**1949**

# New results related to factorials, arithmetic progressions and perfect powers

Thesis for the Degree of Doctor of Philosophy (PhD)

by Ágoston Papp
Supervisor: Dr. Lajos Hajdu

University of Debrecen
Doctoral Council of Natural Sciences and Information Technology
Doctoral School of Mathematical and Computational Sciences
Debrecen, 2024

Hereby I declare that I prepared this thesis withtin the Doctoral Council of Natural Sciences and Infromation Technology, Doctoral School of Mathematical and Computational Sciences, University of Debrecen in order to obtain PhD Degree in Natural Sciences at Debrecen University.

The results published in the thesis are not reported in any other PhD thesis.

Debrecen, 2024. .....

signature of candidate

Hereby I confirm that Ágoston Papp candidate concluded his studies with my supervision within the Constructive and Diophantine Number Theory Doctoral Program of the Doctoral School of Mathematical and Computational Sciences between 2019 and 2023. The independent studies and research work of the candidate significantly contributed to the results published in this thesis.

I also declare that the results published in this thesis are not part of any other thesis.

I support the acceptance of the thesis.

Debrecen, 2024. .....

signature of supervisor

# New results related to factorials, arithmetic progressions and perfect powers

Dissertation submitted in partial fulfilment of the requirements for the doctoral (PhD) degree in mathematics and computing

Written by Ágoston Papp certified mathematician

Prepared in the framework of the Doctoral School of Mathematical and Computational Sciences of the University of Debrecen (Constructive and Diophantine Number Theory programme)

Dissertation advisor: Dr. Lajos Hajdu

The official opponents of the dissertation:

..... .....

..... .....

The evaluation committee:

chairperson: ..... .....

members: ..... .....

..... .....

..... .....

..... .....

The date of the dissertation defence:

## Köszönetnyilvánítás

Köszönöm feleségemnek és családomnak a sok támogatást, amit nyújtottak, köszönöm diáktársaimnak, akik kutatómunkára inspiráltak, és köszönöm kiváló tanáraimnak, akik végigkísérek ezen az úton. SDG.

# Contents

# 1 Introduction

This dissertation contains new results in Diophantine number theory in three topics. The first part is about factorials and densities of certain subsets connected to the classical problem of representing factorials as sums of squares. The second part concerns the problem of common values of products of consecutive terms of arithmetic progressions and various polynomials. The third part deals with the number of powers in a fixed number of consecutive terms of arithmetic progressions. In what follows, we briefly outline the problems considered and the new results obtained. A detailed introduction to the topics, including a thorough survey of the corresponding literature, and the precise formulation of our results, together with their embedding in the literature, is given in the corresponding sections.

The first topic we consider is related to arithmetic properties of factorials. On the one hand, we are interested in the densities of sets of $n$ such that the exponents of given primes in the prime factorization of $n!$ hold certain congruence properties. On the other hand, given $M$, we investigate the behavior of the $M$-free parts of factorials. In fact we study the combination of these two properties. Here the following problem of Erdős and Graham [25] can be considered to be the starting point. Is it true that for any finite set $\{p_1, \dots, p_k\}$ of primes there exist infinitely many $n$, such that the exponents of the $p_i$ $(i = 1, \dots, k)$ in $n!$ are all even? The topic has a huge literature: this and related questions have been studied among others by Berend, Kolesnik, Luca, Stănică and many more (see e.g. [4, 5, 52, 53]). Another question, considered by Deshouillers and Luca [21] concerns the representability of factorials by sums of three squares. To attack this question, one has to make a further step: beside understanding the behavior of the exponent of 2 in $n!$, it is also necessary to describe the behavior of the odd part of $n!$. Our new results (published in [40]) yield a general step into this direction: we are able to describe the simultaneous behavior of the

exponents of $p$ in $n!$ and the $p$-free part of $n!$. We give several related theorems, among others, we are able to sharpen the corresponding result of Deshouillers and Luca [21]. To prove our theorems we need to combine various ideas of algebraic and combinatorial nature.

The second problem we discuss is about the common values of polynomials and products of consecutive terms of arithmetic progressions. Here the starting point is a classical result due to Fermat and Euler, showing that the product of four consecutive terms of an arithmetic progression (apart from trivial cases) cannot be a square. The most important theorem in the field is that of Erdős and Selfridge [27], saying that the product of consecutive positive integers is never a perfect power. This result has been generalized into many directions by several authors, including but not restricted to Shorey, Tijdeman, Saradha, Győry, Bennett, Hajdu, Pintér, Tengely, Siksek (see e.g. [2, 35, 36, 60, 61, 62, 64, 67]). The three most important directions are: omit a few terms from the block of consecutive integers; instead of consecutive integers consider consecutive terms of an arithmetic progression; instead of a perfect power $y^\ell$ on the right hand side take a polynomial $g(y) \in \mathbb{Z}[y]$. We give new results (published in [42] and [44]) which concern a common generalization of these directions, and extend many related papers from the literature. More precisely, we provide various finiteness results for the integer solutions $x, y$ of equations of the form

$$\prod_{a \in A}(x - c - ad) = g(y),$$

where $g(y) \in \mathbb{Z}[y]$, $c, d$ are rationals and $A \subseteq \{1, \dots, n\}$ for some positive integer $n$. That is, the product on the left hand side is taken over some terms of an arithmetic progression. In the special case where $g$ is a shifted power, that is of the form $g(y) = ay^\ell + b$, our results are based upon Baker's method and are effective. In case of general $g$, we use a classical theorem of Bilu and Tichy [8] and our results are ineffective. To make these powerful tools work, we need to use and

combine many ideas. Among others, we discover a deep link between the Prouhet-Terry-Escott problem and indecomposability of polynomials, which play a very important role in the proofs of our ineffective results. Finally, we mention that we also have somewhat related results in our paper [41] concerning equations of the shape $A!B! = C!$. However, we do not include these results here.

The third field we consider concerns upper bounds for the number $P_\ell(N)$ of $\ell$-th powers among $N$ consecutive terms of an arithmetic progression. Here the starting point is a conjecture of Erdős, predicting that $P_2(N) = o(N)$. This conjecture has been proved by Szemerédi [66]. A much stronger conjecture is due to Rudin, saying that $P_2(N) = O(\sqrt{N})$. This conjecture is still open, in spite of several related papers and deep results of Bombieri, Granville, Pintz , Zannier and others (see e.g. [10, 11]). Recently, Hajdu and Tengely [46] extended the research from $\ell = 2$ to the general case, and proved that for any $\ell$ there is a 'best' arithmetic progression, containing the most $\ell$-th powers asymptotically. Our new results from [43] are twofold: on the one hand, we show that these 'best' arithmetic progressions contain only 'slightly more' $\ell$-th powers than the 'average', and on the other hand we give a sharp upper bound for the number of powers (with not fixed exponents) among the first $N$ terms of arithmetic progressions. In our proofs we combine various tools, including a classical result of Wigert [70] concerning the number of divisors of positive integers.

The structure of the dissertation is the following. In the second section we give our results concerning factorials. Then in the third section we provide our theorems related to polynomial values of products of consecutive terms of arithmetic progressions. In the fourth section our results concerning the number of powers in arithmetic progressions are given. Then, after listing the references, we provide Hungarian and English summary, followed by the lists of our publications and conference talks.

# 2 Asymptotic density properties of the sequence (n!)

## 2.1 Introduction

To formalize the questions and results we study in this section consider the following generalized problem, which contains all the related problems studied earlier in the literature.

**Problem 2.1.** Let $p_1, \dots, p_t$ be distinct primes, $m_1, \dots, m_t$ be integers greater than 1, and $r_1, \dots, r_t$ be given integers. Set

$$A = A(p_1, \dots, p_t; m_1, \dots, m_t; r_1, \dots, r_t) =$$

$$\{n : \nu_{p_i}(n!) \equiv r_i \pmod{m_i} \ (i = 1, \dots, t)\},$$

where for $q$ being a prime, $\nu_q(k)$ stands for the exponent of $q$ in the prime factorization of the positive integer $k$. Is it true that for any choice of the parameters $p_i, m_i, r_i$ $(i = 1, \dots, t)$ the set $A$ is non-empty (or even infinite)? Does the set $A$ have a density? Is $A$ relatively dense? (That is, is there an absolute constant $c$ such that the differences of the consecutive elements of $A$ are bounded by $c$?)

A large part of the above problem is already solved. In the special case where $p_1, \dots, p_t$ are the first $t$ primes, $m_1 = \cdots = m_t = 2$ and $r_1 = \cdots = r_t = 0$ (i.e. the original question of Erdős and Graham [25]), Problem 2.1 was answered to the affirmative by Berend [4]. More precisely, Berend showed that in this case $A$ is infinite, and further, it is relatively dense. Berend also provided the same result in the case where $m_1 = \cdots = m_t$ is arbitrary (still only for $p_1, \dots, p_t$ being the first $t$ primes and $r_1 = \cdots = r_t = 0$). Later on, Chen and Zhu [18] formulated Problem 2.1 in the case $m_i = 2$ with $r_i \in \{0, 1\}$ $(i = 1, \dots, t)$. (In fact they asked only about the relative density of $A$, but not about

its density.) Among other results they proved that (in their settings) either $A$ is empty, or it is infinite and even relatively dense. Sander [59] proved that if $t = 1$, then $A$ is infinite, further, it has a density of $1/2$. He also proved that for $t = 2$, $A$ is always infinite. Chen [17] could solve the problem of Chen and Zhu [18] completely, i.e. he proved the relative density of $A$ in Problem 2.1 with $m_i = 2$ and arbitrary $r_i \in \{0, 1\}$ $(i = 1, \dots, t)$. Problem 2.1 with arbitrary moduli $m_i$ $(i = 1, \dots, t)$ was first considered by Luca and Stănică [53] (see their Conjecture 2; in fact they did not ask about relative density). They could prove that under the assumption $p_i \nmid m_i$ $(i = 1, \dots, t)$ the density of $A$ exists and it is equal to $1/m_1 \cdots m_t$. Finally, Berend and Kolesnik [5] proved that Conjecture 2 in [52] is true. That is, they showed that the density of $A$ in Problem 2.1 always exists, and equals to $1/m_1 \cdots m_t$.

We shall also be interested in the behavior of the 'remaining' part of $n!$ (obtained after removing some primes).

**Problem 2.2.** Let $m > 1$ be a positive integer and take an integer $r$ coprime to $m$. Set $M = \prod_{p|m} p$ and for any positive integer $k$, write $k^{(M)}$ for the $M$-free part of $k$, that is, set $k^{(M)} := k/\prod_{p|M} p^{\nu_p(k)}$. Put

$$B = B(m; r) = \{n : (n!)^{(M)} \equiv r \pmod{m}\}.$$

Is it true that for any choice of the integers $m, r$, the set $B$ is non-empty (or even infinite)? Does the set $B$ have a density? Is $B$ relatively dense?

Later we shall see that this problem (in combination with Problem 2.1) has an application concerning the set of those integers $n$ for which $n!$ is representable as a sum of three squares. At this point we mention a related, classical question: what is the distribution of the numbers $1!, 2!, \dots, (p-1)!$ modulo $p$, where $p$ is a prime? (see F11 in [34]) Answering a question of Erdős, Rokowska and Schinzel [57] showed that if $2!, \dots, (p-1)!$ are all distinct modulo $p$, then the missing residue

is $-((p-1)/2)!$, and $p \equiv 5 \pmod 8$ must hold; however, there is no such $p$ with $5 < p \leq 1000$. According to a standard conjecture (see F11 of [34] again), approximately $p/e$ modulo $p$ residue classes are not represented by $n!$, as $p$ tends to infinity. This problem has a huge literature; see e.g. the papers [32, 47] and the references given there. For other related questions, we also refer to the paper [52].

As a combination of Problems 2.1 and 2.2, we also consider

**Problem 2.3.** Let $p_1, \dots, p_t$ be distinct primes, and let $m_1, \dots, m_t$ and $r_1, \dots, r_t$ be integers with $m_i \geq 2$ $(i = 1, \dots, t)$. Further, let $m > 1$ be a positive integer and $r$ be an integer with $\gcd(m, r) = 1$. Put $M = \prod_{p|m} p$. Let

$$C = C(p_1, \dots, p_t; m_1, \dots, m_t; r_1, \dots, r_t; m; r) =$$

$$\{n : \nu_{p_i}(n!) \equiv r_i \pmod{m_i} \, (i = 1, \dots, t) \text{ and } (n!)^{(M)} \equiv r \pmod m\}.$$

Is it true that for any choice of the parameters $p_1, \dots, p_t$, $m_1, \dots, m_t$, $r_1, \dots, r_t$, $m$, $r$ the set $C$ is non-empty (or even infinite)? Does the set $C$ have a density? Is $C$ relatively dense?

We think that in case of all the problems, the answers to all the questions are affirmative. In particular, we conjecture that for any choices of the parameters, both sets $B$ and $C$ have densities, and these are given by $1/\varphi(m)$ and $1/m_1 \cdots m_t \varphi(m)$, respectively. (As we mentioned before, the fact that the density of $A$ exists and equals $1/m_1 \cdots m_t$, was proved by Berend and Kolesnik [5].) This would mean that the properties required in the definitions of the sets $A$ and $B$, are independent. Later on, we shall give some support for this conjecture.

In case of Problems 2.2 and 2.3 we know about only very restricted results, which are related to the problem of representing factorials as sums of squares. This we shall explain a little later.In the case of three squares we arrive at Problem 2.3 with $t = 1$, $p_1 = 2$, $r_1 = 0$, $m = 8$ and

$r = 7$. In this particular case, Deshouillers and Luca [21] proved that the set of values of $n$ satisfying (1) has a density of $7/8$. They also provided an asymptotic formula, saying that the number of $n$ with (1) up to $N$ is $(7/8)N + \mathcal{O}(N^{2/3})$. We mention that for the set of positive integers themselves, the corresponding qualitative results are long known. As one can easily check (and it must also be long known, though unfortunately we could not find any related reference), the density of the set of positive integers with even exponents of 2 in their prime factorization, is $2/3$. Further, the density of the set of positive integers having odd part congruent to 7 modulo 8, is clearly $1/4$. So it is not surprising that the set of positive integers *not* representable as the sum of three squares, is $1/6$. The latter statement was proved by Landau [49]. (See Wagstaff [68] for a similar result concerning the Schnirelmann density of the same set.)

The above problems are strongly related to the question of representing factorials as sum of squares. It is long known that the equation

$$n! = x^2$$

has no solutions in non-negative integers $n, x$ for $n > 1$. This fact follows e.g. from Bertrand's postulate. Thus the question naturally arises: is it still possible to find infinitely many values of $n!$, such that the exponents of all $p \in P$ is even, where $P$ is a given finite set of primes? In the case where $P$ consists of the first $t$ primes for some $t$, this question was posed by Erdős and Graham [25]. The question, and its various extensions have attracted a lot of attention. As it was noted by Erdős and Obláth [26], the equation

$$n! = x^2 + y^2$$

has no solutions in non-negative integers $n, x, y$ for $n > 6$. (We have $6! = 720 = 12^2 + 24^2$.) This follows from the fact that for $n \geq 7$ there exists a prime $p$ of the form $4k + 3$ between $n/2$ and $n$ (see [15] and [22]). On the other hand, by a classical result of Lagrange we know

that the Diophantine equation

$$n! = x^2 + y^2 + z^2 + w^2$$

is solvable for every $n$, in non-negative integers $x, y, z, w$. Consider now the remaining case

$$n! = x^2 + y^2 + z^2 \tag{1}$$

in non-negative integers $n, x, y, z$. By a classical result of Gauss, it is known that an integer is *not* representable as the sum of three squares if and only if it is of the form $2^{2a}(8b + 7)$ with non-negative integers $a, b$. This provides a direct link to Problem 2.3.

As we saw, the question of representing $n!$ as the sum of at most two squares is treated by the knowledge concerning primes in the block of the first $n$ positive integers. The much more general problem of describing the size of the largest prime factor in a block of consecutive integers has been investigated by many authors. For related results, we refer to the excellent, recent survey paper of Shorey and Tijdeman [64], and the references therein.

In this section we prove several new results concerning the most general question formulated, namely Problem 2.3. We take up the problem where the moduli are any powers of some prime $p$. We prove that for all choices of the other parameters, the conjecture formulated above is true; that is, in all cases $C$ has the suspected density. We also prove that $C$ is relatively dense, with an explicit bound for the gaps of the consecutive elements of $C$. In the particular case $p = 2$ and $m_1 = 2$, $m = 8$ we also prove that asymptotically, the error term is $\mathcal{O}(N^{1/2} \log^2 N)$ up to $N$, thus improving the result of Deshouillers and Luca [21]. In this special case we give a sharp upper bound for the gaps between the consecutive elements of $C$, as well.

## 2.2 New results

To formulate our general results, we need some notation. Let $p$ be a prime and $a, b$ be positive integers. Throughout this section we shall always assume that $p, a, b$ are fixed. Put

$$I_1 = \{0, 1, \dots, p^a - 1\},\ I_2 = \{i\ :\ 1 \le i \le p^b,\ p \nmid i\} \quad \text{and} \quad I = I_1 \times I_2.$$

Observe that $|I| = (p-1)p^{a+b-1}$. To simplify the reference to these sets, by writing $\alpha \in I_1$ and $\beta \in I_2$ for arbitrary integers $\alpha$ and $\beta$ with $p \nmid \beta$, we shall always mean the elements $\alpha' \in I_1$ and $\beta' \in I_2$, for which $\alpha' \equiv \alpha \pmod{p^a}$ and $\beta' \equiv \beta \pmod{p^b}$, respectively.

For $(\alpha, \beta) \in I$ put

$$H^{(\alpha,\beta)} = \{n : \nu_p(n!) \equiv \alpha \pmod{p^a},\ (n!)^{(p)} \equiv \beta \pmod{p^b}\}.$$

Finally, we shall use the conventions

$$\nu_p(0) = 0 \quad \text{and} \quad 0^{(p)} = 1.$$

Our first two theorems solve the question of density and relative density in Problem 2.3 for $t = 1$ with $m_1 = p^a$ and $m = p^b$, where $a, b$ are arbitrary positive integers, $p_1 = p$ is an arbitrary prime, and $r_1$ and $r$ are arbitrary integers. The following statement shows that the pairs $(\nu_p(n!) \pmod{p^a}, (n!)^{(p)} \pmod{p^b})$ are uniformly distributed among the possible pairs.

**Theorem 2.1** (Á. Papp, L. Hajdu [40])**.** *For all $(\alpha, \beta) \in I$, the set $H^{(\alpha,\beta)}$ has a density of $1/(p-1)p^{a+b-1}$.*

In case of relative density, we can even give an explicit upper bound for the differences between the consecutive terms of $H^{(\alpha,\beta)}$.

**Theorem 2.2** (Á. Papp, L. Hajdu [40]). *For all $(\alpha, \beta) \in I$, the set $H^{(\alpha,\beta)}$ is relatively dense. Further, if we write $H_1^{(\alpha,\beta)} < H_2^{(\alpha,\beta)} < H_3^{(\alpha,\beta)} < \dots$ for the elements of $H^{(\alpha,\beta)}$, then we have*

$$H_{i+1}^{(\alpha,\beta)} - H_i^{(\alpha,\beta)} \leq 2p^{\max(a,b)+b(p-1)p^{a+2b-2}-b+1}$$

*for all $i \geq 1$.*

As a simple consequence of the above theorems we obtain the following.

**Corollary 2.1.** *For any $\alpha \in I_1$ and $\beta \in I_2$, the sets*

$$\{n : \nu_p(n!) \equiv \alpha \pmod{p^a}\} \quad \text{and} \quad \{n : (n!)^{(p)} \equiv \beta \pmod{p^b}\}$$

*are relatively dense and are of densities $1/p^a$ and $1/(p-1)p^{b-1}$, respectively.*

Note that the fact that the density of the first set is $1/p^\alpha$, follows from the earlier mentioned results of Berend and Kolesnik [5].

Now we give alike, but more precise statements in the special case of $p = 2$, $a = 1$ and $b = 3$. This case is of particular interest, since it describes the density of the set of values of $n$ for which $n!$ is expressible as a sum of three squares.

The next theorem improves and extends an earlier mentioned result of Deshouillers and Luca [21]. In that paper only the case $(\alpha, \beta) = (0, 7)$ has been investigated (though it seems to be clear that the methods applied in [21] are capable to handle all the other choices of $(\alpha, \beta)$). Our result significantly improves the error term $\mathcal{O}(x^{2/3})$ in [21], too.

**Theorem 2.3** (Á. Papp, L. Hajdu [40]). *Let $p = 2$, $a = 1$, $b = 3$ and $(\alpha, \beta) \in I$. Then for all $x > 0$ we have*

$$|H^{(\alpha,\beta)} \cap [0, x)| = (1/8)x + \mathcal{O}(x^{1/2} \log^2 x).$$

Our final theorem gives the precise value for the maximal gap length in the sets $H^{(\alpha,\beta)}$ in this special case.

**Theorem 2.4** (Á. Papp, L. Hajdu [40]). *Let $p = 2$, $a = 1$, $b = 3$ and $(\alpha, \beta) \in I$. Then the set $H^{(\alpha,\beta)}$ is relatively dense, and if we write $H_1^{(\alpha,\beta)} < H_2^{(\alpha,\beta)} < H_3^{(\alpha,\beta)} < \dots$ for the elements of $H^{(\alpha,\beta)}$, then we have $H_{i+1}^{(\alpha,\beta)} - H_i^{(\alpha,\beta)} \leq 42$ for all $i \geq 1$. Further, the upper bound 42 is sharp for all $(\alpha, \beta) \in I$.*

**Remark 2.1.** Clearly, for the special choices $p = 2$, $a = 1$ and $b \leq 3$ Corollary 2.1 also follows from Theorems 2.3 and 2.4.

## 2.3 Proofs

To prove our theorems, we need several lemmas. The first lemma reveals a pattern in the behavior of $\nu_p(n!)$. For similar statements and assertions see [5, 18, 52, 59].

