TWENTY-SIX RESULTS IN NUMBER THEORY AND DISCRETE MATHEMATICS

ANMOL SAHU

Abstract. We present proofs of twenty-six results drawn from problems about integer sequences, divisors, primes, and finite sets. Two divisor-sum questions have a common counterexample: the ratio of the nonsquarefree divisor sum to the squarefree divisor sum can equal 3107^2, even when both sums are squares. A construction with digit blocks gives infinitely many examples in every base whose repetitions, followed by a final digit 1, are always composite. Other sections derive recurrences for binary cellular automata and turn a sieve estimate into an explicit family of sparse rows of Gaussian primes. In additive number theory we give a finite proof of the known sharp threshold between 69 and 70 for sums with at most five ordered representations. Nine of the twenty-six results establish individual steps of larger problems. We identify exactly what each step proves, give references for earlier arguments, and link the corresponding Lean proofs.

Working draft · 28 November 2026

Contents

1. Introduction 1
2. Five questions about integer sequences 2
3. Reading sequences from cellular automata 5
4. What a prime factorization tells us 7
5. Powers modulo primes and polynomial factors 9
6. A sieve bound for rows of Gaussian primes 11
7. Representations by two summands 12
8. Products of consecutive integers 14
9. Forcing an arithmetic progression 15
10. What the proofs leave open 16
Appendix A. Problem sources and formal proofs 17
References 17

1. Introduction

What can a small example tell us about an infinite collection of integers? For one question, the answer is everything. Suppose that every square ratio in a family of divisor sums is conjectured to be the square of a prime. A single ratio equal to 3107^2 settles the question, because $3107 = 13 \cdot 239$. A related question imposes two extra square conditions. Multiplying the original integer by 5 meets both conditions and leaves the ratio unchanged. We prove the two assertions together in Sections 4.2 and 4.3.

For another question, a finite calculation leaves a genuine gap. Suppose that every sufficiently large integer is a sum of two members of a set A. The Erdős–Turán conjecture asks whether some integers must have arbitrarily many such representations. We explain why covering the interval from 0 through 70 forces a sum with at least six ordered representations, and why 69 is the last endpoint

for which five still suffice [1]. The argument establishes one precise threshold. The conjecture asks for a conclusion at every threshold.

These two examples describe the range of this paper. Of the twenty-six formal submissions collected here, seventeen settle the stated problem in the source collection and nine prove an intermediate assertion. Several of the results have published mathematical proofs. For those results our contribution is a machine-checked formulation or proof, and we identify its antecedents where we use them. The all-bases digit-block construction is a companion extension to the decimal result; it is presented separately from the count of twenty-six submissions.

How to read the paper. The sections are largely independent. For a first example in linear algebra, start with the four-by-four determinant in Section 2.1. For a short counting argument, try Section 7.1: it takes only two ordered pairs to explain the effect of adding one element to a set. Readers interested in longer proofs can turn to the sieve estimate in Section 6.1 or the binary encoding in Section 7.3. We introduce the specialized notation at the point where it is needed.

Throughout, a prime is a positive integer greater than one, and divisors of a positive integer are positive unless stated otherwise. The notation $a \mid b$ means that b is an integer multiple of a; the congruence $a \equiv b \pmod{m}$ means that $m \mid (a - b)$. For a finite set S, $|S|$ denotes its number of elements. Labels beginning with A followed by six digits refer to entries in the *On-Line Encyclopedia of Integer Sequences* (OEIS). Appendix A matches each result to its original question and its formal proof.

2. Five questions about integer sequences

What would count as a proof that a sequence has a given pattern? Finding a zero in a determinant sequence takes one linear relation. Ruling out a prime in an infinite list takes an argument that works for every term. Between these examples lie two useful techniques: taking consecutive differences, and separating the possible remainders of an index.

2.1. **A zero in the prime Hankel determinants.** Number the primes starting with $p_0 = 2$. A *Hankel matrix* has constant entries along each diagonal running from the lower left to the upper right. In Jeffrey Shallit's sequence the n by n matrix is $H_n = (p_{i+j})_{0 \le i,j < n}$, with the empty determinant set to 1. Jon Perry asked whether 4 is the only index with zero determinant [2].

Theorem 2.1. *The determinant of H_4 is zero.*

Proof. All that is needed is a nonzero vector that the matrix sends to zero. The first seven primes give

$$\begin{pmatrix} 2 & 3 & 5 & 7 \\ 3 & 5 & 7 & 11 \\ 5 & 7 & 11 & 13 \\ 7 & 11 & 13 & 17 \end{pmatrix} \begin{pmatrix} 6 \\ -3 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

For example, the first row gives $12 - 9 - 10 + 7 = 0$. The vector exhibits the fourth column as -6 times the first plus 3 times the second plus 2 times the third. The columns are dependent, so the determinant vanishes. □

The submitted step uses this null-vector calculation from the existing formal library [3, 4]. The uniqueness assertion would also require proving that every other determinant is nonzero.

2.2. **Five rows of Pascal's triangle.** Write $s_2(m)$ for the number of ones in the binary expansion of m. For example, $s_2(21) = 3$ because $21 = (10101)_2$. In row n of Pascal's triangle, let $a(n)$ be the least j with $1 \le j \le n$ such that $s_2(\binom{n}{j})$ is even. Put $a(n) = 0$ if there is no such j. Vladimir Shevelev proposed that exactly five rows have this value [5].

Theorem 2.2. *For $n \in \{0, 1, 2, 7, 8\}$, every entry of row n has an odd number of binary ones. In particular, $a(n) = 0$ for each of these five indices.*

Proof. Row symmetry gives $\binom{n}{j} = \binom{n}{n-j}$, so we need only read from the first entry to the middle of each row:

n	$\binom{n}{0}, \ldots, \binom{n}{\lfloor n/2 \rfloor}$
$0, 1$	1
2	$1, 2$
7	$1, 7, 21, 35$
8	$1, 8, 28, 56, 70$

The powers of two on the list have one binary one. The six remaining numbers, in the order $7, 21, 35, 28, 56, 70$, have binary expansions 111, 10101, 100011, 11100, 111000, 1000110. Each has three ones. For row zero the defining search is empty as well. □

These computations establish the five examples in the proposed classification [6, 7]. The converse asks for an entry with an even number of binary ones in every other row.

