computing $p(x)$ by multiplying the latter three polynomials using 2 more products. By Proposition 3.5, there exists a constant $c_{Prop3.5} \in \mathbb{Z}$ such that

$$p(x) + c_{Prop3.5} = (x^2-35^2)(x^2-47^2)(x^2-94^2)(x^2-121^2)(x^2-146^2)(x^2-148^2),$$

Hence one last product computes the 24-gems $p(x) \times (p(x) + c_{Prop3.5})$.
The construction for degree 23 uses the PTE
$\{1, 2, 3, -6, 7, -9, -10, 13, 14, -15\}, \{-1, -2, -3, 6, -7, 9, 10, -13, -14, 15\}$ of degree 6, which was found using a Java applet written by Natali Zimmermann. By Proposition 3.5, this means that the polynomial y with 10 zeros and the polynomial for the other 10 zeros differ only in degrees up to 3, which are available in the circuit, furthermore the first half can be split to $\{1, -6, 7, 13, -15\}, \{2, 3, -9, -10, 14\}$ which is a PTE of degree 1, which means that the polynomial z with 5 zeros and the polynomial for the other 5 zeros differ again only in degrees up to 3.
But even if PTE solutions of degree 15 were known, these would need to fulfill additional properties in order to yield 32-gems by the strategy described above. We now show that the "sym-perfect" condition below is a sufficient additional condition imposed on ideal symmetric PTE solutions to yield a normalized 2^n gem.

Definition 3.7. A pair of sets $S, T \subset \mathbb{Z}$ is called *sym-perfect* if $S = \{a\}$ and $T = \{-a\}$, or if $S = \{a_1, -a_1, \ldots, a_{m/2}, -a_{m/2}\}$, $T = \{b_1, -b_1, \ldots, b_{m/2}, -b_{m/2}\}$ and for some $s \in \mathbb{Z}$ the pair of smaller size sets $\{a_1^2 + s, \ldots, a_{m/2}^2 + s\}, \{b_1^2 + s, \ldots, b_{m/2}^2 + s\}$ is sym-perfect.

If a_1^2 and $a_{m/2}^2$ are respectively smallest and largest in $\{a_1^2,\cdots,a_{m/2}^2\}$, then the shift s occurring in the recursive definition of sym-perfect is necessarily equal to $-(a_1^2+a_{m/2}^2)/2$ (which is also the average of all the numbers that occur). An example of a sym-perfect pair is $\{-3,3,-11,11\},\{-7,7,-9,9\}$, since the pair $\{9-65,121-65\},\{49-65,81-65\}$ is sym-perfect by virtue of the pair $\{56^2-1696\},\{16^2-1696\}$ in turn being sym-perfect.

Theorem 3.8. *A set $U\subset\mathbb{Z}$ of size 2^{n+1} can be written as $S\cup T$ for a sym-perfect pair S,T if and only if U is the set of zeros of a normalized 2^{n+1}-gem.*

Proof. The sym-perfect pair $\{a\},\{-a\}$ corresponds to the 2-gem c with $f_c(x)=x^2-a^2$ and $\texttt{izeros}_c=\{a,-a\}$, which forms the induction basis $n=0$.
Now let $U=\{a_1,-a_1,\ldots,a_{2^n/2},-a_{2^n/2}\}\cup\{b_1,-b_1,\ldots,b_{2^n/2},-b_{2^n/2}\}$ where the pair of smaller size sets $S'=\{a_1^2-s,\ldots,a_{2^n/2}^2-s\}$, $T'=\{b_1^2-s,\ldots,b_{2^n/2}^2-s\}$ is sym-perfect. By induction, $S'\cup T'$ is $\texttt{izeros}_{c'}$ for a 2^n-gem c'. Hence U is $\texttt{izeros}_c$ for the 2^{n+1}-gem c with $f_c(x)=f_{c'}(x^2-s)$. Conversely, a normalized 2^{n+1}-gem c with zero set U computes $f_c(x)=g(x)^2+\gamma_{n+1}=f_{c_S}(x)\times f_{c_T}(x)$ for some normalized 2^n-gems c_S and c_T where γ_i is the same as in c for all $i<n$ (see Figure 3). Then $U=S\cup T$ for their zeros $S=\{a_1,-a_1,\ldots,a_{2^n/2},-a_{2^n/2}\}$ and $T=\{b_1,-b_1,\ldots,b_{2^n/2},-b_{2^n/2}\}$. The pair S',T' as above with $s=-\gamma_1$ for γ_1 from the gem c is sym-perfect by induction, which makes S,T sym-perfect. □

Corollary 3.9. *A sym-perfect pair is an ideal, symmetric PTE solution.*

Proof. We have to show that $a_1^i+(-a_1)^i+\ldots a_{m/2}^i+(-a_{m/2})^i=b_1^i+(-b_1)^i+\ldots b_{m/2}^i+(-b_{m/2})^i$ for all $i<m$. For odd i, neighbours cancel. For $i=2i'$, this is double the i'-th equation of the PTE $\{a_1^2,\ldots,a_{m/2}^2\},\{b_1^2,\ldots,b_{m/2}^2\}$ which is just a shift of the sym-perfect PTE from the induction. □

4 The case of real numbers

Exclusively in this section, we extend Definition 1.1 to the case of real numbers, that is, to $\{+,-,\times\}$-circuits having inputs from $\{x\}\cup\mathbb{R}$ and computing polynomials that factor completely over $\mathbb{R}$. We will prove as Theorem 4.4 that any skew 2^n-gems c over $\mathbb{R}$ (these are easily shown to exist for every n) satisfies $c_+\geq n$. It follows as Corollary 4.5 that any skew 2^n-gem c over $\mathbb{Z}$ (these are not known to exist for $n>4$, as we have seen) would require at least n additive gates as well. We will also point out the existence of skew d-gems over $\mathbb{R}$ for every d, which is not a surprise since the τ-conjecture (and a fortiori the $\tau_{\mathbb{Z}}$-conjecture) is known to be false over the reals [9].
The methods used in this section are standard (see [17, Chapter 12]). For instance, Rolle's theorem was already used in [2] to investigate the additive complexity of a polynomial. Lemma 4.3 and Theorem 4.4 below do need a separate proof however because they apply to the specific case of skew circuits. The best lower bound known on c_+ for an arbitrary circuit c over the reals computing a polynomial $p\in\mathbb{R}[x]$ having 2^n distinct real roots follows from Khovanskii's work on multivariate polynomials and is $\frac{1}{4}\sqrt{n}$ (by Grigoriev and Risler, see [17, Theorem 12.12]), while our (much easier) lower bound in the case of skew such circuits is n.
Let A be $\mathbb{Z}$, $\mathbb{Q}$ or $\mathbb{R}$. For short, we will say that a nonzero polynomial $p(x)\in\mathbb{R}[x]$ *splits over* A if $\deg(p)=0$ or if p has $\deg(p)$ distinct roots in A.