**Lemma 2.1** (Á. Papp, L. Hajdu [40]). *Let $a$ be a positive integer. Then for any positive integers $t$ and $k$ with $0 \leq t < p$ and $k \geq a$ we have*

$$\nu_p((tp^k + i)!) \equiv \nu_p(i!) + \frac{t}{1-p} \pmod{p^a} \quad (0 \leq i < p^k)$$

*with the usual convention $0! = 1$.*

*Proof.* We clearly have

$$\nu_p((tp^k+i)!) = \nu_p\left((tp^k)! \prod_{j=1}^{i} (tp^k + j)\right) \equiv \nu_p((tp^k)!)+\nu_p(i!) \pmod{p^a}.$$

Using the Legendre formula this gives

$$\nu_p((tp^k+i)!) \equiv \nu_p(i!)+t(1+p+\dots+p^{k-1}) \equiv \nu_p(i!)+t\frac{p^k-1}{p-1} \pmod{p^a},$$

and the lemma follows. □

Our next two lemmas provide similar information about the behavior of $(n!)^{(p)} \pmod{p^b}$.

**Lemma 2.2** (Á. Papp, L. Hajdu [40]). *Let $b \geq 1$ and further assume that $b \geq 3$ if $p = 2$. Then for any $k \geq b-1$ and for any positive integer $t$ we have*

$$((tp^k)!)^{(p)} \equiv \begin{cases} ((t2^{b-1})!)^{(2)} \pmod{2^b}, & \text{if } p = 2, \\ (-1)^{t(k-b+1)}((tp^{b-1})!)^{(p)} \pmod{p^b}, & \text{otherwise,} \end{cases}$$

*or, shortly*

$$((tp^k)!)^{(p)} \equiv (-1)^{t(k-b+1)p}((tp^{b-1})!)^{(p)} \pmod{p^b}$$

*for all $p \geq 2$.*

*Proof.* We proceed by induction on $k$. For $k = b-1$ the statement is an identity. Suppose that the assertion is valid for some $k \geq b-1$. We can write

$$((tp^{k+1})!)^{(p)} = ((tp^k)!)^{(p)} \prod_{\substack{i=1 \\ p \nmid i}}^{tp^{k+1}} i.$$

Since by the induction hypothesis we have

$$((tp^k)!)^{(p)} \equiv \begin{cases} ((t2^{b-1})!)^{(2)} \pmod{2^b}, & \text{if } p = 2, \\ (-1)^{t(k-b+1)}((tp^{b-1})!)^{(p)} \pmod{p^b}, & \text{otherwise,} \end{cases}$$

we only need to show that

$$\prod_{\substack{i=1 \\ p \nmid i}}^{tp^{k+1}} i \equiv \begin{cases} 1 \pmod{2^b}, & \text{if } p = 2, \\ (-1)^t \pmod{p^b}, & \text{otherwise.} \end{cases} \tag{2}$$

To prove this, observe that

$$\prod_{\substack{i=1 \\ p \nmid i}}^{tp^{k+1}} i \equiv \left( \prod_{\substack{i=1 \\ p \nmid i}}^{p^{k+1}} i \right)^t \pmod{p^{k+1}}.$$

Now in view of $k+1 \geq b \geq 3$ if $p = 2$ and $k+1 \geq b \geq 1$ otherwise, we have that $u^2 \equiv 1 \pmod{p^b}$ if and only if $u \equiv \pm 1, 2^k \pm 1 \pmod{2^{k+1}}$ for $p = 2$, and $u \equiv \pm 1 \pmod{p^{k+1}}$ for $p \geq 3$. Hence we get

$$\prod_{\substack{i=1 \\ p \nmid i}}^{p^{k+1}} i \equiv \begin{cases} 1 \pmod{2^{k+1}}, & \text{if } p = 2, \\ -1 \pmod{p^{k+1}}, & \text{otherwise.} \end{cases}$$

This in view of $k+1 \geq b$ gives (2), which proves the lemma. □

The following lemma plays a key role in our arguments.

**Lemma 2.3** (Á. Papp, L. Hajdu [40]). *Let $b \geq 1$ and $k \geq b-1$. If $p = 2$ and $b = 2$ then assume that $k \geq b$. Then for every $t \geq 1$ with $p \nmid t$ there exist uniquely determined numbers $\gamma_t(j) \in I_2$ $(1 \leq j \leq p^{b-1})$, such that for all $i, j$ with $(j-1)p^{k-b+1} \leq i < jp^{k-b+1}$, $1 \leq j \leq p^{b-1}$ we have*

$$((tp^k + i)!)^{(p)} \equiv \gamma_t(j)(i!)^{(p)} \pmod{p^b}.$$

*Further,*

*i) if $p = 2$, then the numbers $\gamma_t(j)$ $(j = 1, \ldots, 2^{b-1})$ form a permutation of $I_2$ for any odd $t$,*

*ii) if $p \geq 3$ then $\gamma_{p-1}(p^{b-1}) = -1$, and the numbers $\gamma_t(j)$ $(1 \leq t \leq p-1,\ 1 \leq j \leq p^{b-1})$ generate the multiplicative group $\mathbb{Z}_{p^b}^*$ of invertible elements of $\mathbb{Z}_{p^b}$.*

*Proof.* For $p = 2$ and $b = 1$ the lemma is trivial. Further, if $p = 2$ and $b = 2$, then the statement can be easily proved by induction; we get $\gamma_1(1) = 3$ and $\gamma_1(2) = 1$ in this case. So from this point on we shall always assume that if $p = 2$ then $b \geq 3$.

First observe that for any $s$ with $tp^k \leq s < (t+1)p^k$ such that $\nu_p(s) \leq k - b$ we have

$$s^{(p)} \equiv (s - tp^k)^{(p)} \pmod{p^b}.$$

This immediately gives that writing

$$\delta_t(j) = \frac{(tp^k + (j-1)p^{k-b+1})^{(p)}}{((j-1)p^{k-b+1})^{(p)}} \quad (1 \le j \le p^{b-1})$$

with the convention $0^{(p)} = 1$, we have

$$((tp^k + i)!)^{(p)} \equiv (i!)^{(p)} \delta_t(0) \prod_{\ell=1}^{j} \delta_t(\ell) \pmod{p^b} \tag{3}$$

for all $i, j$ with $(j-1)p^{k-b+1} \le i < jp^{k-b+1}$, $1 \le j \le p^{b-1}$, where $\delta_t(0) = (-1)^{t(k-b+1)p}((tp^{b-1})!)^{(p)} t^{-1}$. Here we used that by Lemma 2.2

$$((tp^k-1)!)^{(p)} \equiv ((tp^k)!)^{(p)} t^{-1} \equiv (-1)^{t(k-b+1)p}((tp^{b-1})!)^{(p)} t^{-1} \pmod{p^b}.$$

So by choosing

$$\gamma_t(j) \equiv \delta_t(0) \prod_{\ell=1}^{j} \delta_t(\ell) \pmod{p^b} \quad (1 \le j \le p^{b-1}) \tag{4}$$

and noting that these numbers are clearly uniquely determined, the first part of the statement follows.

To prove the second part of the lemma, we have to distinguish the cases $p = 2$ and $p \ge 3$.

*The case* $p = 2$. To prove the second statement in this case, we show that the $\gamma_t(j)$ are all distinct. This clearly implies $i$). For this, we need to show that the products $\prod_{\ell=\ell_1}^{\ell_2} \delta_t(\ell)$ are all distinct from 1 modulo $2^b$, for $1 < \ell_1 \le \ell_2 \le 2^{b-1}$. (Recall that here we may assume that $b \ge 3$.) Putting $u_{j-1} \equiv ((j-1)^{(2)})^{-1} \pmod{2^b}$ for $j > 1$, we have

$$\delta_t(j) \equiv t2^{b-1-\nu_2(j-1)} u_{j-1} + 1 \pmod{2^b}.$$

Since for any $j > 1$ obviously $\delta_t(j) \not\equiv 1 \pmod{2^b}$, we may assume that $\ell_1 < \ell_2$. Thus we need to show that

$$\prod_{\ell=\ell_1}^{\ell_2} (t2^{b-1-\nu_2(\ell-1)} u_{\ell-1} + 1) \not\equiv 1 \pmod{2^b} \quad (1 < \ell_1 < \ell_2 \le 2^{b-1}).$$

Suppose to the contrary that the above congruence holds for some $\ell_1, \ell_2$ as above. Then for these values of $\ell_1, \ell_2$ we have

$$\prod_{\substack{\ell=\ell_1 \\ 2|\ell-1}}^{\ell_2} (t2^{b-1-\nu_2(\ell-1)}u_{\ell-1}+1) \prod_{\substack{\ell=\ell_1 \\ 2\nmid\ell-1}}^{\ell_2} (t2^{b-1-\nu_2(\ell-1)}u_{\ell-1}+1) \equiv 1 \pmod{2^b}.$$

However, then the same congruence certainly holds also modulo $2^{b-1}$. Thus, observing that the second product is clearly 1 modulo $2^{b-1}$, in case of $2 \mid \ell-1$ letting $\ell'-1=(\ell-1)/2$ we get

$$\prod_{\ell'=\lceil(\ell_1+1)/2\rceil}^{\lfloor(\ell_2+1)/2\rfloor} (t2^{b-2-\nu_2(\ell'-1)}u_{\ell'-1}+1) \equiv 1 \pmod{2^{b-1}}.$$

Here we used that $\nu_2(\ell-1)=\nu_2((\ell-1)/2)+1$ and $((\ell-1)/2)^{(2)} = (\ell-1)^{(2)}$ for $\ell-1$ even, whence $u_{\ell'-1} \equiv u_{\ell-1} \pmod{2^b}$. Since $u_{\ell'-1}\cdot(\ell'-1)^{(2)} \equiv 1 \pmod{2^b}$, we certainly have $u_{\ell'-1}\cdot(\ell'-1)^{(2)} \equiv 1 \pmod{2^{b-1}}$, as well. Finally, observe that $1 < \lceil(\ell_1+1)/2\rceil \le \lfloor(\ell_2+1)/2\rfloor \le 2^{b-2}$. Hence our claim that the products $\prod_{\ell=\ell_1}^{\ell_2}\delta_t(\ell)$ $(1<\ell_1\le\ell_2\le 2^{b-1})$ are all different from 1 modulo $2^b$ follows by induction on $b$. Thus the products $\prod_{\ell=1}^{j}\delta_t(\ell)$ $(j=1,\dots,2^{b-1})$ are pairwise distinct modulo $2^b$. Hence the lemma follows in this case.

*The case $p \ge 3$.* To prove the second statement in this case, first observe that by Lemma 2.2, (3) and (4) we have

$$(-1)^{k-b+1}((p^{b-1})!)^{(p)}\gamma_{p-1}(p^{b-1}) \equiv ((p^k)!)^{(p)}\gamma_{p-1}(p^{b-1}) \equiv$$
$$\equiv ((p^{k+1})!)^{(p)} \equiv (-1)^{k-b+2}((p^{b-1})!)^{(p)} \pmod{p^b}.$$

This gives

$$\gamma_{p-1}(p^{b-1}) \equiv -1 \pmod{p^b}.$$

Now we construct an element of the subgroup generated by the elements $\gamma_t(j)$ $(1\le t\le p-1,\ 1\le j\le p^{b-1})$ which is a generator of $\mathbb{Z}_{p^b}^*$. In fact we shall prove that already the subgroup $G$ of $\mathbb{Z}_{p^b}^*$ generated by

$$\pm\delta_t(0)\delta_t(1) \quad (1\le t\le p-1)$$

contains such an element. It will be sufficient, since

$$\gamma_{p-1}(p^{b-1}) = -1 \quad \text{and} \quad \gamma_t(1) = \delta_t(0)\delta_t(1) \quad (1 \le t \le p-1).$$

As $\delta_t(1) \equiv t \pmod{p^b}$, we have

$$\pm\delta_t(0)\delta_t(1) \equiv \pm((tp^{b-1})!)^{(p)} \pmod{p^b} \quad (1 \le t \le p-1).$$

Observe that for $b = 1$ we have $t! \in G$ for all $t = 1, \dots, p-1$ implying $G = \mathbb{Z}_p^*$. Let now $b \ge 2$. We show that $G$ contains an element which is generator modulo $p^2$. As it is well-known, this element will be a generator also modulo $p^b$. Since by Lemma 2.2 we have

$$((tp^{b-1})!)^{(p)} \equiv \pm((tp)!)^{(p)} \pmod{p^2}$$

for any $b \ge 2$, we need to check the statement only for $b = 2$. That is, it is sufficient to show that for $b = 2$, the group $G$ contains a generator modulo $p^2$. For this, first observe that for any $t$ we have

$$\frac{((tp)!)^{(p)}}{(((t-1)p)!)^{(p)}} = t\prod_{i=1}^{p-1}((t-1)p+i) \in G \quad (t = 1, \dots, p-1).$$

Recall that if $g_1, g_2$ are generators modulo $p$ such that

$$g_1 \equiv g_2 \pmod{p} \quad \text{but} \quad g_1 \not\equiv g_2 \pmod{p^2},$$

then one of $g_1, g_2$ is also a generator modulo $p^2$. Let $g$ be any generator element modulo $p$ with $1 < g < p$. Then by Wilson's theorem we have

$$-g\prod_{i=1}^{p-1}((g-1)p+i) \equiv (p-g)\prod_{i=1}^{p-1}((p-g-1)p+i) \equiv g \pmod{p}.$$

If we would also have

$$-g\prod_{i=1}^{p-1}((g-1)p+i) \equiv (p-g)\prod_{i=1}^{p-1}((p-g-1)p+i) \pmod{p^2},$$

then

$$-g(g-1)\sum_{j=1}^{p-1}\frac{(p-1)!}{j}\equiv(p-1)!+g(g+1)\sum_{j=1}^{p-1}\frac{(p-1)!}{j}\pmod p$$

would also hold. However, this by

$$\sum_{j=1}^{p-1}\frac{1}{j}\equiv\sum_{j=1}^{p-1}j\equiv\frac{p(p-1)}{2}\equiv 0\pmod p$$

is impossible. This implies that one of

$$-g\prod_{i=1}^{p-1}((g-1)p+i),\ (p-g)\prod_{i=1}^{p-1}((p-g-1)p+i)\in G$$

is a generator modulo $p^2$. Hence the lemma follows. □

**Remark 2.2.** The combination of Lemmas 2.1 and 2.3 allows us to follow how the classes corresponding to $(\alpha,\beta)\in I$ 'switch'. To see this, just observe that combining these lemmas, for any $k\geq\max(a,b)$ we have that

$$(\nu_p(i!),(i!)^{(p)})=(\alpha,\beta)$$

if and only if

$$(\nu_p((tp^k+i)!),((tp^k+i)!)^{(p)})=\left(\alpha+\frac{t}{1-p},\gamma_t(j)\beta\right),$$

for any $t,j,i$ with $0\leq t\leq p-1$, $1\leq j\leq p^{b-1}$ and $(j-1)p^{k-b+1}\leq i<jp^{k-b+1}$.

Now we can give the proofs of Theorems 2.3 and 2.4.

*Proof of Theorem 2.3.* First we prove the statement for numbers of the form $x=(2^k)!$. For this, we use linear recurrence sequences of vectors. We need to introduce some notation. For $(\alpha,\beta)\in I$ and $x>y\geq 0$ let

$$H^{(\alpha,\beta)}(y,x):=H^{(\alpha,\beta)}\cap[y,x)\quad\text{ and }\quad h^{(\alpha,\beta)}(y,x):=\frac{|H^{(\alpha,\beta)}(y,x)|}{x-y}.$$

If $y = 0$, then we shall simply write $H^{(\alpha,\beta)}(x)$ and $h^{(\alpha,\beta)}(x)$, respectively. Define the vectors $\vec{v}_k$ $(k \geq 0)$ by

$$\vec{v}_k := \begin{pmatrix} h^{(0,1)}(2^{k+1}) \\ h^{(0,3)}(2^{k+1}) \\ h^{(0,5)}(2^{k+1}) \\ h^{(0,7)}(2^{k+1}) \\ h^{(1,1)}(2^{k+1}) \\ h^{(1,3)}(2^{k+1}) \\ h^{(1,5)}(2^{k+1}) \\ h^{(1,7)}(2^{k+1}) \end{pmatrix}.$$

Any term of the sequence $(\vec{v}_k)$ for $k \geq 4$ can be expressed by the help of the previous four terms. The initial vectors $\vec{v}_0, \vec{v}_1, \vec{v}_2, \vec{v}_3$ can be easily obtained by the help of Table 1. These are the following:

$$\vec{v}_0 = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{v}_1 = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \\ 1 \\ 2 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} 2 \\ 2 \\ 2 \\ 2 \\ 1 \\ 5 \\ 1 \\ 1 \end{pmatrix}.$$

As we already mentioned, for every $k \geq 4$ the coordinates of $\vec{v}_k$ can be given by the help of the previous four vectors. The first entry of $\vec{v}_k$, namely $h^{(0,1)}(2^{k+1})$, can be obtained in the following way. We start with

$$h^{(0,1)}(2^{k+1}) = h^{(0,1)}(2^k) + h^{(0,1)}(2^k, 2^{k+1}).$$

Cutting the interval $(2^k, 2^{k+1})$ into four parts, we get

$$h^{(0,1)}(2^{k+1}) = h^{(0,1)}(2^k) + h^{(0,1)}(2^k, 2^k + 2^{k-2}) + h^{(0,1)}(2^k + 2^{k-2}, 2^k + 2^{k-1})$$
$$+ h^{(0,1)}(2^k + 2^{k-1}, 2^k + 2^{k-1} + 2^{k-2}) + h^{(0,1)}(2^k + 2^{k-1} + 2^{k-2}, 2^{k+1}).$$

| $n$ | $2^{\nu_2(n)} \cdot n^{(2)}$ | $2^{\nu_2(n!)} \cdot n!^{(2)}$ | $(\alpha, \beta)$ | $n$ | $2^{\nu_2(n)} \cdot n^{(2)}$ | $2^{\nu_2(n!)} \cdot n!^{(2)}$ | $(\alpha, \beta)$ |
|---|---|---|---|---|---|---|---|
| 0 | 1 | 1 | (0,1) | 16 | $2^4$ | $2^{15} \cdot 3$ | (1,3) |
| 1 | 1 | 1 | (0,1) | 17 | 1 | $2^{15} \cdot 3$ | (1,3) |
| 2 | 2 | $2 \cdot 1$ | (1,1) | 18 | $2 \cdot 1$ | $2^{16} \cdot 3$ | (0,3) |
| 3 | 3 | $2 \cdot 3$ | (1,3) | 19 | 3 | $2^{16} \cdot 1$ | (0,1) |
| 4 | $2^2$ | $2^3 \cdot 3$ | (1,3) | 20 | $2^2 \cdot 5$ | $2^{18} \cdot 5$ | (0,5) |
| 5 | 5 | $2^3 \cdot 7$ | (1,7) | 21 | 5 | $2^{18} \cdot 1$ | (0,1) |
| 6 | $2 \cdot 3$ | $2^4 \cdot 5$ | (0,5) | 22 | $2 \cdot 3$ | $2^{19} \cdot 3$ | (1,3) |
| 7 | 7 | $2^4 \cdot 3$ | (0,3) | 23 | 7 | $2^{19} \cdot 5$ | (1,5) |
| 8 | $2^3$ | $2^7 \cdot 3$ | (1,3) | 24 | $2^3 \cdot 3$ | $2^{22} \cdot 7$ | (0,7) |
| 9 | 1 | $2^7 \cdot 3$ | (1,3) | 25 | 1 | $2^{22} \cdot 7$ | (0,7) |
| 10 | $2 \cdot 5$ | $2^8 \cdot 7$ | (0,7) | 26 | $2 \cdot 5$ | $2^{23} \cdot 3$ | (1,3) |
| 11 | 3 | $2^8 \cdot 5$ | (0,5) | 27 | 3 | $2^{23} \cdot 1$ | (1,1) |
| 12 | $2^2 \cdot 3$ | $2^{10} \cdot 7$ | (0,7) | 28 | $2^2 \cdot 7$ | $2^{25} \cdot 7$ | (1,7) |
| 13 | 5 | $2^{10} \cdot 3$ | (0,3) | 29 | 5 | $2^{25} \cdot 3$ | (1,3) |
| 14 | $2 \cdot 7$ | $2^{11} \cdot 5$ | (1,5) | 30 | $2 \cdot 7$ | $2^{26} \cdot 5$ | (0,5) |
| 15 | 7 | $2^{11} \cdot 3$ | (1,3) | 31 | 7 | $2^{26} \cdot 3$ | (0,3) |

Table 1: Exponents of 2 and odd parts of $n!$ for $0 \leq n \leq 31$.