2.3. **Repeated partial sums of the Catalan numbers.** The Catalan numbers $C_k = \binom{2k}{k}/(k+1)$ start with $1, 1, 2, 5$ [8]. Their partial sums are $1, 2, 4, 9, \ldots$. Paul Barry's sequence writes each sum twice:

$$a_n = \sum_{j=0}^{\lfloor n/2 \rfloor} C_j, \qquad (a_0, \ldots, a_7) = (1, 1, 2, 2, 4, 4, 9, 9).$$

R. J. Mathar proposed the following recurrence [9].

Theorem 2.3. *For every $n \geq 3$,*

$$(n+2)(a_n - a_{n-1}) = 4(n-1)(a_{n-2} - a_{n-3}).$$

Proof. Write $D_n = a_n - a_{n-1}$ for $n \geq 1$. At odd n, the partial sum has just been repeated, so $D_n = 0$. At $n = 2k \geq 2$ exactly one term has been added, giving $D_{2k} = C_k$. Both sides of the recurrence thus vanish when n is odd. When $n = 2k$ they reduce, after division by two, to

$$(k+1)C_k = 2(2k-1)C_{k-1}.$$

This last identity follows by cancelling factorials: $C_k/C_{k-1} = (2k)(2k-1)/(k(k+1))$. □

2.4. **The real part of a Möbius sequence.** For a positive integer n, the Möbius function $\mu(n)$ is zero if n is divisible by the square of a prime. Otherwise $\mu(n) = (-1)^k$, where k counts its distinct prime factors [10]. In particular $\mu(1) = 1$. Antti Karttunen's sequence has $a_n = 0$ for odd n and $a_n = -\mu(n)$ for even n. Mats Granvik proposed the following single expression for both cases [11]; an earlier project submission also gave the identity [12].

Theorem 2.4. *Let $i^2 = -1$ and write $\mathrm{Re}(z)$ for the real part of a complex number z. For every $n \geq 1$,*

$$a_n = \mathrm{Re}(\mu(n) i^n).$$

Proof. The powers of i repeat $1, i, -1, -i$ starting at exponent zero. At an odd index the real part is zero. If $n \equiv 2 \pmod 4$, then $i^n = -1$, giving $-\mu(n)$ as required. The only other indices are multiples of four. Here $i^n = 1$, but $2^2 \mid n$ forces $\mu(n) = 0$, so the sign makes no difference. □

2.5. **A decimal block that always gives a composite.** Write r copies of the digits of a positive integer m and append a final 1. For $m = 104040$ the first two candidates are 1040401 and 1040401040401. Farideh Firoozbakht conjectured that some repetition would give a prime whenever m is not a power of ten [13]. The user sweety439 had already disproved that assertion using $m = 1010$ [14]. Maxence Faldor's earlier formalization gives the witness and factorization below [15].

Theorem 2.5. *For every $r \geq 1$, writing r copies of* 104040 *and appending the digit* 1 *gives a composite number.*

Proof. Call the resulting number N_r. Each block shifts the preceding digits six places, so

$$N_r = 1 + 1040400 \sum_{j=0}^{r-1} 10^{6j}.$$

Multiplying by $10^6 - 1$ clears the geometric sum and exposes a difference of squares:

$$\begin{aligned} 999999 N_r &= 1040400 \cdot 10^{6r} - 40401 \\ &= (1020 \cdot 1000^r - 201)(1020 \cdot 1000^r + 201). \end{aligned}$$

Both factors exceed 999999: even the smaller one at $r = 1$ is 1019799. Each factor is also less than N_r, since it equals $999999 N_r$ divided by the other factor. If N_r were prime, it would divide one of these two positive factors smaller than itself. This is impossible. □

A companion construction in every base. The decimal block 104040 is one instance of a construction that works in every base. To see what changes with the base, let $C_{b,r}(n)$ be the integer obtained by repeating the base-b digits of n exactly r times and appending one digit 1. For example, $C_{10,2}(12) = 12121$. A word of digits is *primitive* if it cannot be formed by writing a shorter word two or more times. Thus 1212 is a repetition of 12, whereas 121 is primitive.

Theorem 2.6 (Primitive blocks)**.** *Let $b \geq 2$, $k \geq 1$, and $x = b^k \geq 3$. Set $n = (x^3 + 2x)^2/b$. The base-b expansion of n is a primitive block of $6k$ digits. It begins with* 1*, has another nonzero digit, and $C_{b,r}(n)$ is composite for every $r \geq 1$. In particular, every base admits infinitely many such primitive blocks.*

Proof. The quotient defining n is integral since $b \mid x$. Moreover,

$$\frac{n}{b^{6k-1}} = 1 + \frac{4}{x^2} + \frac{4}{x^4}.$$

This lies strictly between 1 and 2, since $x \geq 3$. It gives the length of the block, its leading digit, and the presence of another nonzero digit.

Put $A = x^3 + 2x$, $D = 2x^2 + 1$, and $M = x^6 - 1$. Expanding the squares gives $A^2 - D^2 = M$. One shift by an entire block multiplies by $b^{6k} = x^6$. Summing the geometric progression for the r blocks therefore gives

$$M C_{b,r}(n) = A^2(x^{6r} - 1) + M = (Ax^{3r} - D)(Ax^{3r} + D).$$

The smaller factor already exceeds M at $r = 1$, by $2x^2(x^2 - 1)$, and it grows with r. Both factors are consequently positive and less than $C_{b,r}(n)$: dividing their product by either one gives a quotient greater than M. If $C_{b,r}(n)$ were prime, it would divide one of the factors. Both are too small.

It remains to rule out a shorter repeating word. Suppose that $s \geq 2$ identical copies make the block, each of length $d = 6k/s$. Since $n = b^{2k-1}(x^2 + 2)^2$, the final $2k - 1$ digits are zero. If $s \geq 4$, then $d \leq 3k/2 < 2k$. The last copy would consist entirely of zeros, although the first copy begins with 1.

If $s = 2$, then $x^3 + 1$ divides n, since two identical blocks of length $3k$ have value $u(x^3 + 1)$. But

$$bn = (x^3 + 2x)^2 \equiv (2x - 1)^2 \pmod{x^3 + 1}.$$

The remainder is positive and smaller than x^3+1, as their difference is $x(x-2)^2>0$. Finally, if $s=3$, then x^4+x^2+1 divides n. Modulo that number $x^6\equiv 1$ and $x^4\equiv -x^2-1$, so $bn=x^6+4x^4+4x^2\equiv -3$, again impossible. Increasing k gives different block lengths, proving infinitude. □

For $b=2,k=2$ this gives $n=2592$, whose binary expansion is 101000100000. The choice $b=10,k=1$ recovers 104040. Earlier decimal witnesses and factorizations in other bases are recorded in [14, 15]. The companion Lean file verifies the length, primitivity and compositeness assertions in Theorem 2.6.

3. Reading sequences from cellular automata

Imagine one lamp at each integer position. Initially only the lamp at 0 is on. In each round all lamps update simultaneously. Each new bit is determined by three old bits, in the order left neighbor, center, right neighbor. A rule specifies which of the eight possible triples give output 1. We examine three rules and three ways to read their rows: count the on or off bits, read the central bit over time, or read a whole window as a binary number.

For a sequence $(u_n)_{n\ge 0}$ its *ordinary generating function* is the formal power series $U(x)=\sum_{n\ge 0}u_nx^n$. Multiplying by x shifts the coefficients by one place. For example, the coefficient of x^n in $(1-x^2)U(x)$ is u_n-u_{n-2} for $n\ge 2$. We use these series only to record coefficient identities.