Proposition 4.1. *Let A be $\mathbb{Z}$, $\mathbb{Q}$ or $\mathbb{R}$. Let $p \in \mathbb{R}[x]$ and $q \in \mathbb{R}[x]$. If pq splits over A then both p and q split over A.*

Proof. Let pq split over A. If $\deg(p) = \deg(q) = 0$ then we are done. Otherwise, since $\mathbb{R}$ is an entire ring, each root $a \in A$ of pq satisfies $p(a) = 0$ or $q(a) = 0$. The only way for p and q to account for the $\deg(p) + \deg(q)$ distinct roots of pq in A is for p to absorb its maximum number $\deg(p)$ of such roots and for q to absorb the rest. Hence p and q split over A. □

Theorem 4.2. *(Rolle) Let $p \in \mathbb{R}[x]$; if p splits over $\mathbb{R}$, then so does its derivative p'.*

Note that Rolle only applies over $\mathbb{R}$. For example, $(x-1)(x-2)(x-3)$ splits over $\mathbb{Z}$ but its derivative $3x^2 - 12x + 11$ splits neither over $\mathbb{Z}$ nor over $\mathbb{Q}$. For that reason, the only proof we have of Corollary 4.5 below is via Theorem 4.4.

Lemma 4.3. *Let $e \in \mathbb{R}$ and suppose that a polynomial $p(x) + e \in \mathbb{R}[x]$ of degree 2^n splits over $\mathbb{R}$. Then the following holds:*
H1. *Any skew gem over $\mathbb{R}$ for $p(x)$ has at least $n-1$ additive gates.*
H2. *If $e = 0$ then any skew gem over $\mathbb{R}$ for $p(x)$ has at least n additive gates.*

Proof. We use induction on n. In the base case $n = 0$, there is nothing to prove. So let $n \geq 1$ and consider any polynomial $p(x) \in \mathbb{R}[x]$ of degree 2^n. Let $e \in \mathbb{R}$ be such that $p + e$ splits. We need to prove that H1 and H2 hold for $p(x)$.
Consider any skew gem C over $\mathbb{R}$ for $p(x)$. Then C has the form depicted on the left of Figure 3 where the α_i, β_i are real constants. Let $q(x) \in \mathbb{R}[x]$ be the polynomial computed by the subcircuit C_q rooted at the second $\times$ gate nearest to the *output* of C, with $q(x) = x$ when $n = 1$. Note that $\deg(q) = 2^{n-1}$ and that C_q is a skew circuit having $n-1$ product gates. For some $a, b, c \in \mathbb{R}$, $p(x)$ is computed by C from $q(x)$ as follows:

$$p(x) = (q(x) + a) \times (q(x) + b) + c.$$

Then

$$p + e = q^2 + (a+b)q + ab + c + e$$

and the derivative $[p+e]'$ of $p + e$ with respect to x satisfies

$$\begin{aligned} [p+e]' &= 2qq' + (a+b)q' \\ &= (2q + a + b)q' \\ &= [q + (a+b)/2] \cdot [2q']. \end{aligned}$$

By Rolle, $[p+e]'$ splits. By Proposition 4.1, $q + (a+b)/2$ splits. The inductive H1 therefore implies that C_q contains at least $n-2$ additive gates.
Proving H2 for $p(x)$:
We now assume that $e = 0$.

Case 1: Two or more among a, b and c are nonzero. Then the total number of additive gates in C is at least $2+(n-2) \geq n$.
Case 2: Exactly one among a, b and c is nonzero.
If $c \neq 0$, then $q+(a+b)/2 = q$. Then the inductive H2 implies that C_q has at least $n-1$ additive gates, for a total of at least $1+(n-1) \geq n$ additive gates in C.
If $c = 0$, then assume with no loss of generality that $a = 0$. Since $p+e = q(q+b)$ splits, q splits by Proposition 4.1. The inductive H2 again implies that C_q has at least $n-1$ additive gates, for a total of at least $1+(n-1) \geq n$ additive gates in C.
Case 3: $a = b = c = 0$. Then $p = q^2$ has repeated roots so this case is impossible.
Proving H1 for $p(x)$:
We now assume that $e \neq 0$.
We have just proved that H2 holds for any polynomial of degree 2^n. Hence any gem over the reals for the splitting degree-2^n polynomial $[p(x)+e]+0$ requires at least n additive gates. It follows that any gem over the reals computing $p(x)$ requires at least $n-1$ additive gates, as required. □

Theorem 4.4. *A skew 2^n-gem over the reals has at least n additive gates.*

Proof. This follows by applying Lemma 4.3 with $e = 0$. □

Corollary 4.5. *Any skew 2^n-gem (over $\mathbb{Z}$) requires n additive gates.*

Proof. Let a skew 2^n-gem c compute $p(x) \in \mathbb{Z}[x]$. If c had fewer than n additive gates, then c as a skew gem over the reals would contradict Theorem 4.4. □

Chebyshev polynomials are used in [17, Example 12.2 (1)] to construct circuits c_r over $\mathbb{R}$ having $2r$ product gates and r additive gates, computing polynomials of degree 3^r with all distinct real roots. In [17, Example 12.2 (2)], a direct construction of circuits c_r with similar properties is credited to [2]. Neither of the above constructions can be deemed to provide 3^r-gems unless $2r \leq \ell_{3^r}$, which does not seem immediate for large r. Rojas [19, p. 4] on the other hand constructs 2^n-gems over $\mathbb{R}$ for any n. The following variation constructs skew 2^n-gems: $g_1 = x$ and $g_{i+1}(x) := g_i^2(x) - 2$, $1 \leq i < n$, yields $g_n(x)$ having 2^n distinct roots in $[-2,2]$. We extend this to arbitrary degrees:

