The binomial case of Graham's conjecture on polynomial representations with prescribed sum of reciprocals

Wouter van Doorn

Abstract

Let $f(x)$ be an integer polynomial with positive leading coefficient and such that $\gcd(f(1), f(2), \dots) = 1$. Graham conjectured in [1] that for every positive rational α an $n_0 = n_0(f(x), \alpha)$ exists such that for all $n \geq n_0$ distinct positive integers $a_1, \dots, a_r$ exist with $f(a_1) + \ldots + f(a_r) = n$ and $\frac{1}{a_1} + \ldots + \frac{1}{a_r} = \alpha$. For a polynomial $f(x) = ax^c + b$ with positive coefficients and any $\alpha > 0$, we prove that the existence of such an n_0 can in principle be reduced to a finite computation. For $\alpha = 1$ we thereby confirm Graham's conjecture for thousands of linear and quadratic polynomials.

1 Introduction and notation

For a polynomial $f(x) : \mathbb{Z} \to \mathbb{Z}$ and a positive rational α, let $n_0 = n_0(f(x), \alpha)$ be the smallest integer (if it exists) such that, for all $n \geq n_0$, distinct positive integers $a_1, \dots, a_r$ exist with $f(a_1) + \ldots + f(a_r) = n$ and $\frac{1}{a_1} + \ldots + \frac{1}{a_r} = \alpha$. After showing in [1] that $n_0(x, 1)$ is equal to 78 and, more generally, that $n_0(x, \alpha)$ exists for all $\alpha > 0$, Graham furthermore conjectured that $n_0(f(x), \alpha)$ always exists for all positive rationals α, as long as $f(x)$ has positive leading coefficient and $\gcd(f(1), f(2), \dots) = 1$. For $\alpha = 1$ this is listed as problem #283 at [2]. However, besides near-optimal upper bounds for $n_0(x, \alpha)$ that were given in [3] by the author, the only other case that has been solved is $f(x) = x^2$ for $\alpha = 1$; Alekseyev proved in [4] that $n_0(x^2, 1)$ is equal to 8543.

In this paper we will add to this knowledge, and show that if $f(x) = ax^c + b$ is a binomial with positive leading and constant terms, then for all $\alpha > 0$, a finite computation is potentially sufficient to find the value of $n_0(f(x), \alpha)$. For $a = 1$, $b = 0$ such a result was essentially already obtained in [4] and [5]. As a proof of concept we will then carry out this computation for some specific rationals and (mostly linear) polynomials and, for example, show $n_0(x + b, 1) \leq 172 + 10b$ for all positive integers $b \leq 10^4$.

As for our notation, we use the greek letters α and β to denote positive rationals. The letters A, A_i, B refer to sets of positive integers, while all other letters denote positive integers, unless stated otherwise. For a function $f(x)$ and a set $A = \{a_1, \dots, a_r\}$ of positive integers, $\sum f(A)$ denotes $f(a_1) + \ldots + f(a_r)$, with in particular $\sum A^{-1} = \frac{1}{a_1} + \ldots + \frac{1}{a_r}$ and $\sum A^c = a_1^c + \ldots + a_r^c$. Finally, for a positive integer m, we write mA for the set $\{ma_1, \dots, ma_r\}$.

2 Main theorem

For all polynomials $f(x) = ax^c + b$ with positive (coprime) coefficients and all positive rationals α, our main result will essentially reduce the existence and value of $n_0(f(x), \alpha)$ to a finite computation. This theorem will be modeled after Theorem 2 in [5], so familiarity with its statement and proof is definitely helpful, but by no means required[1].