In what follows, we shall apply Lemmas 2.1 and 2.3 repeatedly, in the way explained in Remark 2.2. In the latter statement, as one can easily check, now we have

$$(\gamma_1(1), \gamma_1(2), \gamma_1(3), \gamma_1(4)) = (3, 7, 5, 1).$$

We get

$$h^{(0,1)}(2^{k+1}) = h^{(0,1)}(2^k) + h^{(1,3)}(2^{k-2}) + h^{(1,7)}(2^{k-2}, 2^{k-1})$$

$$+h^{(1,5)}(2^{k-1}, 2^{k-1} + 2^{k-2}) + h^{(1,1)}(2^{k-1} + 2^{k-2}, 2^k).$$

From this we obtain

$$h^{(0,1)}(2^{k+1}) = h^{(0,1)}(2^k) + h^{(1,3)}(2^{k-2}) + \left(h^{(1,7)}(2^{k-1}) - h^{(1,7)}(2^{k-2})\right)$$

$$+ \left(h^{(0,7)}(2^{k-3}) + h^{(0,3)}(2^{k-3}, 2^{k-2})\right) +$$

$$+ \left(h^{(1,1)}(2^k) - h^{(1,1)}(2^{k-1}) - h^{(1,1)}(2^{k-1}, 2^{k-1} + 2^{k-2})\right) =$$

$$= h^{(0,1)}(2^k) + h^{(1,3)}(2^{k-2}) + \left(h^{(1,7)}(2^{k-1}) - h^{(1,7)}(2^{k-2})\right) +$$

$$+ \left(h^{(0,7)}(2^{k-3}) + h^{(0,3)}(2^{k-2}) - h^{(0,3)}(2^{k-3})\right) +$$

$$\left(h^{(1,1)}(2^k) - h^{(1,1)}(2^{k-1}) - h^{(0,3)}(2^{k-3}) - h^{(0,7)}(2^{k-2}) + h^{(0,7)}(2^{k-3})\right).$$

After rearrangement, this yields

$$h^{(0,1)}(2^{k+1}) = h^{(0,1)}(2^k) + h^{(1,1)}(2^k) - h^{(1,1)}(2^{k-1}) +$$

$$+h^{(1,7)}(2^{k-1}) + h^{(0,3)}(2^{k-2}) - h^{(0,7)}(2^{k-2})$$

$$+h^{(1,3)}(2^{k-2}) - h^{(1,7)}(2^{k-2}) - 2 \cdot h^{(0,3)}(2^{k-3}) + 2 \cdot h^{(0,7)}(2^{k-3}).$$

Thus the first coordinate of $\vec{v}_k$ is given by

$$\vec{v}_k = (1,0,0,0,1,0,0,0) \cdot \vec{v}_{k-1} + (0,0,0,0,-1,0,0,1) \cdot \vec{v}_{k-2} +$$

$$+(0,1,0,-1,0,1,0-1) \cdot \vec{v}_{k-3} + (0,-2,0,2,0,0,0,0) \cdot \vec{v}_{k-4}.$$

Similar calculations yield the other coordinates of $\vec{v}_k$, as well. Altogether, we get the recurrence relation

$$\vec{v}_k = G_1 \cdot \vec{v}_{k-1} + G_2 \cdot \vec{v}_{k-2} + G_3 \cdot \vec{v}_{k-3} + G_4 \cdot \vec{v}_{k-4} \quad (k \geq 4),$$

where

$$G_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$G_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$G_3 = \begin{pmatrix} 0 & 1 & 0 & -1 & 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 & 1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 1 & 0 & -1 & 0 & 1 \\ -1 & 0 & 1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 & 1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 1 & 0 & -1 & 0 & 1 \\ -1 & 0 & 1 & 0 & -1 & 0 & 1 & 0 \end{pmatrix},$$

$$G_4 = \begin{pmatrix} 0 & -2 & 0 & 2 & 0 & 0 & 0 & 0 \\ -2 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & -2 & 0 & 0 & 0 & 0 \\ 2 & 0 & -2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -2 & 0 & 2 \\ 0 & 0 & 0 & 0 & -2 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & -2 \\ 0 & 0 & 0 & 0 & 2 & 0 & -2 & 0 \end{pmatrix}.$$

By a theorem of Cerrucci and Vaccarino [16] we know that the above relation can be written as a system of coordinatewise recurrence relations, and the common generating polynomial $g(x)$ of these relations is the characteristic polynomial of the matrix

$$G = \begin{pmatrix} O & O & O & G_4 \\ E & O & O & G_3 \\ O & E & O & G_2 \\ O & O & E & G_1 \end{pmatrix}.$$

Here $O$ is the $8 \times 8$ zero matrix, and $E$ is the $8 \times 8$ unit matrix. Hence for $g(x)$ we get

$$g(x) = x^{11}(x-2)(x^2+2)(x^4-2x^2+4)(x^4-2x^3+2x^2-4x+4)(x^2-2x+2)^2(x^2-2)^3.$$

Now using the standard theory of recurrence sequences (see e.g. [28]), we get that every coordinate of $\vec{v}_k$ can be written as

$$c_1 2^k + c_2(i\sqrt{2})^k + c_3(-i\sqrt{2})^k + \sum_{i=0}^{3} c_{4+i}\alpha_i^k + \sum_{i=0}^{3} c_{8+i}\beta_i^k +$$
$$(c_{12}k + c_{13})(1+i)^k + (c_{14}k + c_{15})(1-i)^k +$$
$$(c_{16}k^2 + c_{17}k + c_{18})\sqrt{2}^k + (c_{19}k^2 + c_{20}k + c_{21})(-\sqrt{2})^k$$

with complex numbers $c_1, \ldots, c_{21}$, which can be different for different coordinates. Here $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ and $\beta_1, \beta_2, \beta_3, \beta_4$ are the (distinct) roots of $x^4 - 2x^2 + 4$ and $x^4 - 2x^3 + 2x^2 - 4x + 4$, respectively. It is easy to check that all these roots have absolute value $\sqrt{2}$. Calculating the vectors $\vec{v}_k$ for $k = 0, \ldots, 20$, the constants $c_1, \ldots, c_{21}$ can be

obtained for each coordinate by solving a system of linear equations. Using Magma [14] we get that $c_1 = 1/8$ in all cases. Thus

$$h^{(\alpha,\beta)}(2^k) = \frac{2^k}{8} + \mathcal{O}\left(k^2\sqrt{2}^k\right) \tag{5}$$

for every $(\alpha, \beta) \in I$. As $k = \log_2 2^k$ and $\sqrt{2}^k = (2^k)^{1/2}$, hence the formula in the theorem follows for $x = 2^k$. As we have

$$h^{(\alpha,\beta)}(2^k, 2^{k+1}) = h^{(\alpha,\beta)}(2^{k+1}) - h^{(\alpha,\beta)}(2^k),$$

thus also

$$h^{(\alpha,\beta)}(2^k, 2^{k+1}) = \frac{2^{k+1}}{8} - \frac{2^k}{8} + \mathcal{O}\left((k+1)^2\sqrt{2}^{k+1}\right) = \frac{2^k}{8} + \mathcal{O}\left(k^2\sqrt{2}^k\right). \tag{6}$$

Let now $N$ be an arbitrary positive integer. Then we can write

$$N = \sum_{i=1}^{j} 2^{f_i},$$

with $f_1 > f_2 > \cdots > f_j \geq 0$. Clearly, $h^{(\alpha,\beta)}(N)$ can be written as

$$h^{(\alpha,\beta)}(N) = h^{(\alpha,\beta)}(2^{f_1}) + h^{(\alpha,\beta)}(2^{f_1}, 2^{f_1} + 2^{f_2}) + \cdots +$$

$$+h^{(\alpha,\beta)}(2^{f_1} + \cdots + 2^{f_{j-1}}, 2^{f_1} + \cdots + 2^{f_j}).$$

Any term of the above sum can be expressed as

$$h^{(\alpha,\beta)}(2^{f_1} + \cdots + 2^{f_\ell}, 2^{f_1} + \cdots + 2^{f_\ell} + 2^{f_{\ell+1}}) = h^{(\alpha,\beta)}(t2^{f_\ell}, t2^{f_\ell} + 2^{f_{\ell+1}}),$$

where $t$ is odd. By Lemmas 2.1 and 2.3, using (5) and (6), we easily get

$$h^{(\alpha,\beta)}(t2^{f_\ell}, t2^{f_\ell} + 2^{f_{\ell+1}}) = \frac{2^{f_{\ell+1}}}{8} + \mathcal{O}\left(f_{\ell+1}^2\sqrt{2}^{f_{\ell+1}}\right).$$

Thus

$$h^{(\alpha,\beta)}(N) = \left(\frac{2^{f_1}}{8} + \cdots + \frac{2^{f_j}}{8}\right) + \mathcal{O}\left(f_1^2\sqrt{2}^{f_1}\right) + \cdots + \mathcal{O}\left(f_j^2\sqrt{2}^{f_j}\right),$$

whence

$$h^{(\alpha,\beta)}(N) = \frac{N}{8} + \mathcal{O}\left(f_1^2 \frac{\sqrt{2}^{f_1+1} - 1}{\sqrt{2}-1}\right) = \frac{N}{8} + \mathcal{O}\left(f_1^2 \sqrt{2}^{f_1}\right).$$

Now using $f_1 \le \log_2 N$ we get

$$h^{(\alpha,\beta)}(N) = \frac{N}{8} + \mathcal{O}\left(\log^2 N \cdot N^{1/2}\right),$$

and the theorem follows. □

*Proof of Theorem 2.4.* The assertions can be readily checked. For this purpose, we used simple Magma [14] programs. Checking all the values of $n!$ with $0 \le n < 127$, we find that all the intervals $[0, 32)$, $[32, 64)$, $[64, 96)$, $[96, 127)$ contain at least two elements from all the sequences $H^{(\alpha,\beta)}$ $((\alpha, \beta) \in I)$. Obviously, the consecutive elements of any $H^{(\alpha,\beta)}$ inside any of these intervals, have distance less than 42. Further, as one can easily check, all the intervals $[0, 14]$, $[21, 31]$, $[32, 48]$, $[47, 63]$, $[64, 78]$, $[85, 95]$, $[96, 122]$, $[100, 127]$ contain at least one element from each $H^{(\alpha,\beta)}$. This immediately shows that all the differences of the consecutive elements below $2^7$ inside all the sets $H^{(\alpha,\beta)}$ are bounded by 42. Further, in view of that all the numbers $(127-100)+(14-0)+1$, $(31-21)+(48-32)+1$, $(63-47)+(78-64)+1$, $(95-85)+(122-96)+1$ are at most 42, we also have no problem with merging the four intervals, the first two statements follow by induction using Lemmas 2.1 and 2.3. To see that the bound 42 cannot be improved for any $(\alpha, \beta) \in I$, one can check that

$$(4836, 4878),\ (39652, 39694),\ (2788, 2830),\ (23268, 23310),$$

$$(6884, 6926),\ (740, 782),\ (13028, 13070),\ (19172, 19214)$$

are pairs of consecutive elements of the sets

$$H^{(0,1)},\ H^{(0,3)},\ H^{(0,5)},\ H^{(0,7)},\ H^{(1,1)},\ H^{(1,3)},\ H^{(1,5)},\ H^{(1,7)},$$

respectively. In fact, these are the first instances of pairs of consecutive elements with difference 42 in each $H^{(\alpha,\beta)}$. □

For the proof of Theorem 2.1 we need a further lemma.

**Lemma 2.4** (A.Papp, L. Hajdu [40]). *Let $p$ be a prime, $u$ be a positive integer and $a_1, \dots, a_{p^u}$ be real numbers. For $t = 1, \dots, u-1$ put*

$$a_i^{(t)} = \frac{1}{p} \sum_{j=1}^{p} a_{(i-1)p+j}^{(t-1)} \quad (i = 1, \dots, p^{u-t})$$

*with $a_i^{(0)} = a_i$ $(i = 1, \dots, p^u)$. Suppose that*

$$\max_{1 \le i \le p^u} |a_i| > C_1 \tag{7}$$

*and*

$$|a_{\ell p+j_1}^{(t)} - a_{\ell p+j_2}^{(t)}| \le C_2 \tag{8}$$

*for any $0 \le t \le u-1$, $0 \le \ell \le p^{u-t} - 1$ and $1 \le j_1 < j_2 \le p$, where $C_1$ and $C_2$ are some real numbers with $C_1 > (u+1)(p-1)C_2/p > 0$. Then $a_1, \dots, a_{p^u}$ have the same sign, and we have*

$$\min_{1 \le i \le p^u} |a_i| > C_1 - (u+1)(p-1)C_2/p.$$

*Proof.* We proceed by induction on $u$. For $u = 1$ conditions (7) and (8) reduce to

$$\max(|a_1|, \dots, |a_p|) > C_1 \quad \text{and} \quad |a_{j_1} - a_{j_2}| \le C_2 \quad (1 \le j_1 < j_2 \le p),$$

respectively. These simply yield

$$\min(|a_1|, \dots, |a_p|) > C_1 - C_2 \ge C_1 - 2(p-1)C_2/p,$$

and it is also clear that $a_1, \dots, a_p$ are of the same sign. So the lemma follows in this case. Assume now that the statement is valid for some $u \ge 1$, for any real numbers $a_1, \dots, a_{p^u}$. Take any real numbers $a_1, \dots, a_{p^{u+1}}$, satisfying the properties (7) and (8), with $u$ replaced by $u+1$; in particular, with $C_1 > (u+2)(p-1)C_2/p > 0$. Put

$$b_i = \frac{1}{p} \sum_{j=1}^{p} a_{(i-1)p+j} \quad (i = 1, \dots, p^u).$$

By (7) and (8) we get that

$$|b_{\ell p+j_1} - b_{\ell p+j_2}| \le C_2$$

for all $0 \le \ell \le p^{u-1} - 1$ and $1 \le j_1 < j_2 \le p$, whence

$$\max_{1\le i\le p^u} |b_i| > \frac{C_1 + (p-1)(C_1 - C_2)}{p} = C_1 - (p-1)C_2/p.$$

Thus the numbers $b_1, \dots, b_{p^u}$ satisfy the conditions (7) and (8), with $C_1$ replaced by $C_1 - (p-1)C_2/p$. Hence by the induction hypothesis we have

$$\min_{1\le i\le p^u} |b_i| > C_1 - (u+1)(p-1)C_2/p,$$

and that the $b_i$-s are of the same sign.

Let $i$ be the index for which $|a_i|$ is minimal. Without loss of generality we may assume that $i = 1$; all the other cases are similar. Then we have

$$C_1 - (u+1)(p-1)C_2/p < |b_1| = \left|\frac{a_1 + \dots + a_p}{p}\right|.$$

This by (8) immediately gives

$$\min_{1\le i\le p^{u+1}} |a_i| = |a_1| \ge C_1 - (u+2)(p-1)C_2/p.$$

It is also clear that the $a_i$-s are of the same sign. Thus the lemma follows. □

Now we are ready to prove Theorems 2.1 and 2.2.

*Proof of Theorem 2.1.* As in the proof of Theorem 2.3, for $(\alpha, \beta) \in I$ and $x > y \ge 0$ set

$$H^{(\alpha,\beta)}(y,x) = H^{(\alpha,\beta)} \cap [y,x) \quad \text{and} \quad h^{(\alpha,\beta)}(y,x) = \frac{|H^{(\alpha,\beta)}(y,x)|}{x-y}.$$

If $y = 0$, we simply write $H^{(\alpha,\beta)}(x)$ and $h^{(\alpha,\beta)}(x)$, respectively. To prove the theorem, by the above notation we need to show that

$$\lim_{x\to\infty} h^{(\alpha,\beta)}(x) = 1/(p-1)p^{a+b-1}$$

for every $(\alpha,\beta) \in I$. For this, first we prove the following assertion: we have

$$\lim_{k\to\infty} h^{(\alpha,\beta)}((j-1)p^{k-b+1}, jp^{k-b+1}) = 1/(p-1)p^{a+b-1} \tag{9}$$

for any $(\alpha,\beta) \in I$ and $1 \le j \le p^{b-1}$. Note that in particular, this immediately gives

$$\lim_{k\to\infty} h^{(\alpha,\beta)}(p^k) = 1/(p-1)p^{a+b-1}$$

for all $(\alpha,\beta) \in I$. To show (9), for $k \ge \max(a,b)$ and $1 \le j \le p^{b-1}$ set

$$A_j^{(k)} = [(j-1)p^{k-b+1}, jp^{k-b+1}),$$

and for $1 \le t \le p-1$ put

$$B_{t,j}^{(k)} = [tp^k + (j-1)p^{k-b+1}, tp^k + jp^{k-b+1}).$$

Observe that we have

$$A_j^{(k+1)} = \bigcup_{i=(j-1)p}^{(j-1)p+p} A_i^{(k)} \quad (j = 1, \dots, p^{b-2})$$

and

$$A_{tp^{b-2}+j}^{(k+1)} = \bigcup_{i=(j-1)p}^{(j-1)p+p} B_{t,i}^{(k)} \quad (t = 1, \dots, p-1,\ j = 1, \dots, p^{b-2}).$$

Set

$$D(k) = \sum_{(\alpha,\beta)\in I} \sum_{j=1}^{p^{b-1}} (x_j^{(k)}(\alpha,\beta))^2,$$

where

$$x_j^{(k)}(\alpha,\beta) = h^{(\alpha,\beta)}((j-1)p^{k-b+1}, jp^{k-b+1}) - 1/(p-1)p^{a+b-1}.$$

Applying Lemmas 2.1 and 2.3 with $1 \le t \le p-1$, the recursive relation for the sets $A_j^{(k)}$ yields that

$$D(k+1) = \sum_{t=0}^{p-1} \sum_{(\alpha,\beta)\in I} \sum_{j=0}^{p^{b-2}-1} \left( \frac{1}{p} \sum_{i=1}^{p} x_{jp+i}^{(k)}(\alpha_{jp+i}^{(t)}, \beta_{jp+i}^{(t)}) \right)^2.$$

Here $(\alpha_{jp+i}^{(t)}, \beta_{jp+i}^{(t)}) \in I$ is given by

$$\alpha_{jp+i}^{(t)} + \frac{t}{1-p} \equiv \alpha \pmod{p^a},\ \gamma_t(j)\beta_{jp+i}^{(t)} \equiv \beta \pmod{p^b},$$

where the $\gamma_t(j)$ are defined in Lemma 2.3 for $t \ge 1$, and $\gamma_0(j) = 1$. Here one should observe that $\ell \in A_j^{(k)}$ if and only if $tp^k + \ell \in B_{t,j}^{(k)}$, and such an $\ell$ belongs to $H^{(\alpha_{jp+i}^{(t)}, \beta_{jp+i}^{(t)})}((j-1)p^{k-b+1}, jp^{k-b+1})$ if and only if $tp^k + \ell$ belongs to $H^{(\alpha,\beta)}(tp^k + (j-1)p^{k-b+1}, tp^k + jp^{k-b+1})$ $(j = 1, \dots, p^{b-1})$. Hence, as for all possible values of $i, j, t$ the pairs $(\alpha_{jp+i}^{(t)}, \beta_{jp+i}^{(t)})$ yield permutations of $I$, by a simple calculation we obtain

$$D(k) - D(k+1) = \sum_{(\alpha,\beta)\in I} \sum_{j=0}^{p^{b-2}-1} \sum_{1\le i_1 < i_2 \le p} \left( \frac{x_{jp+i_1}^{(k)}(\alpha,\beta) - x_{jp+i_2}^{(k)}(\alpha,\beta)}{p} \right)^2. \tag{10}$$

This immediately implies that the sequence $D(k)$ is monotone decreasing. Since clearly, $D(k) \ge 0$ for all $k$, this sequence is convergent; write $\sigma$ for its limit. If $\sigma = 0$, then assertion (9) immediately follows. So assume that $\sigma > 0$. Observe that the definition of $D(k)$ implies that for all $k \ge \max(a, b)$ we have

$$m_k := \max_{\substack{(\alpha,\beta)\in I \\ 1\le j\le p^{b-1}}} |x_j^{(k)}(\alpha,\beta)| > \sqrt{\sigma}/p^{a+b}.$$

Now choose a $k \ge \max(a, b)$ such that $D(k) < \sigma + \mu$, where $\mu$ is to be chosen later. (For the moment, it is sufficient to consider $\mu$

to be 'very small', also with respect to $\sigma$.) Fix $j_0$ and $(\alpha_0, \beta_0)$ such that $m_k = |x_{j_0}^{(k)}(\alpha_0, \beta_0)|$. We shall assume that here $x_{j_0}^{(k)}(\alpha_0, \beta_0) > 0$, the other case is completely similar. Observe that by (10) we have that $D(k+i) - D(k+i+1) < \mu$ for all $i \geq 0$. This, choosing $\mu$ sufficiently small in terms of $\sigma$, by Lemma 2.4 inductively shows that $x_j^{(k)}(\alpha_0, \beta_0) > c_1(\sigma)$ holds for all $j$ with $1 \leq j \leq p^{b-1}$. Here and later on, $c_\ell(\sigma)$ is an explicitly computable positive constant depending only on $\sigma$ (besides $a, b$ and $p$, which are considered to be fixed). At this point we need to distinguish two cases.

Assume first that $p = 2$. Then by Lemmas 2.1 and 2.3, using the recursive definition of the sequence $x_j^{(k)}$, together with a similar argument as above (choosing $\mu$ to be sufficiently small), we get that there exists a $j$ with $1 \leq j \leq 2^{b-1}$ such that $x_j^{(k+1)}(\alpha_0 - 1, \beta') > c_2(\sigma)$, and then that in fact $x_j^{(k+1)}(\alpha_0 - 1, \beta') > c_3(\sigma)$ for all $j$ with $1 \leq j \leq 2^{b-1}$, where $\beta' \in I_2$. Repeating the argument, we get that for all $\alpha^* \in I_1$ there exists a $\beta^* \in I_2$ such that $x_j^{(k+2^a)}(\alpha^*, \beta^*) > c_4(\sigma)$ for all $j$ with $1 \leq j \leq 2^{b-1}$. Now let $(\hat{\alpha}, \hat{\beta}) \in I$ be arbitrary. Then there exists a $\beta^* \in I_2$ such that $x_j^{(k+2^a)}(\hat{\alpha} + 1, \beta^*) > c_4(\sigma)$. Then, since the $\gamma_1(j)$ in Lemma 2.3 yield a permutation of the invertible elements of $\mathbb{Z}_{2^b}$, we see that for the index $j$ defined by $\gamma_1(j)\beta^* \equiv \hat{\beta} \pmod{2^b}$, with the usual argument we obtain that $x_j^{(k+2^a+1)}(\hat{\alpha}, \hat{\beta}) > c_5(\sigma)$. Then repeating the argument once more, we get that in fact $x_j^{(k+2^a+1)}(\hat{\alpha}, \hat{\beta}) > c_6(\sigma)$ for all $j$ with $1 \leq j \leq 2^{b-1}$. This is already sufficient for our purposes; we shall draw the conclusion a bit later, after examining the case of odd primes $p$ as well.