Proposition 4.6. *For all $d > 0$, there exists a skew d-gem over $\mathbb{R}$.*

Proof. Given a minimal addition chain $a_1, \ldots, a_{l_d}$ with $a_{l_d} = d$, we start with the input x as the gate of degree $a_0 = 1$ and inductively define the subcircuit for degree a_i such that the function has a_i distinct zeros. For each i we consider the gates with degree a_j and a_k for $j,k < i$ with $a_i = a_j + a_k$.
If the zeros of the corresponding functions are disjoint, we simply multiply the output of these gates and obtain a function with a_i distinct zeros.
If $j = k$, we also multiply (which means in this case we square) and obtain a function with a_k double zeros, then we subtract a constant $\delta > 0$, which is smaller than any local maximum of this function. This leads to $a_i = 2a_k$ distinct zeros.
In the remaining case, we subtract a constant $\delta > 0$, which is smaller than any local maximum of the function from the gate of degree a_k and multiply with the gate of degree a_j and obtain a function with a_i distinct zeros.

In both cases, we can easily avoid the choice for δ to result in one of the zeros to be identical with a zero that occurred before in the construction because there are only finitely such bad choices among infinitely many possible choices. □

The construction in the proof of Proposition 4.6 produces at most ℓ_d additions (exactly ℓ_d when $d = 2^n$). In cases like $d = 3,7,9,27,81$, it produces $\ell_d/2$ additions. We conjecture that $c_+ \geq \ell_d/2$ for a skew d-gem c over $\mathbb{R}$.

5 Constructing further gems

In this section we construct gems (over $\mathbb{Z}$), at times with the help of a computer. Yet we do not know whether 32-gems or d-gems beyond $d = 55$ exist.

Lemma 5.1. *For every $d \leq 7$ and $d = 9$, for every d distinct integers $a_1, \ldots, a_d$, there is a d-gem c_d such that $f_{c_d}(x) = (x-a_1)(x-a_2)\cdots(x-a_d)$.*

Proof. The d-gems are: $f_{c_1}(x) = (x-a_1)$, $f_{c_2}(x) = (x-a_1) \times (x-a_2)$,
$f_{c_3}(x) = f_{c_2}(x) \times (x-a_3)$,
$f_{c_4}(x) = f_{c_2}(x) \times (f_{c_2}(x) + \underbrace{(a_1+a_2-a_3-a_4)\cdot x}_{\text{iterated additions of } x} + (a_3a_4 - a_1a_2))$,
$f_{c_5}(x) = f_{c_4}(x) \times (x-a_5)$,
$f_{c_6}(x) = f_{c_3}(x) \times (f_{c_3}(x) + a \cdot f_{c_2}(x) + (a\cdot(a_1+a_2) - a_1a_2 - a_1a_3 - a_2a_3)\cdot x + (a_4a_5a_6 - a_1a_2a_3 - a\cdot a_1a_2))$
with $a = (a_1+a_2+a_3-a_4-a_5-a_6)$,
$f_{c_7}(x) = f_{c_6}(x) \times (x-a_7)$,
$f_{c_9}(x) = f_{c_6}(x) \times (f_{c_3}(x) + a \cdot f_{c_2}(x) + (a\cdot(a_1+a_2) - a_1a_2 - a_1a_3 - a_2a_3)\cdot x + (a_7a_8a_9 - a_1a_2a_3 - a\cdot a_1a_2))$
with $a = (a_1+a_2+a_3-a_7-a_8-a_9)$. □

The number of additive gates used in proving Lemma 5.1 can be reduced when the roots satisfy favorable conditions, such as $a_1 = -a_2$ or $a_i = 0$ for some i. Such relationships between the zeros are exploited extensively in Figure 2.
The case $d = 8$ is missing from Lemma 5.1 because not all polynomials having 8 distinct roots have an 8-gem [13]. For any $0 < a < b < c < d$, we can easily construct an 8-gem for $(x+a)(x-a)(x+b)(x-b)(x+c)(x-c)(x+d)(x-d)$ by prepending the 4-gem (available by Lemma 5.1) for $(y-a^2)(y-b^2)(y-c^2)(y-b^2)$ with "$y \leftarrow x \times x$".
To construct a 16-gem, we seek $0 < a < b < c < d < e < f < g < h$ and an 8-gem for $(y-a^2)(y-b^2)(y-c^2)(y-d^2)(y-e^2)(y-f^2)(y-g^2)(y-h^2)$. We first develop sufficient conditions for the existence of a skew 2^n-gem for any n.

Definition 5.2. (Litter conditions.) Let T be the full ordered binary tree with 2^n leaves labeled $a_1, a_2, \ldots, a_{2^n} \in \mathbb{Z}$ in the natural order. Let each internal node in T be labeled with the product of the labels of the leaves subtended by that node. The sequence $a_1, a_2, \ldots, a_{2^n}$ *satisfies the litter conditions* if, for $1 < i < n$, each of the 2^i nodes at level i has the same *litter sum*, where the *litter sum* of a node is defined as the sum of the labels of its two children.

Example 5.3. The litter conditions are trivial for sequences of length 1 or 2. The litter conditions for the sequence a_1,a_2,a_3,a_4 are the equation $a_1+a_2=a_3+a_4$. The litter conditions for the sequence $a^2,b^2,c^2,d^2,e^2,f^2,g^2,h^2$ are

$$a^2+b^2=c^2+d^2=e^2+f^2=g^2+h^2 \text{ and } a^2b^2+c^2d^2=e^2f^2+g^2h^2. \tag{5.1}$$