3.1. On cells under Rule 1.

Rule 1 gives output 1 only on the triple 000. At time n count the lit lamps in the window $-n,-n+1,\dots,n$, and call their number u_n. This finite window is important: infinitely many lamps are lit at time one. Robert Price recorded the sequence, and Colin Barker proposed closed formulas, a recurrence and a generating function [16].

Theorem 3.1. *One has $u_n=1$ at every even time, and $u_n=2n-2$ at every odd time. For $n\ge 4$,*

$$u_n=2u_{n-2}-u_{n-4},\qquad \sum_{j\ge 0}u_jx^j=\frac{1-x^2+4x^3}{(1-x^2)^2}.$$

Proof. At time one the only off cells are $-1,0,1$, precisely the three centers whose neighborhoods included the initial on cell. These three off cells now form the unique 000 triple, so after one more round only cell 0 is on. This is the initial row, hence the two configurations repeat. At odd time n the window contains all three off cells. It therefore contains $(2n+1)-3=2n-2$ on cells.

Along each parity class u_n is linear in n, so consecutive differences two steps apart are equal. This gives the recurrence. Multiplying the generating function by $(1-x^2)^2=1-2x^2+x^4$ kills every coefficient from degree four onward. The values $u_0,u_1,u_2,u_3=1,0,1,4$ leave the numerator $1-x^2+4x^3$. □

3.2. Off cells in the same window.

Let v_n count off cells in the same window and under the same rule. Price's companion sequence and Barker's formulas are recorded in [17]. Each cell contributes to exactly one count, so $u_n+v_n=2n+1$.

Theorem 3.2. *One has $v_n=2n$ when n is even, and $v_n=3$ when n is odd. For $n\ge 4$,*

$$v_n=2v_{n-2}-v_{n-4},\qquad \sum_{j\ge 0}v_jx^j=\frac{3x+4x^2-3x^3}{(1-x^2)^2}.$$

Proof. Subtract the two parity formulas for u_n from $2n+1$. Again the count is linear along each parity class, giving the recurrence. Multiplication by $(1-x^2)^2$ now gives initial coefficients $0,3,4,3-2\cdot 3=-3$ and all later coefficients zero. □

3.3. **Rule 158 and the central digit.** Rule 158 gives output 1 exactly on the triples $001, 010, 011, 100, 111$. Let d_n be the bit at cell 0 at time n. Write the bits $d_0, \dots, d_n$ as decimal digits of a_n. The first numbers are $1, 11, 111, 1110$. Michael De Vlieger recorded a period-four description for Robert Price's sequence, and Colin Barker proposed the recurrence and generating function below [18, 19].

Theorem 3.3. *The central bits repeat* 1110 *from time zero. The resulting decimal integers satisfy*

$$a_n = 10a_{n-1} + a_{n-4} - 10a_{n-5} \quad (n \geq 5), \qquad \sum_{n\geq 0} a_n x^n = \frac{1 + x + x^2}{(1 - x^4)(1 - 10x)}.$$

Proof. Starting from the single 1, the first three rows in the growing windows are 1, 111, 11101. In general the windows at times $2k$ and $2k + 1$ have respective words

$$1(1101)^k, \qquad 111(0011)^k,$$

where a power of a word means repetition. All cells outside the current window are off, since 000 maps to 0.

Here are details of the induction on these words. To read a row, start with two remembered zeros, read one bit at a time, and output the rule value of the remembered pair followed by that bit. Retain the last two bits for the next triple. Read two extra zeros at the end to produce the new, wider window. Each table entry records all outputs while reading the indicated input block:

remembered pair	input	output	new pair
00	1	1	01
01	1101	1100	01
01	00	11	00
00	111	111	11
11	0011	0111	11
11	00	01	00

For instance, the second line scans $011, 111, 110, 101$. The first three lines send $1(1101)^k$ to $1(1100)^k 11 = 111(0011)^k$. The last three lines send that word to $111(0111)^k 01 = 1(1101)^{k+1}$. This proves both row descriptions.

The center of the time-n word has index n counting from zero on the left. In the even word every zero has odd index, so its center is on. In the odd word the zeros have indices $3, 4, 7, 8, \dots$ within its length. Its odd center index therefore gives zero exactly when $n \equiv 3 \pmod 4$. This proves the period.

Appending the new digit gives $a_n - 10a_{n-1} = d_n$ for $n \geq 1$. The same equality four places earlier has the same right side, giving the recurrence. Finally the digit series is $(1 + x + x^2)/(1 - x^4)$, and $a_n = \sum_{j=0}^{n} 10^{n-j} d_j$. The latter sum multiplies the digit series by $1/(1 - 10x)$. □

3.4. **Rule 157 and the value of a whole row.** Rule 157 gives output 1 on $000, 010, 011, 100, 111$. This time read the window $[-n, n]$ at time n as a binary number w_n, with its leftmost bit the most significant. The first windows $1, 011, 00111, 0101111$ give $w_0, w_1, w_2, w_3 = 1, 3, 7, 47$. Robert Price contributed the sequence and Colin Barker proposed its two formulas [20]. The formal statement and an earlier submission are recorded in [21, 22].

Theorem 3.4. *For* $n \geq 5$,

$$w_n = 3w_{n-1} + 14w_{n-2} - 48w_{n-3} + 32w_{n-4}.$$

Its generating function is

$$\sum_{n\geq 0} w_n x^n = \frac{1 - 16x^2 + 32x^3 - 32x^4}{(1 - x)(1 - 2x)(1 - 4x)(1 + 4x)}.$$

Proof. At every time $n \geq 1$ the entire row has the following description. All positions $j < -n$ and $j \geq 0$ are on. The left endpoint $-n$ is off. In between, for $-n < j < 0$, exactly the even positions are on.

To verify this description, the first update gives it directly. Suppose it holds at time n. The new left endpoint at $-n-1$ sees 110 and turns off. The old endpoint sees 100 when n is even and 101 when n is odd, so its new value agrees with the parity rule for interior cells. An even, strictly negative interior cell sees 010 or 011 and stays on. An odd such cell sees 001 or 101 and stays off. At 0 and to its right the triples end in 11; both 011 and 111 give 1. Farther left than the new endpoint the triple is 111. This verifies all positions.