Let S be a set of positive rationals such that for all $\alpha \in S$ we want to prove that for all large enough n, a set A exists for which $\sum f(A) = n$ and $\sum A^{-1} = \alpha$. To prove this, we choose a suitable property Q that a set can have (e.g. not containing certain integers), a fixed cardinality s divisible by a, an integer $m \geq 2$ coprime to a and, for every $\alpha \in S$, rationals $\beta_1, \ldots, \beta_{m^c}$ and sets of positive integers $A_1, \ldots A_{m^c}$ with the following six properties:

1. For all i, $\beta_i \in S$.
2. For all i, $\sum A_i^c \equiv i \pmod{m^c}$.
3. For all i, $|A_i| = s$.
4. For all i, $\alpha = \sum A_i^{-1} + \frac{\beta_i}{m}$.
5. For all i and for every set B with property Q, we have $A_i \cap mB = \emptyset$.
6. For all i and for every set B with property Q, the union $A_i \cup mB$ also has property Q.

We then have the following theorem.

Theorem 1. *Assume that for all $\alpha \in S$ such β_i and A_i exist, satisfying the above six properties. With M defined as $\max_{i,\alpha} \sum A_i^c$, further assume that positive integers X and r exist with*

$$(m^c - 1)b(r(m^c - 1) - s) \geq a(M - \min_{i,\alpha} \sum A_i^c), \tag{1}$$

and such that for all $\alpha \in S$ and all n with $X \leq n \leq m^c X + aM$ and $n \equiv br \pmod{a}$, sets A exist with property Q, all with the same cardinality r, and for which $\sum f(A) = n$ and $\sum A^{-1} = \alpha$. Then for all $\alpha \in S$ and all $n \geq X$ with $n \equiv br \pmod{a}$, a set A with property Q exists for which $\sum f(A) = n$ and $\sum A^{-1} = \alpha$.

[1] For those indeed familiar with it, the most important difference with the theorem we are about to prove, is that here we need the sets under consideration to all have the same cardinality, due to the non-zero constant term b. However, for any set $A = \{a_1, \ldots, a_r\}$ with cardinality r, we always have $\sum f(A) \equiv br \pmod{a}$. A further annoyance is therefore that we have to deal with all residue classes modulo a separately.

Proof. Since $\sum f(A) = n$ is, by assumption, solvable for some $n < X + a$ and for some set A with cardinality r, we deduce $X > br - a$. With Y defined as $m^c X - b(r(m^c - 1) - s) + aM$, one can now check the inequalities $X < Y < m^c X + aM$ using $X > br - a$, $M \geq m^c$ and (1). By induction it is then sufficient to show that for all $\alpha \in S$ and all n with $Y \leq n \leq m^c Y + aM$ and $n \equiv br \pmod{a}$ sets A with property Q exist, all with the same cardinality $r' := r + s$, and for which $\sum f(A) = n$ and $\sum A^{-1} = \alpha$.

So let $\alpha \in S$ be given, assume $Y \leq n \leq m^c Y + aM$ with $n \equiv br \pmod{a}$, and let i be such that $n \equiv a \sum A_i^c - b(r(m^c - 1) - s) \pmod{m^c}$. This implies that an integer n' exists with which we can write $n = m^c n' + a \sum A_i^c - b(r(m^c - 1) - s)$. With this value of n', we claim the inequalities $X \leq n' \leq m^c X + aM$. Indeed, on the one hand we have

$$\begin{aligned} n' &= \frac{n - a\sum A_i^c + b(r(m^c - 1) - s)}{m^c} \\ &\geq \frac{Y - a\sum A_i^c + b(r(m^c - 1) - s)}{m} \\ &= \frac{m^c X - b(r(m^c - 1) - s) + aM - a\sum A_i^c + b(r(m^c - 1) - s)}{m^c} \\ &\geq X. \end{aligned}$$

While on the other hand we have

$$\begin{aligned} n' &= \frac{n - a\sum A_i^c + b(r(m^c - 1) - s)}{m^c} \\ &\leq \frac{m^c Y + aM - a\sum A_i^c + b(r(m^c - 1) - s)}{m^c} \\ &= \frac{m^c(m^c X - b(r(m^c - 1) - s) + aM) + a(M - \sum A_i^c) + b(r(m^c - 1) - s)}{m^c} \\ &= \frac{m^{2c} X + am^c M + a(M - \sum A_i^c) - (m^c - 1)b(r(m^c - 1) - s)}{m^c} \\ &\leq m^c X + aM. \end{aligned}$$