Lemma 5.4. *If the sequence $a_1,a_2,\ldots,a_{2^n}\in\mathbb{Z}$ satisfies the litter conditions, then the following skew 2^n-gem c computes $p(x)=\prod_{1\le i\le 2^n}(x-a_i)$:*

$$\begin{aligned}
&y_0 \leftarrow x-a_1\\
&y_1 \leftarrow y_0\times(y_0+(a_1-a_2))\\
&y_2 \leftarrow y_1\times(y_1+(a_3a_4-a_1a_2))\\
&y_3 \leftarrow y_2\times(y_2+(a_5a_6a_7a_8-a_1a_2a_3a_4))\\
&\quad\vdots \qquad\qquad \vdots\\
&y_n \leftarrow y_{n-1}\times(y_{n-1}+(\prod_{i=2^{n-1}+1}^{2^n}a_i-\prod_{i=1}^{2^{n-1}}a_i)).
\end{aligned}$$ □

Proof. The circuit correctly computes $y_1(x)=(x-a_1)(x-a_1+(a_1-a_2))=p(x)$ when $n=1$. For the inductive step, let $A=a_1+a_2$. Then

$$\begin{aligned}
p(x) &= [(x-a_1)(x-a_2)]\ [(x-a_3)(x-a_4)]\ \cdots\ [(x-a_{2^n-1})(x-a_{2^n})]\\
&= [x^2-Ax+a_1a_2]\ [\quad]\ \cdots\ [x^2-(a_{2^n-1}+a_{2^n})x+a_{2^n-1}a_{2^n}]\\
&= [x^2-Ax+a_1a_2]\ [x^2-Ax+a_3a_4]\ \cdots\ [x^2-Ax+a_{2^n-1}a_{2^n}]\\
&= q(x^2-Ax)
\end{aligned}$$

where $q(y)=(y+a_1a_2)(y+a_3a_4)\cdots(y+a_{2^n-1}a_{2^n})$ and the third equality follows from the litter conditions $a_1+a_2=a_3+a_4=\cdots=a_{2^n-1}+a_{2^n}$. Now the litter conditions on $a_1,a_2,\ldots,a_{2^n}$ include the conditions on $a_1a_2,a_3a_4,\ldots,a_{2^n-1}a_{2^n}$, which are identical to those on $-a_1a_2,-a_3a_4,\ldots,-a_{2^n-1}a_{2^n}$. By induction, the following skew 2^{n-1}-gem c' thus computes the polynomial $q(y)$:

$$\begin{aligned}
&z_1 \leftarrow y-(-a_1a_2)\\
&z_2 \leftarrow z_1\times(z_1+(-a_1a_2-(-a_3a_4)))\\
&z_3 \leftarrow z_2\times(z_2+(a_5a_6a_7a_8-a_1a_2a_3a_4))\\
&\quad\vdots \qquad\qquad \vdots\\
&z_n \leftarrow z_{n-1}\times(z_{n-1}+(\prod_{i=2^{n-2}+1}^{2^{n-1}}a_{2i-1}a_{2i}-\prod_{i=1}^{2^{n-2}}a_{2i-1}a_{2i})).
\end{aligned}$$

So c' computes $p(x)$ when y is set to x^2-Ax. But when y is set to x^2-Ax, $y_1=y_0\times(y_0+(a_1-a_2))=(x-a_1)((x-a_1)+(a_1-a_2))=x^2-Ax-(-a_1a_2)=z_1$. Thus by inspection for $1\le i\le n$, the gate z_i in c' when y is set to x^2-Ax and the gate y_i in c compute the same polynomial in x. Hence y_n computes $p(x)$. □

We return to our quest for a 16-gem. By Lemma 5.4, any distinct squares $a^2, b^2, c^2, d^2, e^2, f^2, g^2, h^2$ satisfying (5.1) from Example 5.3 are the zeros of an 8-gem $p(y)$. In turn, each such 8-gem prepended with "$y \leftarrow x \times x$" is a 16-gem. A small computer in a few hours found several examples, such as:

Proposition 5.5. *A* 16*-gem with* 4 *additive gates exists to compute the polynomial having the 16 roots* $\{\pm 237, \pm 106, \pm 189, \pm 178, \pm 227, \pm 126, \pm 218, \pm 141\}$.

We note that Bremner [8] focusses on 16-gems and constructs two infinite families. Turning to 32-gems, we were unable to find a 16-gem having 16 distinct squares as zeros, nor to find a 32-gem by appealing to the litter conditions of a sequence of length 32 directly.
The next lemma is our tool to generate d-gems when $d \neq 2^n$:

Lemma 5.6. *Let* $h(x) \in \mathbb{Z}[x]$ *and* $m_1, m_2, \ldots, m_d \in \mathbb{Z}$. *Suppose that each one of the* d *polynomials* $h(x) - m_i$ *is computed by a gem and that no two such polynomials share a root. If* $\ell_d + \ell_{\deg(h)} \leq \ell_{d \cdot \deg(h)}$ *and for some* d*-gem* c, $f_c(y) = (y - m_1)(y - m_2) \cdots (y - m_d)$, *then there is a gem computing* $f_c(h(x))$.

Proof. For any $a \in \mathbb{Z}$, $f_c(h(a)) = 0$ iff $h(a) = m_i$ for some i iff a is a root of one of the polynomials $h(x) - m_i$. No two such polynomials share a root and, being computed by a gem, each such polynomial has distinct roots. Hence $f_c(h(x))$ is a polynomial of degree $d \cdot \deg(h)$ having $d \cdot \deg(h)$ distinct roots. To compute $f_c(h(x))$, we use at most $\ell_{\deg(h)}$ product gates to compute $h(x)$ by adding m_1 to the gem for $h(x) - m_1$, and we use another ℓ_d product gates to feed $h(x)$ into c. In total, at most $\ell_d + \ell_{\deg(h)}$ product gates are used, and this is at most $\ell_{\deg(f_c(h(x)))}$ by hypothesis. □

We illustrate the use of Lemma 5.6 in the following:

Theorem 5.7. *There exist* 36*-gems and* 54*-gems.*

Proof. To get a 36-gem, we apply Lemma 5.6 with $h(x) = (x^2 - 2s)^2$, where s is a positive integer expressible in at least 9 essentially distinct ways as a sum of two nonzero squares (such numbers abound, see [20, 28]): $s = a_1^2 + b_1^2 = a_2^2 + b_2^2 = \cdots = a_9^2 + b_9^2$. Then we let $m_i = 4s^2 - 16a_i^2 b_i^2$ and observe that for each i,

$$\begin{aligned} h(x) - m_i &= x^4 - 4sx^2 + 4s^2 - m_i \\ &= x^4 - 4(a_i^2 + b_i^2)x^2 + 16a_i^2 b_i^2 \\ &= (x^2 - 4a_i^2)(x^2 - 4b_i^2) \\ &= (x + 2a_i)(x - 2a_i)(x + 2b_i)(x - 2b_i). \end{aligned}$$