When the window widens, its new left bit is 0 and its new right bit is 1. The old bits double their weights, and only the old left endpoint can change. It turns on just when n is even, with new weight 2^{2n+1}. Hence for $n \geq 1$,

$$w_{n+1} = 2w_n + 1 + E_n, \qquad E_n = \begin{cases} 2^{2n+1}, & n \text{ even}, \\ 0, & n \text{ odd}. \end{cases}$$

Since $E_{n+2} = 16E_n$, subtracting sixteen times this update from its counterpart at $n+2$ gives

$$w_{n+3} - 2w_{n+2} - 16w_{n+1} + 32w_n = -15.$$

The next such equality has the same right-hand side. Subtracting the two equalities gives the stated fourth-order recurrence. Multiplying the series for w_n by the denominator in the theorem applies precisely that recurrence from degree five onward. The initial terms $1, 3, 7, 47, 95$ give the stated numerator. □

4. What a prime factorization tells us

If $n = p_1^{e_1} \cdots p_r^{e_r}$, each divisor of n is determined by choosing one exponent between 0 and e_i for each p_i. For example, the divisors of $12 = 2^2 \cdot 3$ come from multiplying one of $1, 2, 4$ by one of $1, 3$. We use this observation first to count divisors, then to sum them. Later in the section we use the exponents themselves to define a derivative on the positive integers.

4.1. **Counting prime and composite divisors.**

The positive divisors of 9 are $1, 3, 9$. Exactly one is prime and one is composite. Can the same balance occur in a number with several prime factors? The following classification appears in a 2019 argument of Charlie Neder for Felix Fröhlich's sequence [23].

Theorem 4.1. *A positive integer has equally many prime and composite positive divisors if and only if it is* 1 *or the square of a prime.*

Proof. Suppose $n > 1$ has r distinct prime factors. These are precisely its prime divisors. If it has r composite divisors too, counting 1 gives

$$(e_1 + 1) \cdots (e_r + 1) = 2r + 1.$$

The product is odd, so each factor is odd. Since every $e_i \geq 1$, each factor is at least 3. This forces $3^r \leq 2r + 1$. For $r = 2$ we already have $9 > 5$, and multiplying the left side by 3 at each subsequent step preserves the strict reverse inequality. Thus $r = 1$, and its single exponent satisfies $e_1 + 1 = 3$. Conversely, the divisors of p^2 are $1, p, p^2$, and 1 itself has zero divisors of either counted kind. □

4.2. **A square ratio of divisor sums.** A positive integer is *squarefree* if no prime square divides it. Write $S(n)$ for the sum of the squarefree divisors of n, $T(n)$ for the sum of its other divisors, and $\sigma(n) = S(n) + T(n)$. In the example $n = 12$, the two groups are $1, 2, 3, 6$ and $4, 12$, giving $S(12) = 12$ and $T(12) = 16$. Michel Lagneau asked whether an integer square $T(n)/S(n) = k^2$, with $k > 0$, must have a prime square root [24].

Theorem 4.2. *For $N = 2^7 \cdot 337^3$, we have $S(N) = 1014$ and $T(N)/S(N) = 3107^2$, where $3107 = 13 \cdot 239$.*

Proof. Trial division by the primes up to 17 shows that 337 is prime. For a squarefree divisor we may take each of 2 and 337 either once or zero times. For an arbitrary divisor we may take any permitted power. Consequently

$$\begin{aligned} S(N) &= (1+2)(1+337) = 1014, \\ \sigma(N) &= (1+2+\cdots+2^7)(1+337+337^2+337^3) \\ &= 255(1+337)(1+337^2). \end{aligned}$$

Divide the second formula by the first and subtract 1. This gives $T(N)/S(N) = 85(1+337^2) - 1 = 9\,653\,449 = 3107^2$. The displayed factorization of 3107 completes the counterexample. □

4.3. **Making both divisor sums squares.** Lagneau's second question imposes a stronger hypothesis: both $S(n)$ and $T(n)$ must themselves be squares, and $T(n)/S(n)$ must be an integer [25]. The same ratio can survive the extra conditions.

Theorem 4.3. *For $n = 5N$, where N is as in Theorem 4.2,*

$$S(n) = 78^2, \qquad T(n) = 242346^2, \qquad T(n)/S(n) = 3107^2.$$

Proof. For any prime q not dividing N, the divisors of qN come in pairs d, qd, one pair for every $d \mid N$. Both members are squarefree exactly when d is squarefree. Therefore multiplying N by q multiplies each sum by $q + 1$ and preserves their ratio. To see how to choose q, write $S(N) = 1014 = 6 \cdot 13^2$. Multiplication of this sum by 6 makes it a square, so take $q = 5$. Now $S(5N) = (6 \cdot 13)^2 = 78^2$ and $T(5N) = (78 \cdot 3107)^2$. As before, its ratio to $S(5N)$ has a composite square root. □

4.4. **A denominator and the arithmetic derivative.** For a prime $p \mid n$, let e_p be its exponent in n. Define the *arithmetic derivative* by

$$D(1) = 0, \qquad D(n) = \sum_{p \mid n} e_p \frac{n}{p},$$

where here and below a sum indexed by $p \mid n$ runs over the distinct prime divisors. For instance $D(12) = 2 \cdot 6 + 4 = 16$. An empirical observation about a denominator in Jonathan Sondow's sequence led Antti Karttunen to the following formula [26, 27].

Theorem 4.4. *For $n \geq 1$, let d be the positive denominator in lowest terms of $\sum_{p \mid n} 1/p - 1/n$. Then*

$$\frac{n}{d} = \gcd\bigl(n, 1 + (n-1)D(n)\bigr).$$

Proof. Put $U = \sum_{p \mid n} n/p$. The fraction equals $(U-1)/n$, so n/d is the cancellation factor $\gcd(n, U-1)$. Modulo n, the other quantity inside the claimed gcd is $1 - D(n)$. It suffices to compare $\gcd(n, U-1)$ and $\gcd(n, D(n)-1)$ prime by prime.

If $p^2 \mid n$, then p divides every term n/q of U, including the term with $q = p$. Hence both U and $D(n)$ are divisible by p. Subtracting 1 excludes p from both gcds. If p occurs only once in n, then n/p appears with coefficient 1 in both sums. Every other term is divisible by p, so $U \equiv D(n)$

(mod p). Both gcds therefore contain p or both omit it, and neither can contain p^2. This exhausts the primes in n. At $n = 1$ the fraction is -1 and both sides equal 1. □

For $n = 12$ the fraction is $1/2 + 1/3 - 1/12 = 3/4$, so the cancellation factor is $12/4 = 3$. The derivative formula gives $\gcd(12, 1 + 11 \cdot 16) = \gcd(12, 177) = 3$.

4.5. **Differentiating a hyperfactorial.** The *hyperfactorial* is the product $H(n) = \prod_{k=1}^{n} k^k$, with $H(0) = 1$. Matthew Campbell's sequence records its arithmetic derivative [28, 29]. The next identity permits the calculation using only derivatives of integers up to n.