So let now $p$ be an odd prime. Recall that $x_j^{(k)}(\alpha_0, \beta_0) > c_1(\sigma)$ for all $j$ with $1 \leq j \leq p^{b-1}$, and let $\beta^* \in I_2$ be arbitrary. By Lemma 2.3 there exists an $\ell$ depending only on $p$ and $b$, and $\gamma_{t_1}(j_1), \ldots, \gamma_{t_\ell}(j_\ell) \in I_2$ such that $\gamma_{t_1}(j_1) \cdots \gamma_{t_\ell}(j_\ell)\beta_0 \equiv \beta^* \pmod{p^b}$. This by the usual argument gives that for every $\beta^* \in I_2$ one can find an $\alpha^* \in I_1$ such that $x_j^{(k+s)}(\alpha^*, \beta^*) > c_7(\sigma)$ for all $j$ with $1 \leq j \leq p^{b-1}$, with some $s \geq 0$

depending only on $p$ and $b$. Now applying Lemmas 2.1 and 2.3 with $t = p-1$ and $j = p^b - 1$, by the usual argument again, we get that $x_j^{(k+s+1)}(\alpha^* - 1, -\beta^*) > c_8(\sigma)$ for all $j$ with $1 \le j \le p^{b-1}$. Repeating this argument $2p^a$ times, since $p^a$ is odd, we get that $x_j^{(k+s+2p^a)}(\alpha, \beta^*) > c_9(\sigma)$ for all $\alpha \in I_1$ and for all $j$ with $1 \le j \le p^{b-1}$. Since $\beta^* \in I_2$ is arbitrary, we get that in fact $x_j^{(k+s')}(\hat{\alpha}, \hat{\beta}) > c_{10}(\sigma)$ for all $(\hat{\alpha}, \hat{\beta}) \in I$ and $j$ with $1 \le j \le p^{b-1}$, for some $s' \ge 0$ depending only on $a, b$ and $p$.

So in both cases we get $x_j^{(k+s')}(\hat{\alpha}, \hat{\beta}) > c_{11}(\sigma)$ for all $(\hat{\alpha}, \hat{\beta}) \in I$ and $j$ with $1 \le j \le p^{b-1}$, for some $s' \ge 0$ depending only on $a, b$ and $p$. However, this contradicts the identity

$$\sum_{(\alpha,\beta)\in I} x_j^{(k)}(\alpha, \beta) = 0$$

being valid for all $k \ge \max(a, b)$ and $1 \le j \le p^{b-1}$. This proves that $\sigma = 0$, and consequently, $\lim_{k\to\infty} m_k = 0$ (hence also (9)).

Now choose an arbitrary $\varepsilon > 0$, and let $k_0 \ge \max(a, b)$ be such that

$$\left| h^{(\alpha,\beta)}((j-1)p^{k-b+1}, jp^{k-b+1}) - \frac{1}{|I|} \right| < \varepsilon \tag{11}$$

whenever $k \ge k_0$, for any $(\alpha, \beta) \in I$ and $1 \le j \le p^{b-1}$. By (9) we know that such a $k_0$ exists. In fact (11) is valid for any $j \ge 1$. This follows by induction from the fact that by Lemmas 2.1 and 2.3, for any $k \ge k_0$ we have

$$H^{(\alpha,\beta)}(tp^k + (j-1)p^{k-b+1}, tp^k + jp^{k-b+1}) = H^{(\alpha',\beta')}((j-1)p^{k-b+1}, jp^{k-b+1}),$$

with some $(\alpha', \beta') \in I$ and $t, j$ with $0 \le t \le p-1$ and $1 \le j \le p^{b-1}$.

Let $N_0$ be a positive integer to be specified later, and let $N > N_0$. Write

$$N = \sum_{s=1}^{r} c_s p^{f_s}$$

with integers $f_1 > \cdots > f_s \geq 0$ and $0 < c_1, \ldots, c_s < p$. Then for any $(\alpha, \beta) \in I$ we can clearly write

$$\left|H^{(\alpha,\beta)}(N)\right| = \sum_{\ell=1}^{r} \sum_{i=0}^{c_\ell - 1} \left| H^{(\alpha,\beta)} \left( \sum_{g=1}^{\ell-1} c_g p^{f_g} + i p^{f_\ell}, \sum_{g=1}^{\ell-1} c_g p^{f_g} + (i+1) p^{f_\ell} \right) \right|. \tag{12}$$

Let $M$ be the largest multiple of $p^{k_0-b+1}$ with $M \leq N$, and set $q = M/p^{k_0-b+1}$. Using (11), for every $1 \leq j \leq q$ we have

$$\frac{1}{|I|} - \varepsilon < \frac{\left|H^{(\alpha,\beta)}((j-1)p^{k-b+1}, jp^{k-b+1})\right|}{p^{k-b+1}} < \frac{1}{|I|} + \varepsilon,$$

whenever $k \geq k_0$. Write

$$N = \sum_{s=1}^{r} c_s p^{f_s} = \sum_{s=1}^{r'} c_s p^{f_s} + T,$$

where $T = N - M$; observe that $0 \leq T < p^{k_0-b+1}$. Thus, using (11), we have both

$$|H^{(\alpha,\beta)}(N)| \leq \sum_{s=1}^{r'} c_s \left( \frac{p^{f_s}}{|I|} + \varepsilon p^{f_s} \right) + T \leq \frac{N-T}{|I|} + \varepsilon(N-T) + T,$$

and

$$|H^{(\alpha,\beta)}(N)| \geq \sum_{s=1}^{r'} c_s \left( \frac{p^{f_s}}{|I|} - \varepsilon p^{f_s} \right) \geq \frac{N-T}{|I|} - \varepsilon(N-T).$$

Recalling that $h^{(\alpha,\beta)}(N) = |H^{(\alpha,\beta)}(N)|/N$, taking $N_0$ sufficiently large, the theorem follows. □

*Proof of Theorem 2.2.* The statement that for any $(\alpha, \beta) \in I$, $H^{(\alpha,\beta)}$ is relatively dense, would follow from the arguments given in the proof of Theorem 2.1 (in particular, from (9)). However, to give an explicit bound for the largest gap in $H^{(\alpha,\beta)}$, we follow another (though similar) method.

From the proof of Theorem 2.1, for $k \geq \max(a,b)$ we extend the notation

$$A_j^{(k)} = [(j-1)p^{k-b+1}, jp^{k-b+1})$$

from $1 \leq j \leq p^{b-1}$ to $1 \leq j \leq p^b$. For all such $j$, set

$$T_j^{(k)} = \{(\alpha,\beta) \in I : H^{(\alpha,\beta)} \cap A_j^{(k)} \neq \emptyset\}.$$

Observe that by Lemmas 2.1 and 2.3 (see also Remark 2.2), $T^{(k)}_{tp^{b-1}+j}$ is an injective map of $T_j^{(k)}$, for $t = 0, \ldots, p-1$ and $j = 1, \ldots, p^{b-1}$. Further, for any $k \geq \max(a,b)$ we clearly have

$$T_j^{(k+1)} = \bigcup_{i=1}^{p} T^{(k)}_{(j-1)p+i} \quad (j = 1, \ldots, p^{b-1}).$$

For $k$ as before, let

$$s^{(k)} = |T_1^{(k)}| + \cdots + |T^{(k)}_{p^{b-1}}|.$$

Clearly, $s^{(k)}$ is a positive integer with $s^{(k)} \leq p^{b-1}|I| = (p-1)p^{a+2b-2}$. From what we know about the sets $T^{(k)}$ and $T^{(k+1)}$, we easily deduce that $s^{(k+1)} \geq s^{(k)}$, with equality precisely when $T^{(k)}_{(j-1)p+i_1} = T^{(k)}_{(j-1)p+i_2}$ for all $j = 1, \ldots, p^{b-1}$ and $1 \leq i_1 < i_2 \leq p$. Using this observation together with the inductive definition of the sets $T_j^{(k)}$, we see that if for some $k$ we have

$$s^{(k)} = s^{(k+1)} = \cdots = s^{(k+b)},$$

then in fact

$$T_{j_1}^{(k)} = T_{j_2}^{(k)} \quad (1 \leq j_1 < j_2 \leq p^b).$$

This implies that then, with this $k$, we have $s^{(k+\ell)} = s^{(k)}$, for all $\ell \geq 0$. Summarizing the above arguments, we obtain that there exists a $K$ such that $s^{(K+\ell)} = s^{(K)}$ for all $\ell \geq 0$, and that

$$K = \max(a,b) + b(p-1)p^{a+2b-2}$$

is an appropriate choice.

Now we show that with this $K$ we have $T_j^{(K)} = I$ for all $j = 1, \dots, p^b$. For this, it is in fact sufficient to show that this equality holds for all $j = 1, \dots, p^{b-1}$. Note that we already know that $T_{j_1}^{(K)} = T_{j_2}^{(K)}$ for $1 \le j_1 < j_2 \le p^b$. At this point we split our argument into two parts.

Assume first that $p = 2$. Let $(\alpha, \beta) \in T_1^{(K)}$. Then, choosing the $j$ for which $\gamma_1(j) = 1$ in Lemma 2.3, by Lemma 2.1 we see that $(\alpha - 1, \beta) \in T_1^{(K)}$ is also valid (in view of $T_1^{(K)} = T_j^{(K)}$). This shows that for all $\alpha^* \in I_1$, we have $(\alpha^*, \beta) \in T_1^{(K)}$. Let now $(\hat{\alpha}, \hat{\beta}) \in I$ be arbitrary. Choose the $j$ for which $\gamma_1(j)\beta \equiv \hat{\beta} \pmod{2^b}$. Then, since $(\hat{\alpha} + 1, \beta) \in T_1^{(K)}$, we obtain $(\hat{\alpha}, \hat{\beta}) \in T_1^{(K)}$. This proves our claim for $p = 2$.

Suppose next that $p$ is an odd prime. Let $(\alpha, \beta) \in T_1^{(K)}$. Then, in view of $\gamma_{p-1}(p^{b-1}) \equiv -1 \pmod{p^b}$, applying Lemmas 2.1 and 2.3 with $t = p - 1$ and $j = p^b - 1$, we get that $(\alpha - p^a - 1, (-1)^{p^a+1}\beta) \in T_1^{(K)}$, that is, $(\alpha - 1, \beta) \in T_1^{(K)}$. This shows that in fact for all $\alpha^* \in I_1$, we have $(\alpha^*, \beta) \in T_1^{(K)}$. Let now $(\hat{\alpha}, \hat{\beta}) \in I$ be arbitrary. Based upon Lemma 2.3, choose $\gamma_{t_1}(j_1), \dots, \gamma_{t_\ell}(j_\ell) \in I_2$ in Lemma 2.3 such that $\gamma_{t_1}(j_1) \cdots \gamma_{t_\ell}(j_\ell)\beta \equiv \hat{\beta} \pmod{p^b}$. Then we inductively see that $(\alpha', \hat{\beta}) \in T_1^{(K)}$ with some $\alpha' \in I_1$. Indeed, we know that $(\alpha'', \gamma_{t_\ell}(j_\ell)\beta) \in T_{j'}^{(K+s)}$ for some $s$ and $j'$, but then this pair also belongs to $T_1^{(K)}$ - and so on. By what we have proved so far, this yields $(\hat{\alpha}, \hat{\beta}) \in T_1^{(K)}$. Thus our claim follows in this case, too.

So by Lemmas 2.1 and 2.3 we conclude that any interval of the form $[(j-1)p^{K-b+1}, jp^{K-b+1})$ $(j \ge 1)$ contains all elements of $I$. Thus the largest gap in $H^{(\alpha,\beta)}$ cannot be larger than $2p^{K-b+1}$ for any $(\alpha, \beta) \in I$. Hence the theorem follows. □

# 3 Results related to indecomposability of polynomials and Diophantine equations

In this section we study questions related to Diophantine equations of the shape

$$(x - a_1) \cdots (x - a_t) = g(y)$$

where the set $\{a_1, \ldots, a_t\}$ has certain properties. In the first subsection we study the case where we have products of consecutive terms of arithmetic progressions, with one term missing. Then, in the second subsection, we give a further generalization, with many missing terms from the progression in the left hand side.

## 3.1 Shifted power values of products of terms from an arithmetic progression

### 3.1.1 Introduction

A classical result of Erdős and Selfridge [27] says that the product of consecutive positive integers is never a perfect power, that is, the equation

$$x(x+1)\cdots(x+n-1) = y^\ell \tag{13}$$

has no solutions in positive integers $x, n, y, \ell$ with $n \geq 2$ and $\ell \geq 2$. This result and also equation (13) has been generalized into various directions.

The first extension of the problem we mention is when on the left hand side of (13), we omit a term from the product, that is, we consider the equation

$$x(x+1)\cdots(x+j-1)(x+j+1)\cdots(x+n-1) = y^\ell$$

in positive integers $x, n, y, \ell$ with $n \geq 2$ and $\ell \geq 2$, where $0 \leq j \leq n-1$. Confirming a conjecture of Erdős and Selfridge, Saradha and Shorey [60, 61] proved that the only solutions of the above equation are given by

$$\frac{4!}{3} = 2^3, \quad \frac{6!}{5} = 12^2, \quad \frac{10!}{7} = 720^2.$$

The second direction of extensions we mention (which probably attracted the most attention) is when instead of products of consecutive integers one takes products of terms of an arithmetic progression. More precisely, one considers the equation

$$x(x+d)\cdots(x+(n-1)d) = y^\ell$$

in positive integers $x, d, n, y, \ell$ with $n \geq 2$, $\ell \geq 2$ and $\gcd(x, d) = 1$. Under certain (mild, necessary) conditions Darmon and Granville [19] proved that for fixed $n$ and $\ell$, this equation has only finitely many solutions in $x, d, y$. (See also Győry, Hajdu and Saradha [37] for a further generalization.) Recently, Bennett and Siksek [2] proved that if $n$ is large enough, then this equation has only finitely many solutions in $x, d, y, \ell$. On the other hand, for small values of $n$, namely for $3 < n < 35$, a result of Győry, Hajdu and Pintér [36] in accordance with a conjecture of Erdős says that (under certain trivial necessary restrictions) this equation has no solutions at all. We also mention that combining the two directions mentioned above, Saradha and Shorey [62] provided results for equations of the above shape, with one term of the progression missing from the product on the left hand side.

The third direction of extensions we refer to is when in (13) in place of $y^\ell$ on the right hand side we take an arbitrary polynomial, that is, we consider the equation

$$x(x+1)\cdots(x+n-1) = g(y)$$

in integers $x, n, y$, $n > 0$ where $g(y) \in \mathbb{Q}[y]$. Here we recall a result of Kulkarny and Sury [48] who could completely describe when this

equation can have infinitely many solutions in $x, y$, for $n$ fixed. Further, in the particular case where $g(y)$ is of the shape $ay^\ell + b$, Yuan [71] could give effective upper bounds for $x, y$, while Bilu, Kulkarny and Sury [7] could prove an ineffective finiteness theorem for $x, n, y, \ell$. With some extra conditions on $ay^\ell$ Levin (see Section 5 in [50]) proved somewhat more general theorems. We also mention that if $g(y) = \binom{y}{\ell}$ then all solutions are completely described by results of Erdős [23] and Győry [35].

In this subsection we consider common generalizations of these results. Namely, we consider the equation

$$x(x+d)\cdots(x+(j-1)d)(x+(j+1)d)\cdots(x+(n-1)d) = g(y) \quad (14)$$

in integers $x, d, n, y$ with $d \neq 0$, $n \geq 3$ and $0 \leq j \leq n-1$, where $g(y) \in \mathbb{Q}[y]$. Note that the choice $j = 0$ (or $j = n-1$) gives back the classical case, where we have a full product on the left hand side. We shall prove finiteness result concerning equation (14). In the particular case where $g(y)$ is of the form $g(y) = ay^\ell + b$, we are able to provide effective upper bounds for the solutions $x, y$, as well. We mention that there are many results in the literature which are related in the sense that they concern equal values or polynomial values of terms of families of combinatorial polynomials. We cannot survey the extremely huge literature, we only refer to the papers [1, 6, 13, 38, 39, 45, 65] and the references there.

The main tool we use is Baker's method (through results of Schinzel and Tijdeman [63] and Brindza [12]). However, to make it work we need to combine several arguments of combinatorial nature, as well.

### 3.1.2 New results

For the smooth formulation of our main result, we introduce the following notations. Let $n, j, d$ be integers with $d \neq 0$, $n \geq 3$ and

$0 \le j \le n-1$, and put

$$f_{n,j}(x) = x(x+d)\cdots(x+(j-1)d)(x+(j+1)d)\cdots(x+(n-1)d).$$

Note that in the following theorem also the exponent $\ell$ is a variable, so in fact this theorem concerns families of polynomials $g(y)$.

**Theorem 3.1** (Á. Papp, L. Hajdu [42]). *Let $n \ge 8$, $0 \le j \le n-1$ and let $a, b \in \mathbb{Q}$ with $a \ne 0$. Then for all solutions of the equation*

$$f_{n,j}(x) = ay^{\ell} + b \tag{15}$$

*in integers $x, y, \ell$ with $\ell \ge 2$ we have $\max(|x|, |y|, \ell) < C_0$, where $C_0$ is an effectively computable constant depending only on $n, a, b$. Here we use the convention that for $|y| \le 1$ we have $\ell = 2, 3$.*

**Remark 3.1.** One can easily check that we have

$$f_{7,3}(x) = (x^3 + 9dx^2 + 20d^2x + 6d^3)^2 - 36d^6. \tag{16}$$

This shows that (15) with $n = 7$, $j = 3$ and $a = 1$, $b = -36d^6$, $\ell = 2$ has infinitely many integer solutions $x, y$. Thus the assumption $n \ge 8$ in Theorem 3.1 is necessary.

### 3.1.3 Proofs

In our arguments we shall need several lemmas. The first two concern certain properties of the derivatives and shifts of the polynomials $f_{n,j}(x)$. In fact, we can simplify our treatment due to the observation that in these studies the parameter $d$ does not play an important role. That is, instead of the polynomials $f_{n,j}(x)$ it is sufficient to study the polynomials

$$p_{n,j}(x) = x(x+1)\cdots(x+j-1)(x+j+1)\cdots(x+n-1).$$

The reason why this simplification is possible is that we have

$$f_{n,j}(x) = d^{n-1} p_{n,j}(x/d). \tag{17}$$

We start with describing the root structure of the polynomials $p'_{n,j}(x)$ (which is in fact rather simple).

**Lemma 3.1** (Á. Papp, L. Hajdu [42]). *For every $n \geq 3$ and $0 \leq j \leq n-1$, the roots of the polynomial $p'_{n,j}(x)$ are real and simple, and there is a root in each interval*

$$(-n+1, -n+2), (-n+2, -n+3), \ldots,$$
$$\ldots, (-j-2, -j-1), (-j-1, -j+1), (-j+1, -j+2), \ldots, (-1, 0).$$

*Proof.* The statement is a trivial consequence of Rolle's theorem. Note that the cases $j = 0, n-1$ are already treated in the proof of Proposition 3.4 of [6]. □

Our next lemma concerns the common roots of the derivatives and shifts of the polynomials $p_{n,j}(x)$.

**Lemma 3.2** (Á. Papp, L. Hajdu [42]). *For any $n \geq 3$ and $0 \leq j \leq n-1$ we have*

$$\max_{\lambda \in \mathbb{C}} \deg \gcd(p'_{n,j}(x), p_{n,j}(x) - \lambda) \leq 4.$$

*Proof.* If $n \leq 5$ then $\deg p'_{n,j}(x) \leq 4$ and the statement is trivial. So we may assume that $n \geq 6$. Further, for $j = 0, n-1$ the statement immediately follows from Proposition 3.4 of [6]. Thus we may also assume that $0 < j < n-1$. In what follows, we shall use these assumptions without any further mentioning. Write $\alpha_1, \ldots, \alpha_{n-2}$ for the roots of $p'_{n,j}(x)$. By Lemma 3.1 we know that these roots are

distinct, and (renumbering them if necessary) we have

$$-n+1<\alpha_1<-n+2<\cdots<-j-2<\alpha_{n-j-2}<-j-1<$$
$$<\alpha_{n-j-1}<-j+1<\alpha_{n-j}<-j+2<\cdots<-1<\alpha_{n-2}<0.$$

We give an upper bound for the number of $\alpha_i$-s satisfying

$$p_{n,j}(\alpha_i)=\lambda$$

for any fixed $\lambda\in\mathbb{C}$. For this, put

$$P^*_{n,j}(x):=|p_{n,j}(x)|-|p_{n,j}(x-1)|.$$

We would like to calculate the sign changes of $P^*_{n,j}(x)$ inside certain intervals. We note that we take the polynomial $p_{n,j}$ at $x$ and $x-1$ (not at other shifts of $x$) to make this analysis simple – in this way the problem reduces to the study of a quadratic polynomial. Indeed, as we have

$$P^*_{n,j}(x)=(|(x+j-1)(x+n-1)|-|(x-1)(x+j)|)\cdot$$
$$\cdot|x(x+1)\dots(x+j-3)(x+j-2)(x+j+1)(x+j+2)\dots(x+n-3)(x+n-2)|,$$

we may restrict our attention to

$$P_{n,j}(x):=|(x+j-1)(x+n-1)|-|(x-1)(x+j)|.$$

We need to understand the behavior of $P_{n,j}(x)$ (and ultimately of $P^*_{n,j}(x)$) on certain subintervals of $(-n+1,0)$. A simple consideration gives that for $-n+1<x<0$ we have

$$P_{n,j}(x)=\begin{cases}q_1(x), & \text{if } -j+1<x<0,\\ q_2(x), & \text{if } -j<x<-j+1,\\ -q_1(x), & \text{if } -n+1<x<-j,\end{cases}$$

where

$$q_1(x)=2x^2+(n+2j-3)x+(jn-n-2j+1),\ q_2(x)=(1-n)x+(n-jn-1).$$

This shows that $P_{n,j}(x)$, and hence $P^*_{n,j}(x)$ changes sign on the interval $(-n+1, 0)$ at most three times: at the two roots of $q_1(x)$ and between $-j$ and $-j+1$. (Note that the root of $q_2(x)$ is between $-j$ and $-j+1$.) The relevance of this fact is shown by the following observation. Suppose that $P^*_{n,j}(x)$ is positive on some interval $(-i, -i+1)$, with $0 < i \le n-2$. If $i \ge j$, then $\alpha_{n-i} \in (-i, -i+1)$ and $\alpha_{n-i-1} \in (-i-1, -i)$, and we have

$$0 < |p_{n,j}(\alpha_{n-i-1}+1)| - |p_{n,j}(\alpha_{n-i-1})| \le |p_{n,j}(\alpha_{n-i})| - |p_{n,j}(\alpha_{n-i-1})|,$$

that is,

$$|p_{n,j}(\alpha_{n-i})| > |p_{n,j}(\alpha_{n-i-1})|.$$

A similar phenomenon occurs for $i < j$, while in the case where $P^*_{n,j}(x)$ is negative on some interval $(-i, -i+1)$, the above relation just turns around. Altogether, we see that the sequence

$$|p_{n,j}(\alpha_1)|, |p_{n,j}(\alpha_2)|, \ldots, |p_{n,j}(\alpha_{n-2})|$$

changes strict monotonicity at most three times. In other words, the above sequence is the union of at most four strictly monotone sequences, hence it can contain at most four equal terms. This proves our claim. □

To go on we need some new notation. Let $f(x) \in \mathbb{Z}[x]$ of degree $d$ and height (i.e, maximum of the absolute values of the coefficients) $H$, and $a$ be a non-zero integer. Consider the equation

$$f(x) = ay^\ell \tag{18}$$

in $x, y, \ell \in \mathbb{Z}$ with $\ell \ge 2$. Our first lemma is a result of Bérczes, Brindza and Hajdu [3]. Note that the first result of this type is due to Tijdeman [67] and Schinzel and Tijdeman [63].