Hence the 9 polynomials $h(x) - m_i$ have pairwise disjoint sets of 4 distinct roots, and each has a gem by Lemma 5.1. Since $\ell_4 + \ell_9 = 2 + 4 \leq \ell_{36} = 6$ and Lemma 5.1 also provides a 9-gem c_9 such that $f_{c_9}(y) = \prod_{i=1}^{9}(y - m_i)$, Lemma 5.6 yields a 36-gem for $\prod_{i=1}^{9}(h(x) - m_i)$.
To get a 54-gem, we apply Lemma 5.6 with $h(x) = ((x^2 - s) \times x)^2$, where s is a positive integer expressible in at least 9 essentially distinct ways in the form $(a^2 + b^2 + ab)$. Then we let $m_i = (a_i b_i (a_i + b_i))^2$ and

observe that for each i,

$$\begin{aligned} h(x)-m_i &= ((x^2-s)\times x)^2 - m_i \\ &= ((x^2-(a_i^2+b_i^2+a_ib_i))x)^2 - (a_ib_i(a_i+b_i))^2 \\ &= ((x^2-(a_i^2+b_i^2+a_ib_i))x - a_ib_i(a_i+b_i))((x^2-(a_i^2+b_i^2+a_ib_i))x + a_ib_i(a_i+b_i)) \\ &= (x+a_i)(x+b_i)(x-a_i-b_i)(x-a_i)(x-b_i)(x+a_i+b_i). \end{aligned}$$

Hence the 9 polynomials $h(x)-m_i$ have pairwise disjoint sets of 6 distinct roots, and each has a gem by Lemma 5.1. Since $\ell_6+\ell_9 = 3+4 \le \ell_{54} = 7$ and Lemma 5.1 also provides a 9-gem c_9 such that $f_{c_9}(y) = \prod_{i=1}^{9}(y-m_i)$, Lemma 5.6 yields a 36-gem for $\prod_{i=1}^{9}(h(x)-m_i)$.
To get such an s, we apply the formula $(a^2+b^2+ab)(c^2+d^2+cd) = (ac+bd+bc)^2+(ad-bc)^2+(ac+bd+bc)(ad-bc)$ instead of the Brahmagupta-Fibonachi equation iteratively in several combinations. For example $7 = 2^2+1^2+2$ and $13 = 3^2+1^2+3$ have this form thus $((x^2-7)x)^2$ is 36 for $x = -1,-2,3,1,2,-3$, the function $((x^2-7\cdot 13)x)^2$ is 90^2 for $x = -1,-9,10,1,9,-10$ and 330^2 for $x = -5,-6,11,5,6,-11$, the function $(x^2-7\cdot 7)x$ is 120 for $x = -3,-5,8$, it is -120 for $x = 3,5,-8$ and 0 for $x = 0,-7,7$. The table

x $((x^2-7\cdot 13\cdot 19)x)^2$	$\pm 3,40,43$ $m_1 = 5160^2$	$\pm 8,37,45$ $m_2 = 13320^2$	$\pm 15,32,47$ $m_3 = 22560^2$	$\pm 23,25,48$ $m_4 = 27600^2$
x $((x^2-7^2\cdot 13^2\cdot 19)x)^2$	$\pm 13,390,403$ $m_1 = 2043210^2$	$\pm 35,378,413$ $m_2 = 54663990^2$	$\pm 70,357,427$ $m_3 = 10670730^2$	$\pm 103,335,438$ $m_4 = 15113190^2$
$\pm 117,325,442$ $m_5 = 16807050^2$	$\pm 137,310,447$ $m_6 = 18984090^2$	$\pm 182,273,455$ $m_7 = 22607130^2$	$\pm 202,255,457$ $m_8 = 23540070^2$	$\pm 225,233,458$ $m_9 = 24010650^2$

shows the values for constructing the gems in Figure 2. Note that the product $7\cdot 13\cdot 19\cdot 31$ respectively. $7^2\cdot 13\cdot 19$ would already be sufficient for the construction of a 42-gem respectively. 36-gem and produce smaller numbers there. □

Multiplying our 54-gem with $(x-a_{55})$ leads to the 55-gem $f_{c_9}(h(x))\cdot(x-z_{55})$ which is the highest which we found. We have filled the rest of Figure 2 largely by trying to combine the mentioned methods along possible shortest addition chains for the degree d. Many cases remain open.

6 Conclusion

The following heuristic for factoring a $2n$-bit integer $N = pq$, for primes p and q of comparable size, is inspired by Lipton [16]:

- assume distinct $a_i \in \mathbb{Z}$ and a circuit c computing $f_c(x) = \prod_{i=1}^{2^n}(x-a_i) \in \mathbb{Z}[x]$
- pick $a \in \{0,\ldots,N-1\}$ at random
- compute $d = f_c(a)$ modulo N by evaluating each gate in c modulo N
- output $\gcd(d,N)$.

This is merely a heuristic because its success probability depends on the distribution of the 2^n integers $(a-a_i)$ modulo N. If this is close to uniform, then indeed $Prob[1 < \gcd(d,N) < N]$ is constant. Of course the heuristic runs in time polynomial in the number of bits required to represent c, and the $\tau_{\mathbb{Z}}$-conjecture claims that this number is exponential in n.

Here we introduced *gems*. These are circuits that use an almost optimal number of $\times$-gates to compute, from $\{x\} \cup \mathbb{Z}$, polynomials that factor completely over $\mathbb{Z}$ with distinct roots. A 2^n-gem could thus serve in the heuristic above if its inputs modulo N can be computed in time polynomial in $n = O(\log N)$. But the τ-conjecture [21, 1] further implies prohibitive growth of the number of bits required to express the integer inputs to such gems.