Here, the last inequality uses (1). Moreover, we have

$$\begin{aligned} n' &= \frac{n - a\sum A_i^c + b(r(m^c - 1) - s)}{m^c} \\ &\equiv \frac{br + br(m^c - 1)}{m^c} \pmod{a} \\ &\equiv br \pmod{a}. \end{aligned}$$

Since $X \leq n' \leq m^c X + aM$ and $n' \equiv br \pmod{a}$, we may apply the induction hypothesis on n', so let B be a set with property Q for which $|B| = r$, $\sum f(B) = n'$ and $\sum B^{-1} = \beta_i$. Then we claim that the union $A := A_i \cup mB$ has property Q, $|A| = r'$, $\sum f(A) = n$ and $\sum A^{-1} = \alpha$.

Indeed, A_i and mB are disjoint by the fifth property above, implying $|A| = |A_i| + |mB| = r'$, while A has property Q by the sixth property. We get that the sum $\sum A^{-1}$ of reciprocals is equal to $\sum A_i^{-1} + \sum (mB)^{-1} = \sum A_i^{-1} + \frac{\beta_i}{m} = \alpha$ by the fourth property. And finally, the sum $\sum f(A)$ is equal to

$$\begin{aligned}\sum f(A) &= \sum f(A_i) + \sum f(mB) \\ &= a \sum A_i^c + bs + m^c \sum f(B) - br(m^c - 1) \\ &= a \sum A_i^c + m^c n' - b(r(m^c - 1) - s) \\ &= n.\end{aligned}$$

□

3 Applying our main result

The goal of this section is to apply Theorem 1 in order to conclude that with $\alpha = 1$, Graham's conjecture holds for many linear and quadratic polynomials, at least when a is small. To do this, we start with one crucial but nearly trivial remark: by defining $g(x) := f(x) + 1$, we see that $\sum f(A) = n$ holds for some set A with cardinality r if, and only if, $\sum g(A) = n + r$.

With the above remark in mind, assume that all conditions of Theorem 1 hold, but with the interval $X \leq n \leq m^c X + aM$ slightly widened to $X \leq n \leq m^c X + (m^c - 1)r + aM$. This implies in particular that all conditions still hold with f replaced by g, and the interval replaced by $X + r \leq n \leq m^c(X + r) + aM$. And this in turn implies that for all $\alpha \in S$ and all $n \geq X + r$ with $n \equiv (b+1)r \pmod{a}$, a set A exists for which $\sum g(A) = n$ and $\sum A^{-1} = \alpha$.

More generally, assume that all conditions of Theorem 1 hold, but with the interval $X \leq n \leq m^c X + aM$ widened to $X \leq n \leq m^c X + l(m^c - 1)r + aM$ for some $l \geq 1$. Then for all i with $0 \leq i \leq l$, Graham's conjecture holds for all $n \equiv (b+i)r \pmod{a}$ and all polynomials of the form $f(x) = ax^c + b + i$.

In practice this means that, in order to deduce Graham's conjecture for $f(x) = ax^c + b'$ for all $b' \geq b$ up to some reasonably large threshold, it is often sufficient to find suitable m, r, s, X, β_i, A_i for all α and for all residue classes modulo a, but only for the polynomial $ax^c + b$. To make all of this concrete, let us start off with linear polynomials with leading coefficient equal to 1.

Theorem 2. *Let $f(x)$ be equal to $x+b$ where b is a positive integer with $b \leq 10^4$. Then for all $n \geq 172 + 10b$ there is a set A with $\sum f(A) = n$ and $\sum A^{-1} = 1$.*

Proof. By Theorem 1 we have to define multiple sets, integers and rationals, so let us do so. We choose $S = \{\frac{4}{5}, 1\}$, we let Q be the property of not intersecting $\{13, 23, 57, 65, 85, 115, 117\}$, and we set $r = 10$, $s = 7$, $m = 2$ and $X = 182$. For $\alpha = \frac{4}{5}$ we set $\beta_1 = \frac{4}{5}$, $A_1 = \{3, 27, 119, 135, 170, 234, 273\}$, $\beta_2 = 1$, and $A_2 = \{5, 19, 46, 105, 114, 190, 483\}$. For $\alpha = 1$ we set $\beta_1 = \frac{4}{5}$,

$A_1 = \{3, 5, 21, 130, 230, 273, 299\}$, $\beta_2 = 1$, and $A_2 = \{3, 9, 39, 91, 105, 130, 585\}$.