**Lemma 3.3.** *If $f(x)$ has at least two different roots, then for all solutions $x, y, \ell$ of (18) with $|y| > 1$ we have*

$$\ell < C_1(a, d, H).$$

*Here $C_1(a, d, H)$ is an effectively computable constant depending only on $a, d$ and $H$.*

Our second lemma is the main result of Brindza [12]. To its formulation we need some new notation. Let $S$ be a finite set of primes, and let $\mathbb{Z}_S$ be the set of those rational numbers whose denominators have no prime divisors outside $S$. By the height $h(q)$ of a rational number $q$ we mean the maximum of the absolute value of its denominator and numerator.

**Lemma 3.4.** *Let $f(x) \in \mathbb{Z}[x]$ with*

$$f(x) = a_0 \prod_{i=1}^{m} (x - \gamma_i)^{r_i},$$

*where $a_0$ is the leading coefficient of $f$, and $\gamma_1, \ldots, \gamma_m$ are the distinct complex roots of $f(x)$, with multiplicities $r_1, \ldots, r_m$, respectively. Further, fix $\ell$ with $\ell \geq 2$, and set*

$$t_i = \frac{\ell}{\gcd(\ell, r_i)} \qquad (i = 1, \ldots, m).$$

*Suppose that $(t_1, \ldots, t_m)$ is not a permutation of any of the $m$-tuples*

$$(t, 1, \ldots, 1) \ (t \geq 1), \quad (2, 2, 1, \ldots, 1).$$

*Then for any finite set $S$ of primes, the solutions $x, y \in \mathbb{Z}_S$ of (18) satisfy*

$$\max(h(x), h(y)) < C_2(a, \ell, d, H, S),$$

*where $C_2(a, \ell, d, H, S)$ is an effectively computable constant depending only on $a, \ell, d, H, S$.*

Now we are ready to give the proof of the main result of this subsection.

*Proof of Theorem 3.1.* First observe that in view of (17) Lemma 3.1 implies that $f_{n,j}(x) - b$ has at least two distrinct roots. So Lemma 3.3 shows that $\ell$ is bounded in the requied way. Thus from this point on we may assume that $\ell$ is fixed.

By Lemma 3.4 it is sufficient to prove that the polynomial $f_{n,j}(x) - b$ has more than two zeros of multiplicities coprime to $\ell$. Suppose to the contrary that

$$f_{n,j}(x) - b = p(x) \cdot (q(x))^{\ell} \tag{19}$$

holds with some $p(x), q(x) \in \mathbb{Q}[x]$ and $\deg p \leq 2$. Since by Lemma 3.1 and (17) all the roots of $f'_{n,j}(x)$ are simple, by taking derivative of (19) we immediately get a contradiction for $\ell \geq 3$. That is, we may assume that $\ell = 2$. Then, by taking derivatives of both sides of the above equation, we obtain

$$f'_{n,j}(x) = q(x)(p'(x)q(x) + 2p(x)q'(x)).$$

Let $\alpha_1, \ldots, \alpha_{n-2}$ be the roots of $f'_{n,j}(x)$. Observe that all the roots of $q(x)$ are among them. Further, if $\alpha_i$ is a root of $q(x)$, then (19) yields

$$f_{n,j}(\alpha_i) = b.$$

However, by Lemma 3.2, in view of (17) we see that the above formula may hold for at most four $\alpha_i$-s. That is, we have $\deg q \leq 4$. Hence, we obtain that $n - 1 = \deg f_{n,j}(x) \leq 10$, so $n \leq 11$. That is, we are left with the cases $8 \leq n \leq 11$. Now a simple and tedious computation shows that for these values of $n$, (19) is not possible for any $0 \leq j \leq n - 1$ and $b \in \mathbb{Q}$. We illustrate it by an example. Let $n = 8$, $j = 3$. Then letting $X = x/d$ and $B = b/d^7$ we have

$$f_{8,3}(x) - b = d^7(X(X+1)(X+2)(X+4)(X+5)(X+6)(X+7) - B).$$

The discriminant of $X(X+1)(X+2)(X+4)(X+5)(X+6)(X+7) - B$ (with respect to $X$) is given by

$$-823543B^6 - 116938944B^5 + 40895276544B^4 + 3554646736896B^3$$
$$-448755174604800B^2 - 10577549721600000B + 758487711744000000$$

which is irreducible over $\mathbb{Q}$. That is, $f_{8,3}(x) - b$ has no double roots for any $b \in \mathbb{Q}$, which proves our claim in this case. In all the other cases we came to similar conclusions. Hence, the theorem follows. $\square$

## 3.2 The Prouhet-Tarry-Escott problem, arithmetic progressions, indecomposability of polynomials and Diophantine equations

### 3.2.1 Introduction

The Prouhet-Tarry-Escott problem, shortly PTE, asks to describe disjoint pairs $A$ and $B$ of sets of integers such that their first $k$ power sum symmetric polynomials are equal (cf. [55]). For example, if

$$A = \{2, 3, 7\} \quad \text{and} \quad B = \{1, 5, 6\}$$

then we can take $k = 2$, since we have

$$2 + 3 + 7 = 1 + 5 + 6 \quad \text{and} \quad 2^2 + 3^2 + 7^2 = 1^2 + 5^2 + 6^2. \tag{20}$$

In this subsection we connect the PTE problem, and the question for which polynomials $f(x), g(x) \in \mathbb{Z}[x]$ the equation $f(x) = g(y)$ has infinitely many solutions $(x, y) \in \mathbb{Z}^2$ if the zeros of $f$ are simple and form (almost) an arithmetical progression. Both problems have attracted a lot of attention. Already at this point we mention that the latter question (through a deep result of Bilu and Tichy [8]) is closely related to decomposabilty of polynomials. A polynomial $f(x) \in \mathbb{Q}[x]$ is decomposable if we can write $f(x) = h_1(h_2(x))$ with $h_1, h_2 \in \mathbb{Q}[x]$, in a nontrivial way. (Later we shall give the precise notion.) For example,

$$f(x) = (x-1)(x-2)(x-3)(x-5)(x-6)(x-7)$$

is decomposable, since as one can readily check we have $f(x) = h_1(h_2(x))$ with

$$h_1(x) = (x-2{\cdot}3{\cdot}7)(x-1{\cdot}5{\cdot}6),\ h_2(x) = (x-2)(x-3)(x-7)+2{\cdot}3{\cdot}7. \tag{21}$$

The similarity of (20) and (21) is not a coincidence; in this subsection we show the general connections between these properties. In subsubsection 3.2.2 we form the theorems which describe the connection

between these properties. The proofs of our theorems are given in separate subsections.

There is an extensive literature on binomial coefficients which are equal or differ by a small or fixed constant (see e.g. [31, 65, 69] and the references there). In the latter paper the authors study the related Diophantine equation

$$\binom{f_1(x)}{k} + \binom{x}{2} = \binom{f_2(x)}{2},$$

in polynomials $f_1, f_2 \in \mathbb{Q}[x]$ with $\deg f_1 = 2$, $\deg f_2 = k$. Benne de Weger remarked that this equation leads to the following problem (private communication).

**Problem 3.1.** Let $k \geq 1$. Describe the values of $k$ for which it is possible to partition the set $\{1, \ldots, 2k+1\}$ into a singleton $A_0$ and two sets $A_1$ and $A_2$ with $k = |A_1| = |A_2|$, such that the symmetric polynomials $\sigma_1, \ldots, \sigma_{k-1}$ of the elements of $A_1$ and of $A_2$ coincide.

This is the PTE-problem for $n = 2k+1$. De Weger added that he had solutions for $k = 1, 2, 3$ and had proved that there are none for $4 \leq k \leq 14$. A solution for $k = 3$ is $A = \{2, 3, 7\}$, $B = \{1, 5, 6\}$. Indeed we have

$$2 + 3 + 7 = 1 + 5 + 6,$$

and, by (21),

$$2 \cdot 3 + 2 \cdot 7 + 3 \cdot 7 = \frac{(2+3+7)^2 - (2^2+3^2+7^2)}{2} =$$
$$= \frac{(1+5+6)^2 - (1^2+5^2+6^2)}{2} = 1 \cdot 5 + 1 \cdot 6 + 5 \cdot 6.$$

In this subsection we study the following more general problem.

**Problem 3.2.** Let $r$ be a fixed non-negative integer. Describe those positive integers $n$ for which the set $\{1, \ldots, n\}$ can be partitioned into sets $A_0, A_1, \ldots, A_t$ with $t \geq 2$, $|A_0| = r$ and

$$k := |A_1| = \cdots = |A_t| \geq 2$$

such that all the symmetric polynomials $\sigma_1, \ldots, \sigma_{k-1}$ of the elements of the $A_i$ $(i = 1, \ldots, t)$ coincide.

The problem asks: is it possible to omit a 'few' elements from the set $\{1, \ldots, n\}$ such that the remaining set can be splitted into $t$ subsets which have pairwise the PTE-property? Observe that Problem 3.1 is the special case $r = 1$, $t = 2$.

In Theorem 3.2 we show that if $r$ is small enough with respect to $n$, then only $k = 2$ is possible and $A_1, A_2, \ldots, A_t$ are symmetric. We call a set $A = \{a_1, \ldots, a_k\} \subset \mathbb{R}$ with $a_1 < \cdots < a_k$ symmetric if the sums $a_i + a_{k+1-i}$ $(i = 1, \ldots, k)$ are all equal. It is obvious that such a symmetry implies a PTE-structure.

Next we establish a new link between PTE-problems and the indecomposability of certain polynomials. We recall some standard notions. Let $K$ be a field and $f \in K[x]$. Then $f$ is called decomposable (or composite) over $K$ if there exist $h_1, h_2 \in K[x]$ such that

$$f(x) = h_1(h_2(x)) \quad (h_1, h_2 \in K[x],\ \deg h_1 > 1, \deg h_2 > 1).$$

Otherwise $f$ is called indecomposable. If $f(x) = h_1(h_2(x))$ and $\lambda(x) \in K[x]$ is a linear polynomial, then $f(x) = h_3(h_4(x))$ with $h_3(x) = h_1(\lambda^{-1}(x))$ and $h_4(x) = \lambda(h_2(x))$ is another decomposition of $f(x)$. In the sequel we do not distinguish between such equivalent decompositions. Further, we consider the polynomials $f(x), f(\lambda(x))$, as well as the polynomials $f(x), \lambda(f(x))$ to be equivalent. There is a vast literature on (in)decomposability of polynomials (see e.g. [6, 8, 9, 20, 29, 30, 56] and the references there). In Theorem 3.3 we show that

the studied variant of the PTE problem is equivalent to asking for the indecomposability of certain polynomials.

Using this connection, we show in Corollary 3.1 for given integers $n > r \geq 0$ with $r$ small enough with respect to $n$ that if for $A \subseteq \{1, \ldots, n\}$ with $|A| = n - r$ the polynomial

$$f_{A,c,d}(x) := \prod_{a \in A}(x - c - ad), \quad c, d \in \mathbb{Q},\ d \neq 0 \tag{22}$$

is decomposable over $\mathbb{Q}$ as $h_1(h_2(x))$, then $h_1$ and $h_2$ can be given explicitly. Note that the polynomial $f_{A,c,d}(x)$ represents the product with terms of an arithmetic progression of length $n$ with $r$ terms missing. For example, if

$$f_A(x) := f_{A,0,1}(x) = (x-1)(x-2)(x-3)(x-4)(x-6)(x-7)(x-8)(x-9)$$

is decomposable as $h_1(h_2(x))$, then, apart from equivalence,

$$h_2(x) = x^2 - 10x, \quad h_1(x) = (x+9)(x+16)(x+21)(x+24).$$

Next, using the above results, we establish a finiteness theorem for the number of times that a polynomial $f_{A,c,d}$ of the form (22) assumes a value which is also assumed by a given polynomial $P$ with rational coefficients. In Theorem 3.4 we provide a finiteness result for the number of values of $f_{A,c,d}$ also taken by another polynomial $P(x) \in \mathbb{Q}[x]$. This result, similarly to the above mentioned ones, is ineffective.

Finally, we consider shifted power values (i.e. values of the shape $ay^\ell + b$) of $f_{A,c,d}$. Related problems have been investigated by many authors, the literature has been surveyed in the previous subsection. In particular, we shall generalize our Theorem 3.1. Note however, that in the proof of our more general result, Theorem 3.1 will be play a useful role.

In the equation

$$f_{A,c,d}(x) = ay^\ell + b$$

we give an effective upper bound for the exponent $\ell$ and for the integer values $x, y$ for which this equation holds, in Theorem 3.5. This result implies for example that for every integer $n \geq 24$ and rational numbers $a, b$ with $a \neq 0$ there exists an effectively computable number $C_2$ such that the equation $f_A(x) = ay^\ell + b$ with $A \subset \{1, 2, \ldots, n\}, |A| = n - 2$ implies $\max(|x|, |y|, \ell) < C_2$.

Our results make a step forward towards the solution of the problem how much one can 'mutilate' an arithmetic progression such that the corresponding product of terms still can take only finitely many values of a given polynomial, or shifted power values.

### 3.2.2 New results

In connection with Problem 3.2 we prove the following result.

**Theorem 3.2** (Á. Papp, L. Hajdu, R. Tijdeman [44])**.** *Let $n, r$ be non-negative integers with*

$$n > 2r^{3/2} + 5r + 8. \tag{23}$$

*Then every decomposition of $\{1, \ldots, n\}$ as in Problem 3.2 has the following structure. Putting $A := \{1, \ldots, n\} \setminus A_0$ with $r = |A_0|$, we have $k = 2$, and all classes $A_i = \left\{a_1^{(i)}, a_2^{(i)}\right\}$ $(i = 1, \ldots, t)$ are symmetric with respect to*

$$\bar{a} := \frac{1}{n-r} \sum_{a \in A} a \tag{24}$$

*that is,*

$$a_1^{(i)} + a_2^{(i)} = 2\bar{a} \quad (i = 1, \ldots, t).$$

**Remark 3.2.** Theorem 3.2 yields a complete answer to Problem 3.2 for every $n > 2r^{3/2} + 5r + 8$. On the other hand, for any $r$ and $n$ with $n - r$ even, if $A = \{1, \ldots, n\} \setminus A_0$ is symmetric with respect to $\bar{a}$ (i.e. $a \in A$ implies that $2\bar{a} - a \in A$), then we have a partition as in Problem 3.2 with $k = 2$.

**Remark 3.3.** The following extension of Theorem 3.2 is also valid. Let $b_1, \ldots, b_n$ be a non-constant arithmetic progression in $\mathbb{Q}$. Put $B = \{b_1, \ldots, b_n\}$ and suppose that $B_0, B_1, \ldots, B_t$ is a partition of $B$ such that $r := |B_0|$, $k := |B_1| = \cdots = |B_t|$, $n > 2r^{3/2} + 5r + 8$ and for all $i = 1, \ldots, t$ the symmetric polynomials $\sigma_1, \ldots, \sigma_{k-1}$ of the elements of $B_i$ $(i = 1, \ldots, t)$ are the same. Then $k = 2$ and writing $B_i = \{b_1^{(i)}, b_2^{(i)}\}$ $(i = 1, \ldots, t)$ we have

$$b_1^{(i)} + b_2^{(i)} = b_1^{(j)} + b_2^{(j)} \quad (1 \leq i, j \leq t).$$

Indeed, writing $b_s = c + da_s$ with $a_s \in A \setminus A_0$ and $c, d \in \mathbb{Q}$, $d \neq 0$, it can be easily seen by induction on $n$ that $c$ can be taken to be zero. Then clearly, we may take $d = 1$, and the claim follows.

The next result establishes a link between partitions as in Problem 3.2 and decomposability of certain polynomials.

**Theorem 3.3** (Á. Papp, L. Hajdu, R. Tijdeman [44])**.** *Let $n$ be a positive integer and $r$ a non-negative integer. Then there exists a partition $A_0, A_1, \ldots, A_t$ of $\{1, \ldots, n\}$ as in Problem 3.2 if and only if there exists an $A \subseteq \{1, \ldots, n\}$ with $|A| = n - r$ such that the polynomial*

$$f_A(x) = \prod_{a \in A} (x - a) \tag{25}$$

*is decomposable over $\mathbb{Q}$. In particular, if $A_0, A_1, \ldots, A_t$ is a partition of the required type, then $f_A(x) = h_1(h_2(x))$ with $A = \{1, \ldots, n\} \setminus A_0$ and*

$$h_2(x) = \prod_{a \in A_1} (x - a) - \prod_{a \in A_1} (-a)$$

*and*

$$h_1(x) = \left(x + \prod_{a \in A_1} (-a)\right) \cdots \left(x + \prod_{a \in A_t} (-a)\right).$$

**Remark 3.4.** From the proof of the theorem it will be clear that in fact $h_2$ is independent of which $A_i$ we use in its definition.

As a simple consequence of Theorems 3.2 and 3.3 we obtain the following statement.

**Corollary 3.1** (Á. Papp, L. Hajdu, R. Tijdeman [44]). *Let $A \subseteq \{1, \dots, n\}$ with $|A| = n - r$ where $n$ and $r$ are integers with $r \geq 0$ and $n > 2r^{3/2} + 5r + 8$. Further, let $c, d \in \mathbb{Q}$ with $d \neq 0$. Then the polynomial*

$$f_{A,c,d}(x) = \prod_{a \in A} (x - c - ad) \tag{26}$$

*is decomposable over $\mathbb{Q}$ if and only if $n - r$ is even and $A$ is symmetric with respect to*

$$\bar{a} := \frac{1}{n-r} \sum_{a \in A} a,$$

*when (up to equivalence) the only decomposition of $f_{A,c,d}(x)$ is given by $f_{A,c,d}(x) = \varphi^*((\frac{x-c}{d} - \bar{a})^2)$ with*

$$\varphi^*(x) = d^{n-r} h_1\left(x - \overline{a}^2\right). \tag{27}$$

*Here $h_1$ is the polynomial defined in Theorem 3.3 corresponding to the partition $A_1, \dots, A_t$ of $A$ with $|A_1| = |A_2| = \cdots = |A_t| = 2$.*

Next we apply our results to the equation $f_{A,c,d}(x) = P(y)$ where $P$ is a given polynomial. The first theorem of this type is general, but ineffective: it only guarantees the finiteness of the number of integral solutions.

**Theorem 3.4** (Á. Papp, L. Hajdu, R. Tijdeman [44]). *Let $A \subseteq \{1, \dots, n\}$ with $|A| = n - r$ for integers $r \geq 0$ and $n > 2r^{3/2} + 5r + 8$ and let $c, d \in \mathbb{Q}$ with $d \neq 0$. Let $f_{A,c,d}(x)$ be as in (26) and let $P(y) \in \mathbb{Q}[y]$ with $\deg P \geq 2$. Then the equation*

$$f_{A,c,d}(x) = P(y) \tag{28}$$

*has only finitely many integer solutions $x, y$, unless we are in one of the following cases:*

*(i)* $P(y) = f_{A,c,d}(T(y))$, *where* $T$ *is an arbitrary non-constant polynomial with rational coefficients,*

*(ii)* $P(y) = \varphi^*(Q(y))$, *where* $\varphi^*$ *is given by* (27) *and* $Q$ *is a non-constant polynomial with rational coefficients having at most two roots of odd multiplicities.*

**Remark 3.5.** In cases (i) and (ii) one can easily give examples where equation (28) has infinitely many integer solutions $x, y$.

If the right hand side of (28) is of the shape $ay^\ell + b$ where $\ell$ is also unknown, then we can give an effective result.

**Theorem 3.5** (Á. Papp, L. Hajdu, R. Tijdeman [44])**.** *Let* $A \subseteq \{1, \dots, n\}$ *with* $|A| = n - r$ *with integers* $r \geq 0$ *and* $n > 2r^{3/2} + 5r + 8$ *and let* $c, d \in \mathbb{Q}$ *with* $d \neq 0$. *Let* $f_{A,c,d}(x)$ *be given by* (26) *and let* $a, b \in \mathbb{Q}$ *with* $a \neq 0$. *Then all solutions of the equation*

$$f_{A,c,d}(x) = ay^\ell + b \tag{29}$$

*in integers* $x, y, \ell$ *with* $\ell \geq 2$ *satisfy* $\max(|x|, |y|, \ell) < C_3$ *for some effectively computable constant* $C_3$ *depending only on* $a, b, c, d, n$. *Here we use the convention that for* $|y| \leq 1$ *we have* $\ell \leq 3$.