We exhibited d-gems over $\mathbb{R}$ for every d. But the d-gems we care about (over $\mathbb{Z}$) are elusive. Even more elusive are *skew* d-gems. In particular, constructing skew 32-gems would yield new solutions to the Prouhet-Tarry-Escott problem. These new Prouhet-Tarry-Escot solutions would even fulfill additional conditions. Yet skew 2^n-gems for any $n \geq 5$ cannot currently be ruled out.

Here we constructed d-gems for several d up to $d = 55$. We proved that any skew 2^n-gem requires n additive gates. We showed that for $d \leq 71$, no circuit can compute a degree-d polynomial using fewer than ℓ_d product gates. Hence all the gems constructed so far are $\times$-optimal, and our skew 2^n-gems for $n \leq 4$ are also $\{+,-\}$-optimal among skew gems.

We note that the $\tau_{\mathbb{Z}}$-conjecture is implied by the stronger L-conjecture of Bürgisser [9], which states that for some β and any d, any polynomial $f(x) \in \mathbb{Q}[x]$ has at most $(L(f)+d)^{\beta}$ irreducible factors of degree d or less, where $L(f)$ is the size of a smallest $\{+,-,\times,\div\}$-circuit computing $f(x)$ from $\{x\} \cup \mathbb{Q}$ (see also [10] and [19]). Hence L-conjecture $\Rightarrow$ $\tau_{\mathbb{Z}}$-conjecture $\Rightarrow$ τ-conjecture, and the existence of d-gems for infinitely many d, disproving the $\tau_{\mathbb{Z}}$-conjecture, would also disprove the L-conjecture. To disprove the weaker τ-conjecture, an infinite family of d-gems d would need to fulfill further conditions such as polynomial constructivity of the integer constants involved.

The difficulty of constructing d-gems for $d = 32$ or $d > 55$ can be argued as supporting the L-conjecture and the $\tau_{\mathbb{Z}}$-conjecture. On the other hand, the fact that skew 2^n-gems for $n > 4$ would yield new PTE solutions can (optimistically) be interpreted as an indication that skew 2^n-gems are too demanding to be of any use towards resolving the conjectures. More realistically, not being able to rule out even extreme counter-examples to these conjectures suggest that we are still very far from resolving them.

Our main open question is essentially the challenge we started with: do d-gems exist for infinitely many d? A concrete step would be to undertake a search for d-gems of every type for small values of d, extending the systematic approach used by Bremner to study normalized 16-gems [8]. This would add entries to Figure 2. Hopefully it could help finding skew 2^n-gems for $n = 5$. As seen above, resolving the existence question for skew 2^n-gems seems like a natural baby step towards resolving the $\tau_{\mathbb{Z}}$-conjecture. But even this step would seem to break ground from a number-theoretic perspective.

Acknowledgments. We are grateful to Andrew Granville for insights and for noticing the connection between gems and the Prouhet-Tarry-Escott problem, to Allan Borodin for helpful suggestions and for bringing up Strassen's work, and to Peter Hauck for pointing out the work by Dilcher. We thank Yara Elias, Andreas Krebs, Klaus-Jörn Lange, Ken Regan and the referees for their useful comments.

References

[1] L. BLUM, F. CUCKER, M. SHUB, AND S. SMALE: *Complexity and Real Computation*. Springer-Verlag, 1997. 2, 17

[2] A. BORODIN AND S. COOK: On the number of additions to compute specific polynomials. *SIAM J. Comput.*, 5(1):146–157, 1976. 10, 12

[3] A. BORODIN AND R. MOENCK: Fast modular transforms. *J. Comput. Syst. Sci*, 8(3):366–386, 1974. 2

[4] A. BORODIN AND I. MUNRO: *The computational complexity of algebraic and numeric problems*. Elsevier Computer Science Library, Theory of Computation Series. American Elsevier, 1975. 5

[5] P. BORWEIN: *Computational excursions in analysis and number theory*. Canadian Math. Soc. Books in Mathematics. Springer-Verlag, 2002. 8

[6] P. BORWEIN AND C. INGALLS: The Prouhet-Tarry-Escott problem revisited. *Enseign. Math.*, 40:3–27, 1994. 3, 7, 8, 9

[7] P. BORWEIN, P. LISONEK, AND C. PERCIVAL: Computational investigations of the Prouhet-Tarry-Escott problem. *Math. Comput*, 72(244):2063–2070, 2003. 7

[8] A. BREMNER: When can $(((X^2-P)^2)-Q)^2-R)^2-S^2$ split into linear factors? *Experimental Mathematics*, 17(4):385–390, 2008. 2, 6, 15, 17

[9] P. BÜRGISSER: On implications between P-NP-hypotheses: Decision versus computation in algebraic complexity. In *Mathematical Foundations of Computer Science*, volume 2316 of *LNCS*, pp. 3–17. Springer, 2001. 10, 17

[10] Q. CHENG: Straight line programs and torsion points on elliptic curves. *Comput. Complex.*, 12(3-4):150–161, 2004. 17

[11] R. CRANDALL: *Topics in Advanced Scientific Computation*. TELOS, the Electronic Library of Science, Springer-Verlag, New York, 1996. 1, 2, 6

[12] R. CRANDALL AND C. POMERANCE: *Prime Numbers: A computational perspective*. Springer-Verlag, Berlin-Heidelberg-New York-Barcelona-Hong Kong-London-Milan-Paris-Singapur-Tokyo, 2001. 2, 6

[13] K. DILCHER: Nested squares and evaluations of integer products. *Experimental Mathematics*, 9(3):369–372, 2000. 1, 6, 13

[14] H. DOLWART AND O. BROWN: The Tarry-Escott problem. In *Proc. Amer. Math. Soc.*, volume 44, pp. 613–626, 1937. 8

[15] D. KNUTH: *The Art of Computer Programming, Vol. 2*. Addison-Wesley, Reading, 1981. 5

[16] R. LIPTON: Straight-line complexity and integer factorization. In *ANTS: 1st International Algorithmic Number Theory Symposium (ANTS)*, volume 877 of *LNCS*, pp. 71–79, Berlin, 1994. Springer-Verlag. 2, 16

[17] M. CLAUSEN P. BÜRGISSER AND M. SHOKROLLAHI: *Algebraic Complexity Theory*. volume 315 of *Grundlehren der mathematischen Wissenschaften*. Springer, 1997. 10, 12