One can then check that all six properties from Theorem 1 hold. Moreover, $M = 962$ and $M - \min_{i,\alpha} \sum A_i = 962 - 961 = 1$, so (1) (i.e. the inequality $(m^c-1)b(r(m^c-1)-s) \geq a(M - \min_{i,\alpha} \sum A_i^c))$ simplifies to $3b \geq 1$. It is by the remarks at the start of this section now sufficient to find sets $A = \{a_1, \ldots a_r\}$ with $(a_1+1) + \ldots + (a_r+1) = n$ and $\frac{1}{a_1} + \ldots \frac{1}{a_r} = \alpha$ with property Q for all $\alpha \in S$ and all n with $X \leq n \leq 2X + 10^4 r + M = 101326$. One can find these sets on the author's GitHub page[2]. $\square$

When $\max(a, c) > 1$, the computational requirements in order to apply Theorem 1 quickly add up. If $c > 1$, then X, m^c and M all generally increase substantially, which greatly increases the length of the interval $X \leq n \leq m^c X + aM$ one has to check. And if $a > 1$, then we need to check all residue classes modulo a separately in order to prove that all sufficiently large n are representable. Every residue class requires a different value of r however, and increasing r magnifies the search space dramatically. That being said, as a proof of concept we tackled $(a, c) = (5, 1)$ and $(a, c) = (1, 2)$. The reader is cordially invited to write more efficient code (or to perhaps tweak the requirements of Theorem 1) in order to cover (way) more polynomials, but for now we shall leave it at the two cases we mentioned.

Theorem 3. *If $f(x) = 5x + b$ with $1 \leq b \leq 750$, then for all $n \geq 20000$ there is a set A with $\sum f(A) = n$ and $\sum A^{-1} = 1$.*

Theorem 4. *If $f(x) = x^2 + b$ with $50 \leq b \leq 2000$, then for all $n \geq 50000$ there is a set A with $\sum f(A) = n$ and $\sum A^{-1} = 1$.*

Proof of Theorem 3. We choose $S = \{\frac{4}{5}, 1\}$, we let Q be the property of not intersecting $\{17, 117, 119, 133, 143\}$, and we set $s = 10$, $m = 2$ and $X = 2265$. As we mentioned before, we need a different value of r for every residue class modulo a, and we choose $r \in \{11, 12, 13, 14, 15\}$.

For $\alpha = \frac{4}{5}$ we set $\beta_1 = 1$, $A_1 = \{9, 13, 39, 45, 63, 65, 77, 105, 165, 234\}$, $\beta_2 = \frac{4}{5}$, and $A_2 = \{5, 15, 33, 34, 35, 63, 99, 105, 187, 238\}$. For $\alpha = 1$ we set $\beta_1 = 1$, $A_1 = \{5, 9, 11, 33, 45, 65, 91, 105, 165, 286\}$, $\beta_2 = \frac{4}{5}$, and finally $A_2 = \{3, 13, 15, 21, 39, 63, 65, 95, 234, 266\}$. Once again one can then check that all six properties from Theorem 1 hold, with $M = 815$ and $M - \min_{i,\alpha} \sum A_i = 1$.

On the author's GitHub page one can now find $A = \{a_1, \ldots a_r\}$ with $(5a_1+1) + \ldots + (5a_r+1) = n$ and $\frac{1}{a_1} + \ldots \frac{1}{a_r} = \alpha$ with property Q for all $\alpha \in S$ and all n with $X \leq n \leq 2X + 750r + 5M \leq 20000$. Since (1) is satisfied for $b \geq 2$, this finishes the proof for $b \geq 2$.