### 3.2.3 Proofs of results of Prouhet-Tarry-Escott type

*Proof of Theorem 3.2.* Throughout the proof, we shall use the earlier notation: $r = |A_0|$ stands for the number of 'missing elements' from $\{1, \dots, n\}$, $A_0, A_1, \dots, A_t$ form a partition of $\{1, \dots, n\}$ with the prescribed properties and

$$k = |A_1| = \cdots = |A_t|.$$

In particular, we have $n - r = tk$. Observe that (23) implies that $n - 1 > 2(r-1)^{3/2} + 5(r-1) + 8$ for $r > 0$. Therefore, by induction on $r$, we may assume without loss of generality that $n \in A$.

We shall frequently use the identity

$$\sum_{i=j}^{I} \binom{i}{j} = \binom{I+1}{j+1},$$

valid for all $j \geq 0$. Also, we shall make use of the fact that it follows from the conditions of the theorem by induction on $h$ that

$$\sum_{a \in A_i} \binom{a+\ell}{h} = \sum_{a \in A_j} \binom{a+\ell}{h} \quad \text{for } h = 0, 1, \ldots, k-1 \text{ and all } i, j. \tag{30}$$

As we shall see, our choice of $\ell$ will depend on the parity of $k$.

Suppose first $k$ is odd. Then $k \geq 3$. We choose $\ell = -r-2$ in (30) and let

$$f(x) = \binom{x-r-2}{k-1}.$$

Since $\deg f = k-1$, by our assumptions we have

$$\sum_{a \in A_i} f(a) = \sum_{a \in A_j} f(a) \quad (1 \leq i, j \leq t).$$

Recall that $A = A_1 \cup \cdots \cup A_t$ and $n \in A$. Observe that

$$f(1) = \binom{k+r-1}{k-1},\ f(2) = \binom{k+r-2}{k-1}, \ldots, f(r+1) = \binom{k-1}{k-1}.$$

Thus we have

$$\sum_{a \in A} f(a) \leq \sum_{i=k-1}^{n-r-2} \binom{i}{k-1} + \sum_{i=k-1}^{k+r-1} \binom{i}{k-1} = \binom{n-r-1}{k} + \binom{k+r}{k}.$$

Hence for any $j$ with $1 \leq j \leq t$ we get

$$\sum_{a \in A_j} f(a) \leq \frac{\binom{n-r-1}{k} + \binom{k+r}{k}}{t} = \frac{n-r-1}{n-r}\binom{n-r-2}{k-1} + \frac{k}{n-r}\binom{k+r}{k}. \tag{31}$$

On the other hand, assuming without loss of generality that $n \in A_1$, we also have

$$\sum_{a \in A_1} f(a) \geq \binom{n-r-2}{k-1}. \tag{32}$$

Combining (31) and (32), we obtain

$$k\binom{k+r}{k} \geq \binom{n-r-2}{k-1}.$$

Since $k \geq 3$, we can rewrite this inequality as

$$(r+1)(r+2)(r+3)\prod_{i=1}^{k-3}(r+3+i) \geq$$
$$\geq (n-r-2)(n-r-3)\prod_{i=1}^{k-3}(n-r-k-1+i).$$

We show that it is impossible. On the one hand, in view of $k \leq (n-r)/2$ and (23), we have

$$n-r-k-1+i > r+3+i \quad (i = 1, \ldots, k-3).$$

On the other hand, we get from (23) that

$$(r+1)(r+2)(r+3) < (n-r-2)(n-r-3). \tag{33}$$

This yields a contradiction, which proves our claim for $k$ odd.

Suppose $k$ is even and $k \geq 4$. Here we choose $\ell = -2r-2$ in (30) and let

$$f(x) = \binom{x-2r-2}{k-1}.$$

Since $\deg f = k-1$, by our assumptions we have

$$\sum_{a \in A_i} f(a) = \sum_{a \in A_j} f(a) \quad (1 \leq i, j \leq t).$$

Observe that the negative values of $f$ are

$$f(1) = -\binom{k+2r-1}{k-1},\ f(2) = -\binom{k+2r-2}{k-1}, \ldots, f(2r+1) = -\binom{k-1}{k-1}.$$

Thus we have

$$\sum_{a\in A, f(a)<0} |f(a)| \leq \sum_{i=k-1}^{k+2r-1} \binom{i}{k-1} = \binom{k+2r}{k}. \tag{34}$$

Furthermore,

$$\sum_{a\in A, f(a)\geq 0} f(a) \leq \sum_{i=2r+2}^{n} \binom{i-2r-2}{k-1} = \sum_{j=0}^{n-2r-2} \binom{j}{k-1} = \binom{n-2r-1}{k}.$$

Hence for any $j$ with $1 \leq j \leq t$ we get

$$\sum_{a\in A_j} f(a) \leq \frac{\binom{n-2r-1}{k}}{t} = \frac{n-2r-1}{n-r}\binom{n-2r-2}{k-1}. \tag{35}$$

On the other hand, assuming without loss of generality that $n \in A_1$, we also have, by (34),

$$\sum_{a\in A_1} f(a) \geq \binom{n-2r-2}{k-1} - \binom{k+2r}{k}. \tag{36}$$

Combining (35) and (36), we obtain

$$(n-r)\binom{k+2r}{k} \geq (r+1)\binom{n-2r-2}{k-1}.$$

Since $k \geq 4$, we can rewrite this inequality as

$$(n-r)(2r+1)(2r+2)(2r+3)(2r+4)\prod_{i=1}^{k-4}(2r+4+i) \geq$$

$$\geq k(r+1)(n-2r-2)(n-2r-3)(n-2r-4)\prod_{i=1}^{k-4}(n-2r-k-1+i).$$

We show that it is impossible. In view of $k \le (n-r)/2$ and (23), we have

$$n - 2r - k - 1 + i > 2r + 4 + i \quad (i = 1, \dots, k-4).$$

On using $k \ge 4$ and writing $m = n - 2r$ it follows that

$$(m+r)(2r+1)(2r+3)(r+2) > (m-2)(m-3)(m-4). \tag{37}$$

Since (23) implies $m > 2r^{3/2} + 3r + 8$, this yields a contradiction.

Finally, let $k = 2$. Then we have $t = (n-r)/2$; in particular, $n - r$ is even. That is, we have pairs of elements of $A$ having the same sum. Obviously, this is possible only if we take the largest number with the smallest one, and so on, so the pairs are symmetric with respect to $\bar{a}$. □

**Remark 3.6.** Remark 3.3 implies that the symmetric polynomials $\sigma_1, \sigma_2$ of $1, 2, 6$ and of $0, 4, 5$, and also of $3, 5, 13$ and of $1, 9, 11$, coincide too.

### 3.2.4 Proofs of results on indecomposability

*Proof of Theorem 3.3.* Let $A_0, A_1, \dots, A_t$ be a partition as stated in Problem 3.2. Put $A = \{1, \dots, n\} \setminus A_0$ and let $f_A, h_1, h_2$ be as in the theorem. We want to show that $f_A(x) = h_1(h_2(x))$. If two polynomials of degree $n - r$ have the same values at $n - r + 1$ points, then they coincide. It is clear that $f_A(0) = h_1(h_2(0)) = \prod_{a \in A}(-a)$ and that both $f_A(x)$ and $h_1(h_2(x))$ have all $a \in A_1$ as roots. In view of

$$\prod_{a \in A_1}(x-a) - \prod_{a \in A_1}(-a) = \prod_{a \in A_i}(x-a) - \prod_{a \in A_i}(-a)$$

for $2 \le i \le t$, we see that every $a \in A$ is both a root of $f_A(x)$ and of $h_1(h_2(x))$. Thus $f_A(x)$ and $h_1(h_2(x))$ assume the same value at $n-r+1$ points, hence $f_A = h_1(h_2)$. This proves the "only if" statement and the second statement of the theorem.

To prove the "if" statement, let $A \subseteq \{1, \ldots, n\}$ with $|A| = n - r$, and suppose that $h_1(h_2)$ is a decomposition of $f_A$ with $h_1, h_2 \in \mathbb{Q}[x]$. Clearly, we may assume that both $h_1$ and $h_2$ are monic polynomials. Set $h_1(x) = (x - \alpha_1) \ldots (x - \alpha_t)$ with $\alpha_1, \ldots, \alpha_t \in \mathbb{C}$. Observe that these roots are pairwise distinct. Then

$$\prod_{a \in A} (x - a) = h_1(h_2(x)) = (h_2(x) - \alpha_1) \cdots (h_2(x) - \alpha_t).$$

Let $A_i$ consist of the roots of the polynomial $h_2(x) - \alpha_i$ $(i = 1, \ldots, t)$. Then all the symmetric polynomials of the elements of $A_i$ for $i = 1, \ldots, t$ coincide. So putting $A_0 = \{1, \ldots, n\} \backslash A$, the sets $A_0, A_1, \ldots, A_t$ form a partition as in Problem 3.2. □

*Proof of Corollary 3.1.* Clearly, by

$$f_{A,c,d}(x) = d^{n-r} \prod_{a \in A} \left( \frac{x - c}{d} - a \right) = d^{n-r} f_A \left( \frac{x - c}{d} \right),$$

$f_{A,c,d}$ and $f_A$ are equivalent and therefore have equivalent decompositions. It follows from Theorems 3.3 and 3.2 that $f_{A,c,d}$ is decomposable if and only if $n - r$ is even, each partition set $A_i$ has two elements, $a_1^{(i)}$ and $a_2^{(i)}$ for $i = 1, \ldots, t$, say, the set $A$ is symmetric with respect to $\bar{a}$ and (24) holds.

To get the specific decomposition observe that

$$(x - a_1^{(i)})(x - a_2^{(i)}) - a_1^{(i)} a_2^{(i)} = (x - \overline{a})^2 - \overline{a}^2$$

for $i = 1, \ldots, n$. Thus, using the decomposition $f_A = h_1(h_2)$ with $h_1, h_2$ as in Theorem 3.3, we have

$$f_{A,c,d}(x) = d^{n-r} h_1 \left( h_2 \left( \frac{x - c}{d} \right) \right) = d^{n-r} h_1 \left( \left( \frac{x - c}{d} - \overline{a} \right)^2 - \overline{a}^2 \right),$$

so choosing $\varphi^*(x) = d^{n-r} h_1(x - \overline{a}^2)$ we obtain the decomposition as given in the theorem.

To prove the uniqueness let $f_A(x) = P(Bx^2 + Cx + D)$ any decomposition of $f_A(x)$ with $P(x) \in \mathbb{Q}[x]$ and $B, C, D \in \mathbb{Q}, B \neq 0$. Without loss of generality we may assume $B = 1$. Then

$$f_A(x) = P\left((x + C/2)^2 + D - C^2/4\right).$$

Hence the roots of $f_A$ form a symmetric set with respect to $-C/2$, but they also form a symmetric set with respect to $\overline{a}$. Thus $C = -2\overline{a}$. This proves the uniqueness. □

### 3.2.5 Proofs of results on Diophantine equations

We start with the proof of Theorem 3.5. In addition to Lemma 3.4 and Lemma 3.3 from the previous subsection we need another statement taking care of the cases $r \leq 1$.

**Lemma 3.5.** *Let $n, j$ be integers with $n \geq 8$ and $1 \leq j \leq n$, and put*

$$f_{n,j}(x) = \prod_{\substack{i=1 \\ i \neq j}}^{n} (x - i).$$

*Further, let $a, b \in \mathbb{Q}$ with $a \neq 0$. Then for all solutions of the equation*

$$f_{n,j}(x) = ay^\ell + b$$

*in integers $x, y, \ell$ with $\ell \geq 2$ we have $\max(|x|, |y|, \ell) < C_4$, where $C_4$ is an effectively computable constant depending only on $k, a, b$. Here we use the convention that for $|y| \leq 1$ we have $\ell \leq 3$.*

*Proof.* In case of $j = 1$ or $j = n$ the statement follows from the main result of [71], while in the other cases it is a consequence of Theorem 3.1. □

Now we are ready to give the proof of our effective result.

*Proof of Theorem 3.5.* Consider (29) with fixed $A, c, d, a, b$ in integers $x, y, \ell$ with $\ell \geq 2$. Our proof relies on Lemmas 3.3 and 3.4, hence ultimately on the multiplicities of the roots of $f_{A,c,d}(x)$ and its shifts $f_{A,c,d}(x) - b$. Thus as by a simple rational substitution and multiplication by appropriate rationals we can transform $f_{A,c,d}(x)$ into $f_A(x)$, we may consider $f_A(x)$ in place of $f_{A,c,d}(x)$. In view of Lemma 3.5 and $n \geq 9$, we may assume $r = n - |A| \geq 2$ as well.

As all the roots of $f_A(x)$ are simple and real, the same is valid for the polynomial $f'_A(x)$, and consequently for $(f_A(x) - b)'$. Thus the polynomial $f_A(x) - b$ can have at most double roots. Since its degree is $n - r \geq 22$, the statement immediately follows from Lemmas 3.3 and 3.4, unless $\ell = 2$ and $f_A(x)$ is of the form

$$f_A(x) = p(x)(q(x))^2 + b \tag{38}$$

with some $p, q \in \mathbb{Q}[x]$ with $\deg p \leq 2$. In particular,

$$N := |A| = \deg f_A(x)$$

has the same parity as $\deg p$ has. Write $a_1 < \cdots < a_N$ for the elements of $A$. Taking derivatives, (38) gives

$$f'_A(x) = q(x)(p'(x)q(x) + 2p(x)q'(x)). \tag{39}$$

Let $\alpha_1, \ldots, \alpha_{N-1}$ be the roots of $f'_A(x)$. Then by Rolle's theorem these are distinct real numbers with

$$a_i < \alpha_i < a_{i+1} \quad (i = 1, \ldots, N-1).$$

We only consider the case $\deg p = 2$. In fact, it is the most complicated possibility, the other cases are simpler and can be handled similarly. Then clearly, $\deg q = N/2 - 1$, and (39) shows that the roots of $q(x)$ are among the $\alpha_i$-s. Further, (38) implies that for these $\alpha_i$-s we have $f_A(\alpha_i) = b$. Observe that, by (25), $f_A(\alpha_i) < 0$ for $i$ odd, while $f_A(\alpha_i) > 0$ for $i$ even. Altogether, we have two options:

(a) either the roots of $q(x)$ are given by $\alpha_2, \alpha_4, \ldots, \alpha_{N-2}$ (i.e. all the roots with even indices are involved),

(b) or the $N/2-1$ roots of $q(x)$ are among $\alpha_1, \alpha_3, \ldots, \alpha_{N-1}$ (that is, all the roots with odd indices with one exception are involved).

Put

$$G(x) = f_A(x-2) - f_A(x)$$

and set

$$A^* = \{a \in A \ : \ a+2 \in A\}.$$

Observe that $|A^*| \geq N - r - 2$ and

$$G(x) = H(x) \prod_{a^* \in A^*} (x - a^*)$$

with $\deg H \leq r+2$. Further, among the (not disjoint) quadruples

$$\{2i-2, 2i-1, 2i, 2i+1\} \quad (i = 2, \ldots, \lfloor (n-1)/2 \rfloor)$$

at least $n/2 - 2 - 2r$ are subsets of $A$. So by (23), there is a quadruple $2i-2, 2i-1, 2i, 2i+1$ contained in $A$, such that $H(x)$, and thus $G(x)$ has no root in some interval $(2i, 2i+1)$. However, then the sign of $G(x)$ does not change in this interval. If $G(x) > 0$ for $x \in (2i, 2i+1)$ then $f_A(x-2) > f_A(x)$ and choosing $x = \alpha_{2i}$ we have

$$f_A(\alpha_{2i-2}) \geq f_A(\alpha_{2i} - 2) > f_A(\alpha_{2i}).$$

Here we use that $\alpha_{2i-2}$ is the maximum of $f_A$ on $(2i-2, 2i-1)$. If $G(x) < 0$ for $x \in (2i, 2i+1)$ then $f_A(x-2) < f_A(x)$ and choosing $x = \alpha_{2i-2} + 2$ the same reasoning gives

$$f_A(\alpha_{2i-2}) < f_A(\alpha_{2i-2} + 2) \leq f_A(\alpha_{2i}).$$

Hence in both cases

$$f_A(\alpha_{2i-2}) \neq f_A(\alpha_{2i}), \tag{40}$$

which shows that the option (a) above concerning the roots of $q(x)$ is not possible. On the other hand, among the quadruples

$$\{2i-1, 2i, 2i+1, 2i+2\} \quad (i = 1, \ldots, \lfloor n/2 - 1 \rfloor)$$

at least $n/2 - 2r - 2$ are subsets of $A$. So by (23), there are three quadruples as above contained in $A$, such that $H(x)$, and thus $G(x)$ has no root in three distinct intervals $(2i_j + 1, 2i_j + 2)$ $(j = 1, 2, 3)$. Similarly to (40) we obtain

$$f_A(\alpha_{2i_j-1}) \neq f_A(\alpha_{2i_j+1}) \quad (j = 1, 2, 3),$$

which shows that the option (b) above concerning the roots of $q(x)$ is also impossible. □

Now we give the proof of Theorem 3.4. For this we need some more results and notation.

Let $\delta$ be a non-zero rational number and $\mu$ be a positive integer. Then

$$D_\mu(x, \delta) := \sum_{i=0}^{\lfloor \mu/2 \rfloor} d_{\mu,i} x^{\mu-2i} \quad \text{where } d_{\mu,i} = \frac{\mu}{\mu - i} \binom{\mu - i}{i} (-\delta)^i$$

is the $\mu$-th Dickson polynomial. For properties of these polynomials see e.g. [51].

We shall use a deep result of Bilu and Tichy [8] concerning equations of the type

$$f(x) = g(y) \tag{41}$$

in integers $x, y$, where $f, g$ are polynomials with rational coefficients. To describe this result, we introduce some notation. We say that $F, G \in \mathbb{Q}[x]$ form a standard pair over $\mathbb{Q}$ if either $(F(x), G(x))$ or $(G(x), F(x))$ appears in Table 2.

Now we recall the main result of [8], which will play a key role in the proof of Theorem 3.4.

| Kind | Standard pair | Parameter restrictions |
|---|---|---|
| First | $(x^q, \alpha x^p v(x)^q)$ | $0 \le p < q$, $(p,q)=1$, $p + \deg v(x) > 0$ |
| Second | $(x^2, (\alpha x^2 + \beta) v(x)^2)$ | - |
| Third | $(D_\mu(x, \alpha^\nu), D_\nu(x, \alpha^\mu))$ | $\gcd(\mu, \nu) = 1$ |
| Fourth | $(\alpha^{-\mu/2} D_\mu(x, \alpha), -\beta^{-\nu/2} D_\nu(x, \beta))$ | $\gcd(\mu, \nu) = 2$ |
| Fifth | $((\alpha x^2 - 1)^3, 3x^4 - 4x^3)$ | - |

Table 2: Standard pairs. Here $\alpha, \beta$ are non-zero rational numbers, $\mu, \nu, q$ are positive integers, $p$ is a non-negative integer, $v(x) \in \mathbb{Q}[x]$ is a non-zero, but possibly constant polynomial.

**Lemma 3.6.** *Let $f(x), g(x) \in \mathbb{Q}[x]$ be non-constant polynomials. Then the following two statements are equivalent.*

*(I) Equation* (41) *has infinitely many rational solutions $x, y$ with a bounded denominator.*

*(II) We have $f = \varphi \circ F \circ \lambda$ and $g = \varphi \circ G \circ \kappa$, where $\lambda(x), \kappa(x) \in \mathbb{Q}[x]$ are linear polynomials, $\varphi(x) \in \mathbb{Q}[x]$, and $F(x), G(x)$ form a standard pair over $\mathbb{Q}$ such that the equation $F(x) = G(y)$ has infinitely many rational solutions $x, y$ with a bounded denominator.*

*Proof of Theorem 3.4.* By Lemma 3.6, if $f_{A,c,d}(x) = P(y)$ has infinitely many integer solutions then $f_{A,c,d} = \varphi \circ F \circ \lambda$ and $P = \varphi \circ G \circ \kappa$, where $\varphi, \lambda, \kappa$ are rational polynomials with $\deg \lambda = \deg \kappa = 1$, and $F$ and $G$ form a standard pair. By Corollary 3.1, $\deg \varphi \in \{n - r, (n-r)/2, 1\}$. Observe that since the decompositions of the polynomials $f_{A,c,d}$ and $f_A$ are equivalent, we may assume that $c = 0$ and $d = 1$, that is, it is enough to deal with $f_A(x)$. Further, since all quadratic polynomials are equivalent, in view of the case $\ell = 2$ in Theorem 3.5, we may assume without loss of generality that $\deg P \ge 3$. The condition $r > 0$ and

equation (23) implies

$$\deg f_A(x) = N = n - r \geq 15. \tag{42}$$

If $\deg \varphi = n - r$ then $\deg F = 1$, and we easily get that we are in case (i) of Theorem 3.4.

Suppose $\deg \varphi = (n - r)/2$. Then we have $\deg F = 2$. By Corollary 3.1 the decomposition is given up to a linear transformation: $f_A(x) = \varphi^*((x - \overline{a})^2)$. If we have infinitely many solutions then by Lemma 3.6 we have $\varphi^*((x - \overline{a})^2) = P(y) = \varphi^*(G(y))$ for some $G(y) \in \mathbb{Q}[y]$ such that $(x - \overline{a})^2 = G(y)$ has infinitely many solutions. Lemma 3.4 implies that we must be in case (ii).

Finally, consider the case $\deg \varphi = 1$. Then $\deg F = n - r$, and we have

$$f_A(x) = aF(sx + t) + b$$

where $F$ is a member of a standard pair. We check the possible cases.