[18] M. PATERSON AND L. STOCKMEYER: On the number of nonscalar multiplications necessary to evaluate polynomials. *SICOMP: SIAM Journal on Computing*, 2(1):60–66, 1973. 3, 5, 6

[19] M. ROJAS: A direct ultrametric approach to additive complexity and the Shub-Smale tau conjecture, 2001. ⟨http://arxiv.org/abs/math/0304100⟩. 12, 17

[20] K. H. ROSEN: *Elementary Number Theory and its Applications, 3rd ed.* Addison-Wesley, 1993. 15

[21] M. SHUB AND S. SMALE: On the intractability of Hilbert's Nullstellensatz and an algebraic version of "NP≠P". A celebration of John F. Nash, Jr. *Duke Mathematical Journal*, 81:47–54, 1996. 2, 17

[22] C. SHUWEN: The PTE problem, 2001. ⟨http://euler.free.fr/eslp/TarryPrb.htm⟩. 8, 9

[23] S. SMALE: Mathematical problems for the next century. In V. ARNOLD, M. ATIYAH, P. LAX, AND B. MAZUR, editors, *Mathematics: Frontiers and Perspectives 2000*. AMS, 2000. 2

[24] V. STRASSEN: Einige Resultate über Berechnungskomplexität. *Jahresbericht der Deutsche Math. Verein.*, 78:1–8, 1976. 2

[25] J. V. USPENSKY: *Theory of Equations*. McGraw Hill, 1948. 8

[26] J. VON ZUR GATHEN AND J. GERHARD: *Modern Computer Algebra, 2nd edition*. Cambridge University Press, Cambridge, 2003. 2

[27] E. WEISSTEIN: The Prouhet-Tarry-Escott problem. from MathWorld–A Wolfram Web Resource [online], 2009. ⟨http://mathworld.wolfram.com/Prouhet-Tarry-EscottProblem.html⟩. 8

[28] WIMS: ⟨http://wims.unice.fr/wims/en_tool∼number∼twosquares.en.html ⟩. 15

AUTHORS

Bernd Borchert
WSI, Universität Tübingen, Sand 13, 72076 Tübingen, Germany
borchert@informatik.uni-tuebingen.de
http://www-fs.informatik.uni-tuebingen.de/mitarbeiter/borchert/

Pierre McKenzie
DIRO, Univ. de Montréal, C.P. 6128 Centre-Ville, Montréal (Qc), H3C 3J7 Canada
mckenzie@iro.umontreal.ca
http://www.iro.umontreal.ca/~mckenzie/

Klaus Reinhardt
WSI, Universität Tübingen, Sand 13, 72076 Tübingen, Germany
reinhard@informatik.uni-tuebingen.de
http://www-fs.informatik.uni-tuebingen.de/mitarbeiter/reinhard/

Appendix: Proof of Lemma 2.1

The idea of the proof is that if the circuit does not produce the degrees according to the addition chain, it must use a cancellation in the degrees of two intermediate results. This would require that some degree would have to be produced by two different $\times$-gates. A case distinction on their input degrees shows that this can not lead to a smaller number of $\times$-gates.

Let a circuit c with a minimum number $c_\times < l_d$ of $\times$-gates compute a polynomial of degree d. We must prove that $d \le 2^{l_d-3}$. Fix a topological ordering of the circuit gates. Then every additive gate computes a linear combination $\sum \alpha_i q_i$ of the polynomials q_i with degree d_i computed by the $\times$-gates that occur before it (we let $q_{-1} = 1$ and $q_0 = x$ and allow the sum to run from $i = -1$). At the expense of an increase in the number of additive gates, we assume a normal form in which every $\sum \alpha_i q_i$ is obtained by a sequence of additive gates that start with the constant 0 and add or subtract the desired q_i, one at a time, in descending order of degrees and without ever decreasing an $|\alpha_i|$.

If we assume that the degree of an additive gate is always the maximum of its two inputs (which would for example be the case if different q_i always have different degrees d_i), then each degree of an additive gate is the degree of the first (and thus of highest degree) $\times$-gate in such a sequence of additive gates. This means the degrees would form an addition chain and thus $c_\times \ge l_d$ since the degree of any product is the sum of the degrees of its factors.

Since $c_\times < l_d$, in the contrary, a new degree must have been reached in an additive (respectively subtractive) gate which can only happen if the so far highest degree was canceled[2]. In other words, there must be in the ordering a first additive gate g that computes a polynomial $\sum \gamma_i q_i = (\sum \alpha_i q_i) \pm q_k$ for some degrees d_l and d_k such that $d_k = \deg(\sum \alpha_i q_i) = \deg q_k > \deg(\sum \gamma_i q_i) = d_l$ and let the $\times$-gates g_1 and g_2 compute q_{k-1} and q_k.

Given our normal form, we can assume w.l.o.g. an ordering in which the degrees of the intermediate results did not decrease until d_k. Let $c_\times^{low} = k-2$ be the number of the $\times$-gates occurring before g_1 in the ordering. Furthermore, q_i can have degree at most 2^i for $i \le c_\times^{low}$ since each $\times$-gate can at most double the degree. Thus $d_k = d_{k-1} = \deg(\alpha_{k-1} q_{k-1}) \le 2^{1+c_\times^{low}}$. Let furthermore $c_\times^{up} = c_\times - c_\times^{low} - 2$ be the number of $\times$-gates occurring after g_2. Thus we can already estimate $d < d_k 2^{c_\times^{up}} \le 2^{c_\times^{up}+1+c_\times^{low}} \le 2^{c_\times - 1} \le 2^{l_d-2}$. But we want a better estimate.

We can assume that g_1 and g_2 are the only $\times$-gates computing polynomials of degree d_k, otherwise the inequality chain for d above would extend to show $d \le 2^{l_d-3}$. In the same way we can assume that there are no other two $\times$-gates computing polynomials of the same degree. This means that all other such cancellations of the degree can only occur in a sequence of additive gates after adding q_{k-1} and q_k and w.l.o.g. the next cancellation must be by adding some q_l of degree d_l. If q_l is not used in the subcircuit for calculating q_{k-1} or q_k, that means d_l does not occur in the corresponding addition chain to d_k, then no further cancellations (except with the same ratio between q_{k-1} and q_k) occur.