Theorem 4.5. *For every nonnegative integer n,*

$$D(H(n)) = H(n) \sum_{k=1}^{n} D(k).$$

Proof. Dividing the definition of $D(m)$ by m gives $L(m) = D(m)/m = \sum_{p|m} e_p/p$. For a product ab, the exponent of each prime is the sum of its exponents in a and b. Thus $L(ab) = L(a) + L(b)$. Applying this to k factors equal to k gives $L(k^k) = kL(k) = D(k)$. Applying it once more to the factors of $H(n)$ and multiplying by $H(n)$ proves the identity. At $n = 0$ the sum is empty and $D(1) = 0$. □

4.6. **Doubling a counterexample to a totient inequality.** Euler's totient $\varphi(n)$ counts integers from 1 to n coprime to n. Dedekind's function $\psi(n) = n \prod_{p|n}(1 + 1/p)$ also depends on the primes in n. In 1988 József Sándor asked whether $\varphi(\psi(n)) \le n$ for all $n \ge 2$. Sándor credits László Tóth for the counterexample 39270, and records its first double among the examples found by Kovács Lehel István [30]. The same calculation gives the exact value along every doubling.

Theorem 4.6. *For every integer $k \ge 0$, $\varphi(\psi(39270 \cdot 2^k)) = 41472 \cdot 2^k > 39270 \cdot 2^k$.*

Proof. Write $n_0 = 39270 = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 17$. Since n_0 is squarefree, its psi value is the product of the numbers $p + 1$ for these six primes. Hence

$$\psi(n_0) = 3 \cdot 4 \cdot 6 \cdot 8 \cdot 12 \cdot 18 = 124416 = 2^9 \cdot 3^5.$$

Multiplying n_0 by 2^k leaves its set of prime factors unchanged, so $\psi(n_0 \cdot 2^k) = 124416 \cdot 2^k$. The latter integer has only the prime factors 2 and 3. In every block of six consecutive integers, precisely two, those congruent to 1 or 5, are coprime to it. Thus its totient is one third of its value, as asserted. □

5. Powers modulo primes and polynomial factors

For integers a, b and $d \ge 2$, the notation $a \equiv b \pmod{d}$ means that d divides $a - b$. The point of working with remainders is that a short cycle can control infinitely many exponents. The first two results use this to force a common divisor. The last two use it to describe the roots of a polynomial.

5.1. **Six primes cover all exponents.** Hans Riesel's 1956 example is an odd number k for which $k2^m - 1$ is composite for every $m \ge 1$ [31]. The associated sequence asks whether 509203 is the *least* such odd k [32]. The assertion proved here is that it has the required property.

Theorem 5.1. *The integer $509203 \cdot 2^m - 1$ is composite for every integer $m \ge 1$.*

Proof. Let $k = 509203$. The table gives a divisor of $k2^m - 1$ according to the residue class of m. In a row with modulus t and representative a, the last two entries show that $k2^a \equiv 1$ and $2^t \equiv 1$ modulo the prime in that row.

p	$m \equiv a \pmod t$	$k \bmod p$	$2^a \bmod p$	$2^t \bmod p$
3	$0 \pmod 2$	1	1	1
5	$1 \pmod 4$	3	2	1
7	$2 \pmod 3$	2	4	1
13	$7 \pmod{12}$	6	11	1
17	$7 \pmod 8$	2	9	1
241	$3 \pmod{24}$	211	8	1

For example, $211 \cdot 8 = 7 \cdot 241 + 1$. To check coverage, consider the residues modulo 24, a multiple of all six periods. The first row covers all even residues. The second covers the six odd residues $1, 5, 9, 13, 17, 21$. Of the remaining six, $11, 23$ are covered by 7, $7, 19$ by 13, 15 by 17, and 3 by 241. In every case the divisor is at most 241, whereas $k2^m - 1 \geq 1\,018\,405$. □

5.2. **A progression of exceptions for two shifted powers.** For $n \geq 1$, let $a(n)$ be the smallest positive m for which $2^n - m$ and $3^n - m$ are coprime. Clark Kimberling proposed that $a(n) = 3$ holds exactly when $n = 330k - 209$ for some integer $k \geq 1$ [33]. For the following exponents, $m = 3$ fails even the coprimality test.

Theorem 5.2. *For each integer $t \geq 0$, the exponent $n = 3438721 + 5366790t$ belongs to Kimberling's progression, and* 487891 *divides both $2^n - 3$ and $3^n - 3$.*

Proof. Set $u = 3438721$, $h = 5366790$, and $d = 487891$. The modular calculations needed for all t are

b	$b^u \bmod d$	$b^h \bmod d$
2	3	1
3	3	1

They can be reproduced by repeated squaring: begin with $r = 1$ and read the exponent in binary from left to right. At a digit $e \in \{0, 1\}$, replace r by the remainder of $r^2 b^e$ modulo d. The new r then represents the power specified by the digits read so far. All four calculations are also certified in the formal proof [34].

For $b = 2, 3$ it follows that $b^{u+ht} = b^u (b^h)^t \equiv 3 \pmod d$. Thus $d > 1$ is a common divisor of the two shifted powers. Finally,

$$u + ht = 330(10421 + 16263t) - 209.$$

The positive step h gives infinitely many distinct exceptions. □

5.3. **When a polynomial has ten distinct roots.** Let $\mathbb{F}_p$ denote the field of residues modulo a prime p. Federico Provvedi proposed a polynomial description of the primes $p \equiv 1 \pmod{11}$ in N. J. A. Sloane's sequence [35]. For example, a nonzero residue z has *order* k if k is the least positive integer with $z^k = 1$. Its k powers form a subgroup. The nonzero residues partition into translates of this subgroup, each containing k elements, so k divides $p - 1$. The quotient in the following theorem is polynomial division, so its value at $X = 0$ is well defined.

Theorem 5.3. *The polynomial $Q(X) = ((X + 1)^{11} - 1)/X$ has ten distinct roots in $\mathbb{F}_p$ if and only if $p \equiv 1 \pmod{11}$.*

Proof. The geometric sum gives $Q(X) = \sum_{j=0}^{10} (X + 1)^j$, so $Q(0) = 11$. If $p \neq 11$, adding 1 gives a bijection between the roots of Q and the solutions $y \neq 1$ of $y^{11} = 1$. Each such y has multiplicative

order 11: its order divides 11 and exceeds 1. The order of an element divides the size $p-1$ of the group of nonzero residues. Thus even one such root forces $11 \mid p-1$.

Conversely, if $11 \mid p-1$, then $Y^{11}-1$ divides $Y^{p-1}-1$. By Fermat's little theorem the latter polynomial already has the $p-1$ distinct nonzero residues as roots. Its degree is $p-1$, so it is a product of distinct linear factors. Its divisor $Y^{11}-1$ therefore has eleven distinct roots, ten of which differ from 1. For $p=11$, the binomial theorem gives $Q(X)=X^{10}$, which has only one distinct root. □

5.4. **When a quartic splits into two quadratics.** A polynomial over a field is *irreducible* if it cannot be written as a product of two polynomials of positive degree over that field. It is *monic* if its leading coefficient is 1. Provvedi proposed the following degree-four counterpart, and Jianing Song recorded the equivalent formulation before the shift $X \mapsto X+1$ [36].