As $\deg f_A \geq 15$, $F$ cannot come from a standard pair of the fifth kind. Since we assumed that $\deg P \geq 3$, the polynomial $F$ cannot belong to a standard pair of the second type, either.

Assume that $F$ belongs to a standard pair of the first kind. Since all the zeros of $f_A(x)$ are real and simple, hence by Rolle's theorem all the roots of $f'_A(x)$ are real and simple, $F(x) = x^q$ is not possible. On the other hand, if $F(x) = x^p(v(x))^q$, then $f_A$ is of the form

$$f_A(x) = a(s_1x + s_2)^p(v(s_1x + s_2))^q + b$$

with some $s_1, s_2 \in \mathbb{Q}$, $s_1 \neq 0$. Using again that the roots of $f'_A(x)$ are simple, we get $q \leq 2$. However, then in view of that the other term in the standard pair in question is $x^q$, we see that $\deg P \leq 2$, which is excluded.

Finally, assume that $F$ belongs to a standard pair of the third or fourth kind. Then $f_A(x)$ should be a linear transform of a Dickson polynomial. More precisely, with some rationals $s_1, s_2, t_1, t_2$ ($s_1 t_1 \neq 0$) and non-negative integer $N$ we can write

$$t_1 f_A(s_1 x + s_2) + t_2 = D_N(x, \delta),$$

where $D_N(x, \delta)$ is the $N$-th Dickson polynomial, with non-zero parameter $\delta \in \mathbb{Q}$. (Here we apply the inside and outside linear transformations to $f_A$ rather than to $D_N$. In fact, writing $f_A = \varphi \circ D_N \circ \lambda$, $t_1 x + t_2$ and $s_1 x + s_2$ are the inverses of the linear polynomials $\varphi(x)$ and $\lambda(x)$, respectively.) Observe that here $N = \deg f_A(x) = |A|$ must hold. Then, by the well-known identity (see. e.g. formula (2.2) on p. 9 of [51])

$$D_N\left(y + \frac{\delta}{y}, \delta\right) = y^N + \left(\frac{\delta}{y}\right)^N$$

we obtain

$$t_1 \prod_{a \in A} \left(s_1 \left(y + \frac{\delta}{y}\right) + s_2 - a\right) + t_2 = y^N + \left(\frac{\delta}{y}\right)^N.$$

Hence as $|A| = N$,

$$\prod_{a \in A} \left(y^2 + \frac{s_2 - a}{s_1} y + \delta\right) = y^{2N} - t_2 y^N + \delta^N$$

follows. Here we used by comparing the leading coefficients, that $t_1 s_1^N = 1$ must hold. Write $\zeta, \xi$ for the roots of the polynomial $y^2 - t_2 y + \delta^N$. Clearly, $\zeta, \xi$ are algebraic numbers of degree at most two. Further, we have

$$\prod_{a \in A} \left(y^2 + \frac{s_2 - a}{s_1} y + \delta\right) = (y^N - \zeta)(y^N - \xi). \tag{43}$$

If $\zeta_0$, $\xi_0$ are roots of $y^N - \zeta$ and $y^N - \xi$, respectively, then all the roots of these polynomials are given by

$$\zeta_0 \varepsilon^i \quad \text{and} \quad \xi_0 \varepsilon^i \quad (i = 0, 1, \ldots, N-1),$$

respectively, where $\varepsilon$ is a primitive $N$-th root of unity. By (43) we see that all these roots are algebraic numbers of degrees at most two. This immediately gives that the degree of $\varepsilon$ is at most four, hence $\varphi(N) \leq 4$. We conclude that $N \leq 12$. This contradicts (42). $\square$

# 4 Uniform bounds for the number of powers in arithmetic progressions

## 4.1 Introduction

In this section we consider the problem of determining the number of powers among the first $N$ terms of an arithmetic progression.

Let $a, b, \ell$ be integers with $a > 0$ and let $\ell \geq 2$. Write $P_{a,b;N}(\ell)$ for the number of $\ell$-th powers among the first $N$ terms of the arithmetic progression $ax+b$ ($x \geq 0$). Denote by $P_N(\ell)$ the maximum of these values taken over all arithmetic progressions $ax+b$. (Note that this maximum obviously exists.) The case of squares (i.e, $\ell = 2$) has been studied by many authors. Erdős [24] conjectured and Szemerédi [66] proved that $P_N(2) = o(N)$. Later, by deep tools (such as e.g. elliptic and higher genus curves, Faltings' theorem, the distribution of primes etc.) Bombieri, Granville and Pintz [10] proved $P_N(2) < O(N^{2/3+o(1)})$, which subsequently was improved to $P_N(2) < O(N^{3/5+o(1)})$ by Bombieri and Zannier [11]. See also Granville [33] for related results and remarks. A strong conjecture of Rudin (see [58], end of paragraph 4.6) predicts that $P_N(2) = O(\sqrt{N})$, or in an even more precise form, that

$$P_N(2) = P_{24,1;N}(2) = \sqrt{\frac{8}{3}N} + O(1) \quad (N \geq 6) \tag{44}$$

should hold.

In case $\ell \geq 3$ there is hardly anything known. The authors of [10] noted (without proof) that their methods probably make it possible to prove $P_N(3) \ll N^{3/5+\varepsilon}$ and $P_N(\ell) \ll N^{1/2+\varepsilon}$ ($\ell \geq 4$).

In this section we give sharp, in some sense uniform bounds for the number of $\ell$-th powers and arbitrary (mixed) powers among the first $N$ terms of an arithmetic progression, for $N$ large enough.

The origin of this result is a paper by Hajdu and Tengely [46]. They showed that (up to equivalence) for any $\ell \geq 2$ there is a unique arithmetic progression $ax + b$ which contains the most $\ell$-th powers asymptotically, that is, which maximizes the expression

$$\lim_{N\to\infty} \frac{|\{x : ax + b \text{ is an } \ell\text{-th power}, \ 0 \leq x < N\}|}{\sqrt[\ell]{N}}.$$

(In fact, for $\ell = 4$ there are two such progressions.) They could describe these arithmetic progressions $a_\ell x + b_\ell$ explicitly. Based upon their results, they extended Rudin's conjecture (44) for any $\ell \geq 2$ (by replacing $24x + 1$ by $a_\ell x + b_\ell$ and changing the right hand side accordingly), and proved that for $\ell = 3, 4$ for certain small values of $N$. Note that this asymptotic ('global') version of the problem is simpler than the original 'local' one, namely when we concentrate on a finite part of the progressions. The reason is that the asymptotic approach brings in an 'averaging' effect, which roughly speaking makes it possible to concentrate on a complete (finite) period of a progression $ax + b$ modulo $a$.

In this section we prove that for any positive $\varepsilon$ there is an $\ell_0$ depending only on $\varepsilon$ such that for $\ell > \ell_0$ the number of $\ell$-th powers among the first $N$ terms of any integral arithmetic progression is below $(1+\varepsilon)\sqrt[\ell]{N}$, provided that $N$ is large enough in terms of $\varepsilon, \ell$ and the parameters of the progression. The important feature of $\ell_0$ is that it is uniform in the sense that it depends only on $\varepsilon$, it is independent of the progression. This result is sharp in the sense that for infinitely many $\ell$, one can find a constant $c_1 = c_1(\ell) > 1$ and an arithmetic progression having more than $c_1\sqrt[\ell]{N}$ $\ell$-th powers among its first $N$ terms, for all $N$ large enough. We also give a sharp upper bound for the number of powers (with not fixed exponents) among the first $N$ terms of arithmetic progressions. In our proofs we combine a classical result of Wigert [70] concerning the number of divisors of positive integers, the above mentioned result of Hajdu and Tengely [46] concerning arithmetic progressions containing the most $\ell$-th powers asymptotically, and a new assertion answering a

question of Hajdu and Tengely from [46].

We also give an upper bound for the number of powers in arithmetic progressions. For this, let $P_{a,b;N}(*)$ denote the number of (arbitrary) powers among the first $N$ terms of the arithmetic progression $ax + b$ ($x \geq 0$). It will turn out that – as one would predict – here the number of squares is the decisive factor.

## 4.2 New results

Now we give our main results. We use the notation from the introduction.

**Theorem 4.1** (Á. Papp, L. Hajdu [43])**.** *For every $\varepsilon > 0$ there is an $\ell_0$ depending only on $\varepsilon$ such that for any $\ell > \ell_0$ we have $P_{a,b;N}(\ell) \leq (1+\varepsilon)\sqrt[\ell]{N}$, whenever $N > N_0$. Here $N_0 = N_0(\varepsilon, \ell, a, b)$ depends on $\varepsilon, \ell, a, b$.*

**Remark 4.1.** The above theorem is sharp in the sense that $1+\varepsilon$ cannot be replaced by 1, and $\ell > \ell_0$ is also necessary. Indeed, Theorem 1 of [46] (see also the Remarks after it) implies that for infinitely many exponents $\ell \geq 2$ there exists a $\delta_\ell > 0$ and an arithmetic progression $a_\ell x + b_\ell$ with $P_{a_\ell,b_\ell;N}(\ell) > (1+\delta_\ell)\sqrt[\ell]{N}$ for all $N > N_0$. Here $N_0 = N_0(\ell)$ depends only on $\ell$.

It is clear that if an arithmetic progression $ax + b$ contains an $\ell$-th power then it contains infinitely many, and we have

$$P_{a,b;N}(\ell) > \frac{1}{2a}\sqrt[\ell]{N}$$

for $N > N_0$, where $N_0$ depends on $a, b$.

We also mention that on our way to prove Theorem 4.1, we answer a question of Hajdu and Tengely [46] (see Proposition 4.1).

Now we give the theorem concerning the case of mixed powers.

**Theorem 4.2** (Á. Papp, L. Hajdu [43]). *Let $ax + b$ $(x \geq 0)$ be an arithmetic progression. Then for any $\varepsilon > 0$ there exists an $N_0$ such that*

$$P_{a,b;N}(*) < \left(\sqrt{\frac{8}{3}} + \varepsilon\right)\sqrt{N} \tag{45}$$

*for any $N > N_0$. Here $N_0 = N_0(\varepsilon, a, b)$ depends only on $\varepsilon, a, b$.*

**Remark 4.2.** One can easily check (see also e.g. Theorem 1 of [46]) that

$$\lim_{N\to\infty} \frac{P_{24,1;N}(2)}{\sqrt{N}} = \sqrt{\frac{8}{3}}.$$

This shows that the above result is sharp.

Further, it is also easy to see that if $\gcd(a, b) = 1$ then there exist infinitely many exponents $\ell$ such that $ax + b$ contains $\ell$-th powers. Note that here the condition $\gcd(a, b) = 1$ cannot be dropped: for example, the arithmetic progression $4x + 2$ $(x \geq 0)$ contains no powers at all.

## 4.3 Proofs

To prove Theorem 4.1 we shall need some known and new assertions. The next lemma is a result of Hajdu and Tengely [46]. For its formulation, we need to introduce some new notions and notation (which will play important roles also later on). For any $\ell \geq 2$ and arithmetic progression $ax + b$ put

$$M_{a,b}(\ell) := |\{u : 0 \leq u < a,\ u^\ell \equiv b \pmod{a}\}|$$

and $S_{a,b}(\ell) := M_{a,b}(\ell)a^{\frac{1}{\ell}-1}$.

**Lemma 4.1.** *For any $\ell \geq 2$ and for any arithmetic progression $ax+b$ we have $S_{a,b}(\ell) \leq S(\ell)$, where*

$$S(\ell) = \begin{cases} \sqrt{\frac{8}{3}}, & \text{if } \ell = 2, \\ \prod\limits_{p \text{ prime},\ p-1|\ell,\ \frac{\log p}{\log p - \log(p-1)} > \ell} (p-1)p^{\frac{1}{\ell}-1}, & \text{otherwise.} \end{cases}$$

*Proof.* The statement is the first half of Theorem 1 of [46]. □

**Remark 4.3.** The inequality $S_{a,b}(\ell) \leq S(\ell)$ is sharp: for any $\ell$, by an appropriate choice of $ax+b$ (given in [46]) we get equality. Observe that for $\ell$ odd, we have $S(\ell) = 1$.

In the proofs of Theorems 4.1 and 4.2 we shall need the following new assertion. This answers a question of Hajdu and Tengely concerning the limit of the sequence $S(\ell)$ (see the 'concrete question' on p. 966 in the Remarks after Theorem 1 in [46]), and we find it of possible independent interest.

**Proposition 4.1.** *By the notation of Lemma 4.1, for any $\gamma > 0$ there exists an $\ell_1 = \ell_1(\gamma)$ depending only on $\gamma$ such that for $\ell > \ell_1$ we have*

$$S(\ell) < \exp\left(\ell^{-1+\gamma}\right). \tag{46}$$

*In particular, $\lim\limits_{\ell \to \infty} S(\ell) = 1$ holds.*

**Remark 4.4.** One can easily check that (46) implies that

$$S(\ell) < 1 + 2\ell^{-1+\gamma}$$

for $\ell$ large enough.

To prove the above statement, we need the next classical theorem concerning the number of divisors $d(n)$ of a positive integer $n$.

**Lemma 4.2.** *If $\varepsilon > 0$, $X > X_0(\varepsilon)$ then we have*

$$\max_{n \le X} d(n) < \exp\left( (\log 2 + \varepsilon) \frac{\log X}{\log \log X} \right).$$

*Proof.* This is a classical result of Wigert [70]. Note that in [54], p. 56 a stronger form of this assertion is given, however, the above inequality is sufficient for our present purposes. □

*Proof of Proposition 4.1.* As one can easily check by a direct calculation, the function $(t-1)t^{1/\ell-1}$ is strictly monotone increasing for $t > 0$, for any fixed $\ell \ge 3$. Thus, as for $\ell \ge 3$ the product appearing in $S(\ell)$ has at most $d(\ell)$ terms and in every term $p \le \ell + 1$ holds, we have

$$1 \le S(\ell) \le \left( \ell(\ell+1)^{1/\ell-1} \right)^{d(\ell)} < \left( \sqrt[\ell]{\ell} \right)^{d(\ell)}.$$

Here we also used that by the condition $\frac{\log p}{\log p - \log(p-1)} > \ell$, the terms appearing in $S(\ell)$ are greater than 1. Now by Lemma 4.2 we get that

$$S(\ell) < \exp\left( \frac{d(\ell) \log \ell}{\ell} \right) < \exp\left( \frac{\exp\left( \frac{\log \ell}{\log \log \ell} \right) \log \ell}{\ell} \right) =$$

$$= \exp\left( \exp\left( \frac{\log \ell}{\log \log \ell} + \log \log \ell - \log \ell \right) \right) < \exp(\ell^{-1+\gamma})$$

hold, for any $\gamma > 0$ with $\ell > \ell_1$, where $\ell_1 = \ell_1(\gamma)$ depends only on $\gamma$. Thus the first part of the statement is proved. The second part of the claim, taking any $\gamma$ with $0 < \gamma < 1$, from this immediately follows. □

Now we can give the

*Proof of Theorem 4.1.* To bound $P_{a,b;N}(\ell)$, we need to give an upper bound for the number of $\ell$-th powers among the numbers

$$b, a + b, \ldots, a(N-1) + b.$$

In view of that $N_0$ depends on $a, b$, we may assume that $a(N-1)+b \geq 0$. An $\ell$-th power $u^\ell$ belongs to the above terms if its size is 'between' $b$ and $a(N-1)+b$, and $u^\ell \equiv b \pmod{a}$. Thus we see that

$$P_{a,b;N}(\ell) \leq \left(\sqrt[\ell]{aN+|b|} + \sqrt[\ell]{|b|}\right) \frac{M_{a,b}(\ell)}{a} + M_{a,b}(\ell).$$

Here the term in brackets on the right hand side provides an upper bound for the number of (consecutive) integers (forming an interval $I$) with $\ell$-th power of the 'appropriate' size, the factor $M_{a,b}(\ell)/a$ is the ratio of $\ell$-th powers in the residue class $b \pmod{a}$, while the last term is to bound the number of possible $\ell$-th powers in the progression coming from the last part of $I$ (having less than $a$ elements). This yields

$$P_{a,b;N}(\ell) \leq M_{a,b}(\ell) a^{\frac{1}{\ell}-1} \sqrt[\ell]{N} \left( \sqrt[\ell]{1+\frac{|b|}{aN}} + \sqrt[\ell]{\frac{|b|}{aN}} + \frac{a}{\sqrt[\ell]{aN}} \right). \quad (47)$$

Let $\varepsilon > 0$ arbitrary. Clearly, there exists an $N_1 = N_1(\varepsilon, \ell, a, b)$ depending on $\varepsilon, \ell, a, b$ such that for $N > N_1$ we have

$$\sqrt[\ell]{1+\frac{|b|}{aN}} + \sqrt[\ell]{\frac{|b|}{aN}} + \frac{a}{\sqrt[\ell]{aN}} < 1 + \frac{\varepsilon}{2}.$$

By Lemma 4.1 this together with (47) implies

$$P_{a,b;N}(\ell) < \left(1+\frac{\varepsilon}{2}\right) S(\ell) \sqrt[\ell]{N}. \quad (48)$$

In view of Proposition 4.1 we can take an $\ell_0$ such that

$$S(\ell) < \frac{2+2\varepsilon}{2+\varepsilon}$$

for $\ell > \ell_0$. This by (48) yields that

$$P_{a,b;N}(\ell) < (1+\varepsilon)\sqrt[\ell]{N}$$

under the assumptions made for $\ell$ and $N$. Hence our claim follows. □

Now we give the

*Proof of Theorem 4.2.* Throughout the proof we use the phrase '$N$ is large enough' to express that $N$ is larger than an appropriate bound depending only on $\varepsilon, a, b$.

Combining (47) and Lemma 4.1 we obtain

$$P_{a,b;N}(2) < \left(\sqrt{\frac{8}{3}} + \frac{\varepsilon}{2}\right)\sqrt{N} \tag{49}$$

for $\ell = 2$ and

$$P_{a,b;N}(\ell) < S(\ell)(a+3)\sqrt[\ell]{N}$$

for $\ell \geq 3$, respectively, for $N$ large enough. In view of Proposition 4.1, the latter assertion implies that there exists an absolute constant $C_5$ such that

$$P_{a,b;N}(\ell) < C_5(a+3)\sqrt[\ell]{N} \tag{50}$$

for any $\ell \geq 3$, for $N$ large enough. Further, if $N$ is large enough then we have $aN + b \geq |b|$. Hence if $u^\ell$ with $|u| > 1$ belongs to $ax + b$ $(0 \leq x < N)$ then we have

$$\ell \leq \frac{\log(aN + |b|)}{\log 2}.$$

(If $u \in \{-1, 0, 1\}$, then we may assume that $\ell \leq 3$.) This together with (49) and (50) gives

$$P_{a,b;N}(*) \leq \sum_{2 \leq \ell \leq \frac{\log(aN+|b|)}{\log 2}} P_{a,b;N}(\ell) = P_{a,b;N}(2) + \sum_{3 \leq \ell \leq \frac{\log(aN+|b|)}{\log 2}} P_{a,b;N}(\ell) <$$

$$< \left(\sqrt{\frac{8}{3}} + \frac{\varepsilon}{2}\right)\sqrt{N} + C_5(a+3)\frac{\log(aN+|b|)}{\log 2}\sqrt[3]{N} < \left(\sqrt{\frac{8}{3}} + \varepsilon\right)\sqrt{N}$$

for $N$ large enough. This proves the statement. □

# References


[1] A. Bazsó, A. Bérczes, L. Hajdu, F. Luca, *Polynomial values of sums of products of consecutive integers*, Monat. Math. **187** (2018), 21–34.

[2] M. Bennett, S. Siksek, *A conjecture of Erdős, supersingular primes and short character sums*, Annals of Mathematics **191** (2020), 355–392

[3] A. Bérczes, B. Brindza, L. Hajdu, *On the power values of polynomials*, Publ. Math. Debrecen **53** (1998), 375–381.

[4] D. Berend, *On the parity of exponents in the factorization of* $n!$, J. Number Theory **64** (1997), 13–19.

[5] D. Berend and G. Kolesnik, *Regularity of patterns in the factorization of* $n!$, J. Number Theory **124** (2007), 181–192.

[6] F. Beukers, T. N. Shorey and R. Tijdeman, *Irreducibility of polynomials and arithmetic progressions with equal product of terms*, in: Number Theory in Progress (Proc. Internat. Conf. in Number Theory in Honor of A. Schinzel, Zakopane, 1997), K. Győry, H. Iwaniec and J. Urbanowicz (eds.), de Gruyter, 1999, pp. 11–26.

[7] Yu. Bilu, M. Kulkarni, B. Sury, *The Diophantine equation* $x(x+1)\dots(x+(m-1))+r=y^n$, Acta Arith. **113** (2004), 303–308.

[8] Yu. Bilu, R. Tichy, *The Diophantine equation* $f(x)=g(y)$, Acta Arith. **95** (2000), 261–288.

[9] R. Blankertz, *A polynomial time algorithm for computing all minimal decompositions of a polynomial*, ACM Comm. Computer Algebra 48:1, Issue 187 (2014), 13–23.

[10] E. Bombieri, A. Granville and J. Pintz, *Squares in arithmetic progressions*, Duke Math. J. **66** (1992),369–385.

[11] E. Bombieri and U. Zannier, *A note on squares in arithmetic progressions. II*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei **13** (2002), 69–75.

[12] B. Brindza, *On S-integral solutions of the equation* $y^m = f(x)$, Acta Math. Hungar. **44** (1984), 133–139.

[13] B. Brindza, Yu. Bilu, P. Kirschenhofer, Á. Pintér, R. Tichy, *Diophantine equations and Bernoulli polynomials*. With an appendix by A. Schinzel. Compositio Math. **131** (2002), 173–188.

[14] W. Bosma, J. Cannon and C. Playoust, *The Magma algebra system. I. The user language*, J. Symbolic Comput. **24** (1997), 235–265.

[15] R. Breusch, *Zur Verallgemeinerung des Bertrandschen Postulates, das zwischen* $x$ *und* $2x$ *stets Primzahlen liegen*, Math. Zeitschrift **34** (1932), 505–526.