For $b = 1$, notice that by applying the induction step once, we go from the interval $X \leq n \leq 20000$ with $11 \leq r \leq 15$ to (an interval containing) $2X + M \leq$

[2]See https://github.com/Woett/Binomial-representation-data.

$n \leq 40000$ with $r' = r + s \geq 21$. This interval is still sufficiently large for the purposes of Theorem 1, and with this new value of r, (1) is also satisfied for $b = 1$. $\square$

Proof of Theorem 4. We choose $S = \{\frac{4}{5}, 1, \frac{6}{5}\}$ this time around, we let Q be the property of not intersecting $\{13, 17, 19, 39\}$, $r = 11$, $s = 10$, $m = 2$ and $X = 28000$.

For $\alpha = \frac{4}{5}$ we set $\beta_1 = \beta_3 = 1$ and $\beta_2 = \beta_4 = \frac{4}{5}$ with the following sets:

$$\begin{aligned} A_1 &= \{11, 21, 26, 33, 35, 55, 63, 77, 99, 143\} \\ A_2 &= \{7, 9, 23, 45, 55, 63, 69, 77, 99, 115\} \\ A_3 &= \{9, 26, 34, 35, 45, 51, 63, 65, 78, 153\} \\ A_4 &= \{5, 15, 26, 55, 63, 65, 77, 78, 99, 105\}. \end{aligned}$$

For $\alpha = 1$ we set $\beta_1 = \beta_3 = 1$ and $\beta_2 = \beta_4 = \frac{6}{5}$ with the following sets:

$$\begin{aligned} A_1 &= \{5, 9, 15, 21, 55, 77, 78, 91, 99, 105\} \\ A_2 &= \{7, 9, 23, 45, 55, 63, 69, 77, 99, 115\} \\ A_3 &= \{5, 7, 26, 35, 38, 57, 65, 78, 95, 133\} \\ A_4 &= \{5, 15, 26, 55, 63, 65, 77, 78, 99, 105\}. \end{aligned}$$

Finally, for $\alpha = \frac{6}{5}$ we set $\beta_1 = \beta_3 = 1$ and $\beta_2 = \beta_4 = \frac{4}{5}$ with the following sets:

$$\begin{aligned} A_1 &= \{3, 7, 15, 21, 34, 35, 51, 63, 105, 153\} \\ A_2 &= \{3, 5, 7, 21, 55, 65, 77, 91, 99, 117\} \\ A_3 &= \{3, 5, 26, 34, 35, 51, 63, 65, 78, 153\} \\ A_4 &= \{3, 5, 9, 21, 26, 45, 65, 77, 78, 165\}. \end{aligned}$$

We now have $M = 46725$ and $M - \min_{i,\alpha} \sum A_i = 3346$.

On the author's GitHub page one can find $A = \{a_1, \ldots a_r\}$ with $(a_1^2 + 1) + \ldots + (a_r^2 + 1) = n$ and $\frac{1}{a_1} + \ldots \frac{1}{a_r} = \alpha$ with property Q for all $\alpha \in S$ and all n with $X \leq n \leq (m^c - 1)X + 2000(m^c - 1)r + M \leq 2 \cdot 10^5$. Since (1) is satisfied for $b \geq 50$, this finishes the proof. $\square$

References

[1] R.L. Graham, *A theorem on partitions*. Journal of the Australian Mathematical Society, vol. 3, 435–441, 1963. Also available here.

[2] T. F. Bloom, Erdős Problem #283, https://www.erdosproblems.com, accessed.

[3] W. van Doorn, *Partitions with prescribed sum of reciprocals: asymptotic bounds*. Available here.

[4] M. A. Alekseyev, *On Partitions into Squares of Distinct Integers Whose Reciprocals Sum to 1*. The Mathematics of Various Entertaining Subjects: Volume 3: The Magic of Mathematics, 213–221, 2019. Also available here.

[5] W. van Doorn, *Partitions with prescribed sum of reciprocals: computational results*. Available here.