Theorem 5.4. *The polynomial $F(X)=((X+1)^5-1)/X$ over $\mathbb{F}_p$ is the product of two distinct monic irreducible quadratics if and only if $p \equiv 9 \pmod{10}$.*

Proof. Set $G(Y)=1+Y+Y^2+Y^3+Y^4$. Then $F(X)=G(X+1)$, and translation preserves degrees and factorizations. At $p=5$ one has $F(X)=X^4$, so suppose $p \neq 5$. In a field containing all its roots, G has four distinct roots of multiplicative order 5: adjoining 1 gives all five roots of Y^5-1. These roots are distinct because the derivative $5Y^4$ is nonzero at each one.

Suppose G has an irreducible quadratic factor. Adjoining a root z of that factor gives a field with p^2 elements: each element is uniquely $a+bz$ with $a,b \in \mathbb{F}_p$. The order of z divides p^2-1, so $p^2 \equiv 1 \pmod 5$, or $p \equiv \pm 1 \pmod 5$. If $p \equiv 1 \pmod 5$, then Y^5-1 divides $Y^{p-1}-1$. The latter already splits into linear factors over $\mathbb{F}_p$, as in the preceding proof. This leaves $p \equiv -1 \pmod 5$.

Conversely, for $p \equiv -1 \pmod 5$ and any root z of G, we have $z^p = z^{-1}$. The two roots z, z^{-1} are distinct. The map $x \mapsto x^p$ interchanges them, so it fixes $z+z^{-1}$. The only elements it fixes are the p roots of X^p-X, namely the elements of $\mathbb{F}_p$. Hence

$$(Y-z)(Y-z^{-1}) = Y^2-(z+z^{-1})Y+1$$

has coefficients in $\mathbb{F}_p$. Neither root belongs to $\mathbb{F}_p$, since 5 does not divide $p-1$. This gives an irreducible quadratic for each of the two inverse pairs, accounting for all four roots. Finally, a prime congruent to 4 modulo 5 is odd and therefore ends in 9. □

For a concrete instance take $p=19$. In $\mathbb{F}_{19}[Y]$,

$$G(Y) = (Y^2-4Y+1)(Y^2+5Y+1).$$

Multiplying recovers G because $2-4\cdot 5 \equiv 1 \pmod{19}$. Neither quadratic has a root in $\mathbb{F}_{19}$, whose nonzero elements form a group of order 18.

6. A sieve bound for rows of Gaussian primes

6.1. **An explicit family of sparse rows.** A Gaussian integer has the form $n+si$ with integers n, s and $i^2=-1$. Its norm is n^2+s^2. Away from the coordinate axes, such an integer is a Gaussian prime precisely when its norm is an ordinary prime. Fixing n and varying s therefore turns a question about lattice points into one about the numbers $n^2+1, n^2+4, \dots, n^2+(n-1)^2$.

Write $a(n)$ for the number of primes in this list and $\pi(n)$ for the number of primes at most n. For example, when $n=5$ the list is $26, 29, 34, 41$, giving $a(5)=2$. Noe proposed, in particular, that $5a(n) \geq \pi(n)$ for every $n>1$ [37]. Earlier work already disproved the inequality [38, 39, 40, 41]. The following result gives an explicit choice of rows where it fails.

Theorem 6.1. *For any integer $Y \geq 2^{400000001}$, let S be the set of primes $p \leq Y$ with $p \equiv 3 \pmod 4$ and let $M=\prod_{p\in S} p$. Then $5a(M^{10}) < \pi(M^{10})$.*

The large cutoff serves a specific purpose: it gives a proof that the row works without requiring us to enumerate its primes. The input is the finite sieve estimate of [38]. For a finite set S of primes congruent to 3 modulo 4, put $P = \prod_{p\in S} p/(p-1)$. That estimate states that $M \geq 200000$ and $P > 10000$ suffice. We prove both numerical conditions for the stated S.

Proof. Write $H(Y) = \sum_{1\leq n\leq Y,\ n \text{ odd}} 1/n$. We first show that $H(Y) \leq P^2$. Define $\chi(n)$ to be 0 on even numbers, 1 on numbers congruent to 1 modulo 4, and -1 on numbers congruent to 3 modulo 4. Define $w(1) = 1$, assign to each positive power of a prime p weight 2 when $p \equiv 3 \pmod 4$ and weight 0 otherwise, and multiply the weights belonging to distinct prime factors. We claim that

$$\sum_{d|n} w(d)\chi(n/d) = \begin{cases} 1, & n \text{ odd}, \\ 0, & n \text{ even}. \end{cases}$$

Both sides multiply across relatively prime inputs, so it suffices to check prime powers. At a positive power of 2 the sum is zero. At p^e with $p \equiv 1 \pmod 4$ only $d = 1$ contributes. At p^e with $p \equiv 3 \pmod 4$ the sum is $(-1)^e + 2\sum_{j=0}^{e-1}(-1)^j = 1$.

Divide by n and sum up to Y. After writing $n = dm$, the inner sum is $\sum_{m\leq Y/d} \chi(m)/m$. Its nonzero terms begin $1 - 1/3 + 1/5 - 1/7 + \cdots$, so each partial sum is at most 1. The weights are nonnegative. This gives

$$H(Y) \leq \sum_{d\leq Y} \frac{w(d)}{d} \leq \prod_{p\in S} \left(1 + \frac{2}{p} + \frac{2}{p^2} + \cdots\right) = \prod_{p\in S} \frac{p+1}{p-1} \leq P^2.$$

The product includes every d with positive weight in the finite sum. The last inequality follows from $(p+1)(p-1) \leq p^2$.

Now double the number of odd denominators from m to $2m$. Each of the m new denominators is less than $4m$, so the reciprocal sum increases by at least $1/4$. After k doublings, the first 2^k odd reciprocals sum to at least $1 + k/4$. Take $k = 400000000$. The largest denominator is $2^{k+1} - 1 \leq Y$, and $1 + k/4 > 10000^2$. Thus $P > 10000$. Finally, $(p/(p-1))^2 \leq p$ for $p \geq 3$, so $M \geq P^2 > 200000$. The cited sieve estimate now applies. □

7. Representations by two summands

How many ways can an integer be written as a sum of two numbers from a fixed set? For a set A of nonnegative integers, put

$$r_A(t) = \#\{(x, y) \in A \times A : x + y = t\}.$$

This is an *ordered* count. Thus $2 + 3$ and $3 + 2$ contribute twice, whereas $2 + 2$ contributes once. The Erdős–Turán conjecture asks whether $r_A(t)$ must be unbounded whenever every sufficiently large integer has a representation [42]. Our first two results give a way to modify A and a sharp finite test for its representation counts.

7.1. **Adjoining one summand.**

Theorem 7.1. *For every $a, t \geq 0$,*

$$r_A(t) \leq r_{A\cup\{a\}}(t) \leq r_A(t) + 2.$$

The constant two is attained.