[16] U. Cerucci, F. Vaccarino, *Vector Linear Recurrence Sequences in Commutative Rings*, Applications of Fibonacci Numbers, 1996, pp. 63–72.

[17] Y.-G. Chen, *On the parity of exponents in the standard factorization of* $n!$, J. Number Theory **100** (2003), 326–331.

[18] Y.-G. Chen and Y.-C. Zhu, *On the prime power factorization of* $n!$, J. Number Theory **82** (2000), 1–11.

[19] H. Darmon, A. Granville, *On the equations* $z^m = F(x, y)$ *and* $Ax^p + By^q = Cz^r$, Bull. London Math. Soc. **27** (1995), 513–543.

[20] H. Davenport, D.J. Lewis, A. Schinzel, *Equations of the form* $f(x) = g(y)$, Quart. J. Oxford Ser. **12** (1961), 304–312.

[21] J.-M. Deshouillers, F. Luca, *How often is $n!$ a sum of three squares?*, The Legacy of Alladi Ramakrishnan in the Mathematical Science (2010), pp. 243–251.

[22] P. Erdős, *Über die Primzahlen gewisser arithmetischer Reihen*, Math. Zeitschrift **39** (1935), 473–491.

[23] P. Erdős, *On a Diophantine equation*, J. London Math. Soc. **26** (1951), 176–178.

[24] P. Erdős, *Quelques problémes de théorie des nombres*, Monographies de L'Enseignement Mathématique **6** (1963), pp. 81–135.

[25] P. Erdős and R. L. Graham, *Old and New Problems and Results in Combinatorial Number Theory*, L'Enseignement Mathématique, Imprimerie Kundig, Geneva, 1980.

[26] P. Erdős and R. Obláth, *Über diophantische Gleichungen der Form $n! = x^p \pm y^p$ und $n! \pm m! = x^p$*, Acta Litt. Sci. Szeged **8** (1937), 241–255.

[27] P. Erdős, J. L. Selfridge, *The product of consecutive integers is never a power*, Ill. J. Math. **19** (1975), 292–301.

[28] G. Everest, A. van der Poorten, I. Shparlinski, T. Ward, *Recurrence sequences*, Third Edition, American Matematical Society (2015), pp. 318.

[29] M. Fried, *On a theorem of Ritt and related Diophantine problems*, J. Reine Angew. Math. **264** (1973), 40–55.

[30] M. Fried, *Variables separated polynomials, the genus 0 problem and moduli spaces*, Number Theory in Progress, vol. 1, Walter de Gruyter, Berlin, pp. 169–228 (1999).

[31] H. R. Gallegos-Ruiz, N. Katsipis, Sz. Tengely, M. Ulas, *On the Diophantine equation $\binom{n}{k} = \binom{m}{l} + d$*, J. Number Theory **208** (2020), 418–440.

[32] M. Z. Garaev, F. Luca and I. Shparlinski, *Character sums and congruences with $n!$*, Trans. Amer. Math. Soc. **356** (2004), 5089–5102.

[33] A. Granville, *Squares in Arithmetic Progressions and Infinitely Many Primes*, Amer. Math. Monthly **124** (2017), 951–954.

[34] R. K. Guy, *Unsolved Problems in Number Theory*, Third Edition, Springer (2004).

[35] K. Győry, *On the Diophantine equation $\binom{n}{k} = x^l$*, Acta Arith. **80** (1997), 289–295.

[36] K. Győry, L. Hajdu, Á. Pintér, *Perfect powers from products of consecutive terms in arithmetic progression*, Compositio Math. **145** (2009), 845–864.

[37] K. Győry, L. Hajdu, N. Saradha, *On the Diophantine equation $n(n+d)\cdots(n+(k-1)d) = by^l$*, Canad. Math. Bull. **47** (2004), 373–388. Correction: ibid. **48** (2005), 636.

[38] K. Győry, T. Kovács, Gy. Péter, Á. Pintér, *Equal values of standard counting polynomials*, Publ. Math. Debrecen **84** (2014), 259–277.

[39] L. Hajdu, S. Laishram, Sz. Tengely, *Power values of sums of products of consecutive integers*, Acta Arith. **172** (2016), 333–349.

[40] L. Hajdu, Á. Papp, *On asymptotic density properties of the sequence $(n!)_{n=0}^{\infty}$*, Acta Arith. **184** (2018), 317–340.

[41] L. Hajdu, Á. Papp, T. Szakács, *On the equation $A!B! = C!$*, J. Number Theory **187** (2018), 160–165.

[42] L. Hajdu, Á. Papp, *Polynomial values of products of terms from an arithmetic progression*, Monatshefte für Mathematik **193 (3)** (2020), 637–655.

[43] L. Hajdu, Á. Papp, *Uniform bounds for the number of powers in arithmetic progressions*, Rev. Real Acad. Cienc. Exactas Fis. Nat. Ser. A-Mat. **116** (2022), 169.

[44] L. Hajdu, Á. Papp, R. Tijdeman, *The Prouhet-Tarry-Escott Problem, Indecomposability of Polynomials and Diophantine Equations*, The Ramanujan J. **58** (2022), 1075–1093.

[45] L. Hajdu, Á. Pintér, Sz. Tengely, N. Varga, *Equal values of figurate numbers*, J. Number Theory, **137** (2014), 130–141.

[46] L. Hajdu and Sz. Tengely, *Powers in arithmetic progressions*, The Ramanujan J. **55** (2021), 965–986.

[47] O. Klurman and M. Munsch, *Distribution of factorials modulo p*, J. Th. Nomb. Bordeaux **29** (2017), 169–177.

[48] M. Kulkarni, B. Sury, *On the Diophantine equation* $x(x+1)(x+2)\cdots(x+(m-1)) = g(y)$, Indag. Math. **14** (2003), 35–44.

[49] E. Landau, *Über die Einteilung der positiven ganzen Zahlen in vier Klassen nach der Mindestzahlder zu ihrer additiven Zusammensetzung erforderlichen Quadrate*, Arch. Math. Phys. **13** (1908), 304–312.

[50] A. Levin, *Variations on the theme of Runge: effective determination of integral points on certain varieties*, J. Th. Nomb. Bordeaux **20** (2008) no 2, 385–417.

[51] R. Lidl, G. Mullen, G. Turnwald, *Dickson polynomials*, Pitman Monographs and Surveys in Pure and Applied Mathematics, Longman Scientific & Technical, Harlow **65** (1993).

[52] F. Luca and P. Stănică, *Products of factorials modulo p*, Colloq. Math. **96** (2003), 191–205.

[53] F. Luca and P. Stănică, *On the prime power factorization of n!*, J. Number Theory **102** (2003), 298–305.

[54] H. L. Montgomery and R. C. Vaughan, *Multiplicative Number Theory, I. Classical Theory*, Cambridge Univ. Press, Cambridge (2007).

[55] S. Raghavendran, V.Varayanan, *The Prouhet Tarry Escott problem: A review*, MDPI, Mathematics **7** (2019), 227.

[56] J. F. Ritt, *Prime and composite polynomials*, Trans. Amer. Math. Soc. **23** (1922), 51–66.

[57] B. Rokowska and A. Schinzel, *Sur un problème de M. Erdős*, Elem. Mat. **15** (1960), 84–85.

[58] W. Rudin, *Trigonometric series with gaps*, J. Math. Mech. **9** (1960), 203–227.

[59] J. W. Sander, *On the Parity of Exponents in the Prime Factorization of Factorials*, J. Number Theory **90** (2001), 316–328.

[60] N. Saradha, T. N. Shorey, *Almost perfect powers in arithmetic progression*, Acta Arith. **99** (2001), 363–388.

[61] N. Saradha, T. N. Shorey, *Almost squares and factorizations in consecutive integers*, Compositio Math. **138** (2003), 113–124.

[62] N. Saradha, T. N. Shorey, *On the equation* $n(n+d)\cdots(n+(i_0-1)d)(n+(i_0+1)d)\cdots(n+(k-1)d) = y^l$ *with* $0 < i_0 < k-1$, Acta Arith. **129** (2007), 1–21.

[63] A. Schinzel, R. Tijdeman, *On the equation* $y^m = P(x)$, Acta Arith. **31** (1976), 199–204.

[64] T. Shorey and R. Tijdeman, *Arithmetic Properties of Blocks of Consecutive Integers*, in: From Arithmetic to Zeta-Functions, Number Theory in Memory of Wolfgang Schwarz (J. Sander, J. Steuding, R. Steuding, eds) (2017), pp. 455–471.

[65] T. Stoll, R. F. Tichy, *The Diophantine equation* $\alpha\binom{x}{m} + \beta\binom{y}{n} = \gamma$, Publ. Math. Debrecen **64** (2004), 155–165.

[66] E. Szemerédi, *The number of squares in an arithmetic progression*, Stud. Sci. Math. Hungar. **9** (1974), 417.

[67] R. Tijdeman, *Applications of the Gel'fond-Baker method to rational number theory*, Topics in Number Theory, Proceedings of the Conference at Debrecen 1974, Colloq. Math. Soc. János Bolyai **13**, pp. 399–416, North-Holland, Amsterdam, 1976.

[68] S. S. Wagstaff, *The Schnirelmann density of the sums of three squares*, Proc. Amer. Math. Soc. **52** (1975), 1–7.

[69] B. de Weger, *Equal Binomial Coefficients: Some Elementary Considerations*, J. Number Theory **63** (1997), 373–386.

[70] S. Wigert, *Sur l'ordre de grandeur du nombre des diviseurs d'un entier*, Ark. Mat. **3** (1906/7), 1–9.

[71] P.-Z. Yuan, *On a special diophantine equation* $a\binom{x}{m} = by^r + c$, Publ. Math. Debrecen **44** (1994), 137–143.

# Összefoglaló

Jelen disszertáció a diofantikus számelmélet három témakörében mutat be új eredményeket. Az első rész faktoriálisokról, különböző részhalmazaikról és ezek három négyzetszámként való előállításáról szól. A második számtani sorozatok egymást követő tagjai szorzatai és polinomok közös értékeivel foglalkozik. A harmadik pedig számtani sorozatok adott számú egymást követő tagjaiban szereplő hatványok számára vonatkozó eredményeket mutat be.

Elsőként faktoriálisok aritmetikai tulajdonságait vizsgáltuk. Leírtuk $n!$-ban a $p$ prím kitevőjének (modulo $p^a$) és a $p$-mentes résznek a (modulo $p^b$) viselkedését ($a, b$ egészek mellett). Bebizonyítottuk, hogy az $a, b$ paraméterek tetszőleges megválasztása esetén azon $n$-ek halmazának aszimptotikus sűrűsége, melyekre $n!$ adott mintázatú éppen $1/(p-1)p^{a+b-1}$. Szintén bebizonyítottuk, hogy ugyanezen halmaz relatív sűrű, és a halmaz egymást követő tagjainak távolsága legfeljebb $2p^{\max(a,b)+b(p-1)p^{a+2b-2}-b+1}$. A $p = 2, a = 1, b = 3$ eset kiemelten fontos, hisz kapcsolatban van $n!$ három négyzetszám összegeként való előállításával. Deshouillers és Luca megmutatták, hogy a $(0 <)x$-nél nem nagyobb, ilyen módon előállítható $n$-ek sűrűsége $(1/8)x + \mathcal{O}(x^{2/3})$. Munkájukban szóelméleti eszközöket használtak. Mi valamivel általánosabb és pontosabb eredményt igazoltunk: a paraméterek tetszőleges megválasztása mellett a keresett halmaz sűrűsége $1/8x+\mathcal{O}(x^{1/2}\log^2 x)$. Igazoltuk továbbá, hogy ez esetben a szomszédos tagok különbsége legfeljebb 42. Mindezekhez algebrai és kombinatorikai módszereket használtunk.

A második vizsgált probléma polinomok és számtani sorozatok egymást követő tagjai szorzatainak közös értékei. A témakör legfontosabb tétele Erdős és Selfridge eredménye, mely szerint számtani sorozatok egymást

követő tagjainak szorzata sosem teljes hatvány. Ezt sokan sok irányban általánosították, többek között Shorey, Tijdeman, Saradha, Győry, Bennett, Levin, Hajdu, Pintér, Tengely és Siksek. Az általunk bemutatott új eredmények a korábbi eredmények közös általánosítását jelentik. Nevezetesen, különböző végességi eredményeket adtunk a

$$\prod_{a\in A}(x-c-ad)=g(y),$$

egyenlet $x, y$ egész megoldásaira, ahol $g(y) \in \mathbb{Z}[y]$, $c, d \in \mathbb{Q}$ és $A = \{1, \dots, n\} \setminus A_0$, $|A_0| = r$ valamely $n$ pozitív egészre. Azaz, az egyenlet bal oldalán egy számtani sorozat egymást követő tagjainak szorzata áll, esetleg néhány tag elhagyásával. Először $\max(|x|, |y|, \ell)$-re adtunk effektív felső korlátot abban az esetben, amikor a bal oldalon egy elem hiányzik és $g(y) = ay^\ell + b$ ($a, b, \ell$ egészek, $\ell \geq 2$). Ezután bemutattunk egy újonnan felfedezett szoros kapcsolatot a Prouhet-Terry-Escott probléma és a polinomok dekomponálhatóságának kérdése között. Egyfelől bebizonyítottuk, hogy $A$ minden PTE-particionálása csak a triviális szimmetrikus lehet, ha $n > 2r^{3/2} + 5r + 8$. Másfelől, pontosan akkor létezik PTE-particionálás, ha a $\prod_{a\in A}(x-a)$ polinom dekomponálható $\mathbb{Q}$ felett. Egy alkalmazásként effektív végességi eredményt igazolunk abban az esetben, ha a fenti egyenlet bal oldalán legfeljebb $r$ elem hiányzik melyre $n > 2r^{3/2}+5r+8$, és a jobb oldalon $g(y) = ay^\ell+b$, valamint ineffektív végességi eredményt néhány feltétel mellett, ha a jobb oldalon tetszőleges (legalább másodfokú) polinom áll.

A harmadik terület $\ell$-edik hatványok számtani sorozatok $N$ egymást követő tagjaiban való $P_\ell(N)$ számáról szól. Itt is egy Erdős-sejtés a kiindulópont, mely szerint $P_2(N) = o(N)$. Ezt Szemerédi igazolta. Rudin egy sokkal erősebb sejtése, hogy $P_2(N) = O(\sqrt{N})$. Bár a sejtés még nyitott, sok mély eredmény született már vele kapcsolatban Bombieri, Granville, Pintz, Zannier és mások által. Nemrégiben Hajdu és Tengely kiterjesztette a kutatásokat $\ell = 2$-ről az általános esetre bizonyítva, hogy minden $\ell$-re létezik egy ,,legjobb" számtani sorozat, mely aszimptotikusan a legtöbb $\ell$-edik hatványt tartalmazza. Ezt az irányt foly-

tatva először bizonyítottuk, hogy minden pozitív $\varepsilon$ esetén van egy csak tőle függő $\ell_0$, hogy ha $\ell > \ell_0$, akkor az $\ell$-edik hatványok száma bármely számtani sorozat első $N$ tagja között legfeljebb $(1+\varepsilon)\sqrt[\ell]{N}$, feltéve, hogy $N$ elég nagy $\varepsilon, \ell$ és a sorozat függvényében. Fontos, hogy $\ell_0$ csak $\varepsilon$-tól függ, a sorozattól nem. Másodszor korlátot adtunk (tetszőleges) hatványok számára egy számtani sorozat első $N$ tagja között. Ez a korlát $\left(\sqrt{8/3}+\varepsilon\right)\sqrt{N}$, ha $N$ elég nagy $\varepsilon$ és a sorozat függvényében. Bizonyításunkban Wiegert egy klasszikus, osztók számát becslő eredményét, Hajdu és Tengely korábban említett eredményét, és egy új becslést használtunk.

# Summary

This dissertation contains new results in Diophantine number theory in three topics. The first part is about factorials and densities of certain subsets connected to the classical problem of representing factorials as sums of three squares. The second part concerns the problem of common values of products of consecutive terms of arithmetic progressions and various polynomials. The third part deals with the number of powers in a fixed number of consecutive terms of arithmetic progressions.

The first topic we considered was related to arithmetic properties of factorials. We described the simultaneous behavior of the exponent of $p$ in $n!$ (modulo $p^a$) and the $p$-free part of $n!$ (modulo $p^b$) for prime $p$ and $a, b$ positive integers. We proved that for any choice of $a$ and $b$ the sets of $n$ such that $n!$ have a valid pattern, have equal assymptotic density $1/(p-1)p^{a+b-1}$. We also proved that the corresponding sets are relative dense, that is, the difference between consecutive elements is bounded by $2p^{\max(a,b)+b(p-1)p^{a+2b-2}-b+1}$. The case $p = 2$ with $a = 1, b = 3$ has special importance as this is connected to the representation of $n!$ as a sum of three squares. Deshouillers and Luca showed that the set of such $n$ up to $x > 0$ has density $1/8x + \mathcal{O}(x^{2/3})$ using word theory. We proved a somewhat more general result saying that for any choice of residues the corresponding set has density $1/8x + \mathcal{O}(x^{1/2}\log^2 x)$. We also proved that the set is relative dense with upper bound 42. In the proofs we combined various ideas of algebraic and combinatorial nature.

The second problem we discussed is about the common values of polynomials and products of consecutive terms of arithmetic progressions. The most important theorem in the field is that of Erdős and Self-

ridge, saying that the product of consecutive positive integers is never a perfect power. This result has been generalized into many directions by several authors including but not restricted to Shorey, Tijdeman, Saradha, Győry, Bennett, Hajdu, Pintér, Tengely, Siksek, Levin. We gave new results which concern a common generalization of these directions, and extend many related papers from the literature. Namely, we provided various finiteness results for the integer solutions $x, y$ of the generalized problem, for equations of the form

$$\prod_{a \in A}(x - c - ad) = g(y),$$

where $g(y) \in \mathbb{Z}[y]$, $c, d$ are rationals and $A = \{1, \ldots, n\} \setminus A_0$, $|A_0| = r$ for some positive integer $n$. That is, the product on the left hand side is taken over some terms of an arithmetic progression but skipping one or a few terms. First, we gave an effective bound for $\max(|x|, |y|, \ell)$ in the case where one term is missing on the left side and $g(y) = ay^\ell + b$ (and $a, b, \ell$ are integers, $\ell \geq 2$). Secondly, we discovered a deep link between the Prouhet-Terry-Escott problem and the indecomposability of polynomials. On the one hand, we proved that every PTE-partition of $A$ is the trivial symmetrical if $n > 2r^{3/2} + 5r + 8$. On the other hand, there exists a PTE-partition if and only if the polynomial $\prod_{a \in A}(x - a)$ is decomposable over $\mathbb{Q}$. As an application we gave an effective result on the equation above if on the left hand side no more than $r$ terms are missing with $n > 2r^{3/2} + 5r + 8$, and an effective bound on the solutions if $g(y) = ay^\ell + b$.

The third field concerns upper bounds for the number $P_\ell(N)$ of $\ell$-th powers among $N$ consecutive terms of an arithmetic progression. Here, the starting point is a conjecture of Erdős, predicting that $P_2(N) = o(N)$, which has been proved by Szemerédi. A much stronger conjecture is due to Rudin, saying that $P_2(N) = O(\sqrt{N})$. This conjecture is still open, in spite of several related papers and deep results of Bombieri, Granville, Pintz, Zannier and others. Recently, Hajdu and Tengely extended the research from $\ell = 2$ to the general case, and

proved that for any $\ell$ there is a 'best' arithmetic progression, containing the most $\ell$-th powers asymptotically. Extending their results we first proved that for any positive $\varepsilon$ there is an $\ell_0$ depending only on $\varepsilon$ such that for $\ell > \ell_0$ the number of $\ell$-th powers among the first $N$ terms of any integral arithmetic progression is below $(1+\varepsilon)\sqrt[\ell]{N}$, provided that $N$ is large enough in terms of $\varepsilon, \ell$ and the parameters of the progression. The important feature of $\ell_0$ is that it is uniform in the sense that it depends only on $\varepsilon$, but it is independent of the progression. Secondly, we gave sharp upper bound for the number of powers (with not fixed exponents) among the first $N$ terms of arithmetic progressions showing that their number is bounded by $\left(\sqrt{8/3}+\varepsilon\right)\sqrt{N}$ if $N$ is large enough depending on $\varepsilon$ and the progression. In our proofs we combined a classical result of Wigert concerning the number of divisors of positive integers, the above mentioned result of Hajdu and Tengely concerning arithmetic progressions containing the most $\ell$-th powers asymptotically, and a new assertion answering a question of Hajdu and Tengely.

# Publications of Ágoston Papp

1. L. Hajdu, Á. Papp, *On asymptotic density properties of the sequence* $(n!)_{n=0}^{\infty}$, Acta Arith, **184** (2018), 317–340.

2. L. Hajdu, Á. Papp, T. Szakács, *On the equation* $A!B! = C!$, J. Number Theory **187** (2018), 160–165.

3. L. Hajdu, Á. Papp, *Polynomial values of products of terms from an arithmetic progression*, Monatshefte für Mathematik **193 (3)** (2020), 637–655.

4. L. Hajdu, Á. Papp, R. Tijdeman, *The Prouhet-Tarry-Escott Problem, Indecomposability of Polynomials and Diophantine Equations*, The Ramanujan J. **58** (2022), 1075–1093.

5. L. Hajdu, Á. Papp, *Uniform bounds for the number of powers in arithmetic progressions*, Rev. Real Acad. Cienc. Exactas Fis. Nat. Ser. A-Mat. **116**, 169 (2022)

# Conference talks of Ágoston Papp

1. 24th Central European Number Theory Conference, *On asymptotic density properties of the sequence* $(n!)_{n=0}^{\infty}$ (2019)

2. Number Theory Conference in honour of professors Kálmán Győry, János Pintz and András Sárközy, *Uniform bounds for the number of powers in arithmetic progressions* (2022)