The gate g outputs $p(x) = \alpha q_{k-1}(x) + \beta q_k(x)$ for some $\alpha, \beta \in \mathbb{Z}$, where $d_l = \deg p < \deg q_{k-1} = \deg q_k = d_k$. Furthermore all other additive gates in which degree d_k is cancelled produce just a multiple of g.

Let the inputs to gate g_1 have the degrees d_m and d_n, so that $d_k = d_m + d_n$. Assume w.l.o.g. $d_m \ge d_n$. If

[2]For example $(x^2+x)^2 - (x^2)^2 = x^4 + 2x^3 + x^2 - x^4 = 2x^3 + x^2$ reaches degree 3 only in the subtraction where the highest degree 4 is canceled.

the inputs to gate g_2 also happen to have the degrees d_m and d_n, then

$$\begin{aligned}
q_{k-1} &= [a_m q_m(x) + a_{m-1}q_{m-1}(x) + \ldots + a_j q_j(x) + \ldots] \;\times\; [b_n q_n(x) + b_{n-1}q_{n-1}(x) + \ldots + b_i q_i(x) + \ldots] \\
&= [a_m(q_m(x) + \frac{a_{m-1}}{a_m}q_{m-1}(x) + \ldots) + a_j q_j(x) + \ldots] \;\times\; [b_n(q_n(x) + \frac{b_{n-1}}{b_n}q_{n-1}(x) + \ldots) + b_i q_i(x) + \ldots] \\
&= [a_m P(x) + a_j q_j(x) + \ldots] \;\times\; [b_n Q(x) + b_i q_i(x) + \ldots] \\
q_k &= [a'_m q_m(x) + a'_{m-1}q_{m-1}(x) + \ldots + a'_j q_j(x) + \ldots] \;\times\; [b'_n q_n(x) + b'_{n-1}q_{n-1}(x) + \ldots + b'_i q_i(x) + \ldots] \\
&= [a'_m(q_m(x) + \frac{a'_{m-1}}{a'_m}q_{m-1}(x) + \ldots) + a'_j q_j(x) + \ldots] \;\times\; [b'_n(q_n(x) + \frac{b'_{n-1}}{b'_n}q_{n-1}(x) + \ldots) + b'_i q_i(x) + \ldots] \\
&= [a'_m P(x) + a'_j q_j(x) + \ldots] \;\times\; [b'_n Q(x) + b'_i q_i(x) + \ldots]
\end{aligned}$$

where $d_j < d_m$ is maximal such that $\frac{a_m}{a_j} \neq \frac{a'_m}{a'_j}$ and $d_i < d_n$ is maximal such that $\frac{b_n}{b_i} \neq \frac{b'_n}{b'_i}$. Then

$$\begin{aligned}
q_{k-1} &= \overbrace{a_m b_n P\cdot Q}^{\deg = d_m+d_n} + \overbrace{a_m b_i P\cdot q_i(x)}^{\deg = d_m+d_i} + \overbrace{b_n a_j Q\cdot q_j(x)}^{\deg = d_n+d_j} + \overbrace{a_j b_i q_j(x) q_i(x) + \ldots}^{\deg \leq d_j+d_i} \\
q_k &= a'_m b'_n P\cdot Q + a'_m b'_i P\cdot q_i(x) + b'_n a'_j Q\cdot q_j(x) + a'_j b'_i q_j(x) q_i(x) + \ldots \\
p &= (\alpha a_m b_i + \beta a'_m b'_i)\cdot P\cdot q_i(x) + (\alpha b_n a_j + \beta b'_n a'_j)\cdot Q\cdot q_j(x) + (\alpha a_j b_i + \beta a'_j b'_i) q_j(x) q_i(x) + \ldots
\end{aligned}$$

Note that $\alpha a_m b_n = -\beta a'_m b'_n$ and $\frac{b_n}{b_i} \neq \frac{b'_n}{b'_i}$ imply $\alpha a_m b_i + \beta a'_m b'_i \neq 0$. By a similar reasoning, $\alpha b_n a_j + \beta b'_j a'_j \neq 0$. Hence $q_l = \deg p = \max\{d_m + d_i, d_n + d_j\}$.

If no further cancellations occur, only the degree matters in the upper part of the circuit and we can modify the circuit, preserving $c_\times$ and the degree d of the final polynomial computed, in a way such that no cancellations of degree $\geq d_k$ occur: First we remove the $\times$-gate g_2 that computed q_k and replace it by g_1. Then we replace the additive gate g that computed p by a $\times$-gate computing a polynomial of degree $\deg p = \max\{d_m + d_i, d_n + d_j\}$.

If degree d_l cancels in a later addition, then some q_l with $l < k$ must have been added, this means degree q_l was already available. Furthermore the new degree after this cancellation is again the sum of two available degrees. Cancellations of available degrees may continue until at most one new degree is reached, which must be the sum of two available degrees and the corresponding gate can be replaced by one $\times$-gate. In this way, no new cancellation of degree $\geq d_k$ is introduced past gate g in the ordering (with the new gates included) and in particular, the final polynomial computed by the new circuit still has degree d. But we have removed the cancellation of degree d_k that occurred at g in the old circuit. This completes the proof of the present case since repeating this process will eliminate all the cancellations, contradicting $c_\times < \ell_d$.

Now let g_2 happen to have input degrees $(d_{m'}, d_{n'}) \neq (d_m, d_n)$, yet $d_k = d_{m'} + d_{n'} = d_m + d_n$. Then w.l.o.g. $d_{m'} > d_m > d_n > d_{n'}$. As above, there must be two $\times$-gates g_3 and g'_3 producing the degrees d_m and $d_{m'}$ directly. Thus four gates g_1, g_2, g_3, g'_3 are not counted in $c_\times^{up} + c_\times^{low} = c_\times - 4$ where $c_\times^{low} = k-2$ is now the number of the $\times$-gates occurring before g_3 and g'_3 in the ordering. Thus we can estimate $d < d_k 2^{c_\times^{up}} < d_m 2^{c_\times^{up}+1} \leq 2^{c_\times^{up}+2+c_\times^{low}} \leq 2^{c_\times - 2} \leq 2^{l_d - 3}$.