Proof. All old pairs remain available. A new pair must have a as one of its two coordinates. If $a > t$, this is impossible. Otherwise its partner must be $t - a$, leaving just $(a, t - a)$ and $(t - a, a)$. These coincide if $t = 2a$. For equality at the upper bound, take $A = \{2\}$, $a = 3$ and $t = 5$. The count rises from zero to two. □

Adding k elements can therefore increase each count by at most $2k$. In particular, a set with bounded representation counts keeps that property after filling a finite initial gap. This observation is one step in TheWaler's reduction of the conjecture. The same lemma appears in an earlier submission [43, 44].

7.2. **The sharp endpoint for five representations.** Suppose we require every integer from 0 to L to occur as a sum, and we allow at most five representations of any of these integers. How large can L be? Grekos, Haddad, Helou and Pihko determined the endpoint in 2003 [1, §6.4]. We give an explicit set on one side of the endpoint and a finite certificate on the other. This result enters the reduction in [45].

Theorem 7.2. *If $r_S(t) \geq 1$ for every $0 \leq t \leq 70$, then $r_S(t) \geq 6$ for some t in that range. There is a set S for which $1 \leq r_S(t) \leq 5$ for every $0 \leq t \leq 69$.*

Proof. For the existence assertion take

$$S = \{0, 1, 3, 4, 5, 6, 10, 15, 17, 23, 31, 39, 47, 55, 63\}.$$

For each t, count the $a \in S$ for which $t - a \in S$. The resulting frequency table is

j	1	2	3	4	5
$\#\{0 \leq t \leq 69 : r_S(t) = j\}$	4	42	3	17	4.

For example, $t = 6$ has five representations, with first summands $0, 1, 3, 5, 6$. The frequencies add to 70, covering the whole interval.

For the upper bound, only elements at most $L = 70$ can contribute. Decide in order whether $0, 1, \ldots, L$ belong to S. Given a current choice $B \subseteq \{0, \ldots, x-1\}$, write $c(t) = r_B(t)$ for each $0 \leq t \leq L$. Omitting x leaves every count unchanged. Including x gives the update

$$c'(t) = c(t) + 2\,\mathbf{1}_{\{t-x \in B\}} + \mathbf{1}_{\{t=2x\}},$$

where $\mathbf{1}$ is one when its condition holds and zero otherwise. After either choice, discard a branch if any count exceeds five, or if the count at x is zero. An excessive count can never fall. An uncovered x can never be supplied by a larger nonnegative summand.

These two tests give a complete finite algorithm. Start with the empty set and zero counts, visit both choices at every surviving state, and count the states that survive the decision at L. Every admissible set gives one surviving path, and every surviving path gives an admissible set. The enumeration gives 29 surviving sets for $L = 69$ and none for $L = 70$. In the archived Lean calculation the search is divided into 42 directly evaluated subtrees. When a branch is closed for excessive multiplicity, the checker exhibits six distinct first summands and tests their partners. The calculation thus excludes every possible set. □

The theorem forces six ordered representations once sums cover the initial interval. Proving the Erdős–Turán conjecture would require controlling arbitrarily large multiplicities.

7.3. **An encoding for dyadic intervals.** An infinite *Sidon set* is a set whose unordered pairs have distinct sums, allowing the two elements of a pair to coincide. Erdős problem 1191(ii) asks whether such a set A can have at least $C\sqrt{x}/(\log x)^c$ elements up to x for all sufficiently large x, for some $C, c > 0$ [46]. In mucoucah's reduction, one task is to tell which two intervals supplied a sum [47]. Nicholas Wierenga used the indices of the intervals and their squares in an earlier related construction [48]. We give an explicit bound for the size of a code carrying this information.

Theorem 7.3. *For $a \geq 1$ put $J(a) = \lfloor \log_2 a \rfloor$, and put $J(0) = 0$. There is a map E from the nonnegative integers to themselves such that $E(a) + E(b)$ determines both $a + b$ and the unordered pair $J(a), J(b)$, including repetitions. For every $j \geq 0$,*

$$2^j \leq a < 2^{j+1} \quad \Longrightarrow \quad 1 \leq E(a) \leq 2^j (j+2)^{16}.$$

Proof. Write the binary digits of a from right to left in blocks of lengths $1, 2, 4, 8, \ldots$. Block t, starting at $t = 0$, holds a number x_t in the range $0 \le x_t < B_t$, where $B_t = 2^{2^t}$. Its place value in a is

$$P_t = \prod_{u<t} B_u = 2^{2^t-1}.$$

Let c_t be digit t of $j = J(a)$ in base two and let s_t be digit t of j^2 in base four. Store the three digits together as

$$d_t = 64x_t + 8s_t + c_t, \qquad E(a) = \sum_{t\ge 0} d_t\, 128^t P_t.$$

Only finitely many digits are nonzero. The new place values have successive ratios $128B_t$. This is positional notation in which different positions have different bases.

Consider a sum of two codes. The sum of two c_t digits is at most 2, and the sum of two s_t digits is at most 6. Their combined contribution is at most $8 \cdot 6 + 2 = 50 < 64$. The two data digits contribute at most $64(2B_t - 2)$, so the whole sum at position t is less than $128B_t$. Consequently no carry reaches the next position. Dividing the value at each position successively by 8 and 8 recovers the separate sums of its three digits. Using their original place values now gives

$$a + b, \qquad h + \ell, \qquad h^2 + \ell^2, \qquad h = J(a), \quad \ell = J(b).$$

The last two numbers give $h\ell = ((h + \ell)^2 - (h^2 + \ell^2))/2$. The roots of $X^2 - (h + \ell)X + h\ell$ are h and ℓ, so their order is the only information lost.

For the size bound, fix a in the stated interval and put $k = \lfloor \log_2(j + 1) \rfloor$. Blocks 0 through k hold $2^{k+1} - 1 \ge j + 1$ binary digits of a. Also $j < 2^{k+1}$ and $j^2 < 4^{k+1}$, so all digits used for the index and its square fit. Since $8s_t + c_t \le 25$, we obtain

$$E(a) \le 128^k \left(64a + 25 \sum_{t=0}^{k} P_t \right).$$

The place values at least double at each step. Their sum is less than $2P_k$, and $P_k = 2^{2^k-1} \le 2^j$. Substituting $a < 2^{j+1}$ gives $E(a) < 178 \cdot 128^k 2^j$. Finally, $128^k = (2^k)^7 \le (j+2)^7$ and $178 \le (j+2)^9$. A positive input has a nonzero data digit, so its code is positive. □

For example, $0 + 2$ and $1 + 1$ have the same ordinary sum. Their interval pairs are $(0, 1)$ and $(0, 0)$. In the construction $E(0) = 0$, $E(1) = 64$ and $E(2) = 9 + 64 \cdot 256$, so the two sums of codes are different. The theorem supplies this encoding. Producing a Sidon set of the required density needs the other parts of the reduction.

8. Products of consecutive integers

A product of consecutive integers often has many small prime factors. Two questions of Erdős ask whether such products can divide one another. The first fixes the start of two adjacent blocks and lets their common length vary. The second prescribes the quotient and allows the two blocks to move.

8.1. **Comparing prime exponents in adjacent blocks.**

For positive integers n, k, write

$$U = \prod_{i=0}^{k-1} (n + i), \qquad V = \prod_{i=0}^{k-1} (n + k + i).$$

Erdős problem 389 asks whether each n admits a k for which U divides V [49]. For example, with $n = 2$ and $k = 3$ the products are 24 and 210. The first contains three factors of 2, while the second contains only one. This value of k fails a test that also gives a sufficient condition.

Theorem 8.1. *For a prime p let $v_p(M)$ be its exponent in the prime factorization of the positive integer M. If $v_p(U) \le v_p(V)$ for every prime p, then U divides V.*

Proof. Let P be the finite set of primes dividing UV. Each exponent difference is a nonnegative integer, so

$$Q = \prod_{p \in P} p^{v_p(V) - v_p(U)}$$

is a positive integer. For every $p \in P$, the exponent of p in UQ is $v_p(V)$. All other prime exponents vanish in both numbers. Unique factorization gives $UQ = V$. □

This is the prime-exponent step of Ciphore's reduction [50]. Genbox also supplied an earlier independent proof [51]. Finding a successful k for each n is the part that the exponent comparisons leave open.

8.2. **A prescribed nonsquare quotient from blocks of length two.** Erdős problem 686 asks whether every $N \geq 2$ occurs as the ratio of two disjoint equal-length blocks of consecutive positive integers, each of length at least two [52]. For the nonsquare values two consecutive factors in each block always suffice. Erdős and Straus gave the Pell-equation argument in 1977 [53, p. 63]. The submitted formalization builds on the proof in the Formal Conjectures library [54].

Theorem 8.2. *For every nonsquare integer $N \geq 2$ there are integers $n \geq 0$ and $m \geq n + 2$ such that*

$$N = \frac{(m+1)(m+2)}{(n+1)(n+2)}.$$

Proof. Set $X = 2m + 3$ and $Y = 2n + 3$. Then $4(m+1)(m+2) = X^2 - 1$, so the desired identity becomes

$$X^2 - NY^2 = 1 - N.$$

We need positive odd solutions with $Y \geq 3$ and $X - Y \geq 4$. The pair $(1, 1)$ solves the equation but is too small.

Pell's theorem gives positive integers u, v with $u^2 - Nv^2 = 1$. Write $(u + v\sqrt{N})^2 = U + V\sqrt{N}$. Here $U = 1 + 2Nv^2$ is odd, $V = 2uv$ is even, and $U^2 - NV^2 = 1$. Multiply $X + Y\sqrt{N}$ by $U + V\sqrt{N}$ to get the new pair

$$(X', Y') = (UX + NVY,\ VX + UY).$$

Its norm $X'^2 - NY'^2$ is $(U^2 - NV^2)(X^2 - NY^2) = 1 - N$. Starting at $(1, 1)$, all pairs remain positive and odd. At each step $Y' \geq UY$ with $U > 1$, so Y tends to infinity. The equation gives $X/Y \to \sqrt{N} > 1$, and hence $X - Y \to \infty$. Eventually both size conditions hold. Putting $m = (X - 3)/2$ and $n = (Y - 3)/2$ finishes the proof. □

For instance $5 = (5 \cdot 6)/(2 \cdot 3)$. The nonsquare hypothesis is exactly what allows the use of Pell's theorem. The square quotients form a separate case of problem 686.

9. Forcing an arithmetic progression

9.1. **Three fixed red sites on a board of length 158.** Color a row of integers red and blue. An arithmetic progression of length m has the form $a, a + d, \ldots, a + (m-1)d$ with $d \geq 1$. Write $W(3, 13)$ for the least board length at which every coloring contains three equally spaced reds or thirteen equally spaced blues. Landman, Robertson and Culver determined $W(3, 13) = 160$ [55]. The following restricted statement on a shorter board is one branch of the formal project associated to Green's problem 14 [56, 57].

Theorem 9.1. *Color $0, 1, \ldots, 157$ red or blue, with $76, 78, 79$ red. There is either a red three-term arithmetic progression or a blue thirteen-term arithmetic progression.*

Proof. Suppose neither progression occurs. When two sites $x < y$ are red, each of $2x - y$, $(x + y)/2$, and $2y - x$ would complete a red triple if it were an integer on the board. Such a site is forced blue. For example, red sites 76 and 78 force $74, 77, 80$ blue.

There is a corresponding test in the opposite direction. Every thirteen-term progression must contain red. For spacing $1 \le d \le 13$ its starting site ranges from 0 to $157 - 12d$. These ranges give $\sum_{d=1}^{13}(158 - 12d) = 962$ progressions. If the sites of any one are all forbidden to be red, the current partial coloring is impossible. If only one site can still be red, that site is forced red. Otherwise a site in an unresolved progression may be assigned each color in turn, producing two smaller cases.

The certificate records these choices and refers back to cases already closed [58]. A contradiction closes a leaf. A forced assignment transfers the result from its one remaining child, and a split transfers it when both children are closed. Induction on these steps proves that a closed initial state has no admissible coloring. The Lean lemma `visit_sound` verifies this inference. The 27 checks `c774` through `c800` verify the certificate pieces, and `solvedB` joins them at the initial state with the three prescribed red sites. Hence every completion has one of the two progressions. □

The checker adapts earlier propagation work by MrSeizy, Vincent T. Mossman and PanPapag [59]. The reduction uses work by NickLeko, ZedingZhang and Anmol Sahu [57]. A separate branch of that reduction prescribes two blue sites.

10. What the proofs leave open

The two divisor-sum counterexamples illustrate a useful way to compare conjectures. Multiplication by a new prime rescales two divisor sums by the same factor. A suitable choice of that factor turns a counterexample to the first question into one satisfying the stronger assumptions of the second. The digit-block construction uses a different kind of persistence: a single difference-of-squares identity works for every repetition count. In each argument an algebraic identity explains why an example continues to work when a parameter changes.

The limitations of a finite argument are equally informative. The computation at 70 rules out every possible set at that endpoint. Raising the required multiplicity from six to an arbitrary integer would require a new argument about how these endpoints grow. Likewise, the encoding in Section 7.3 distinguishes the scales of two summands while preserving their sum, but by itself does not construct the dense infinite Sidon set asked for in Erdős problem 1191(ii). For the prime determinant in Section 2.1, one explicit null vector accounts for the zero at size four; excluding further zeros remains a distinct question.

Contributions and verification. Anmol Sahu directed the project and submitted the results under the account anmolspace. The formal work was developed with OpenAI Codex, ChatGPT and Claude Code; the original submissions and portfolio identify the recorded agent for each result. OpenAI Codex assisted with drafting and checking this exposition. References at each result credit earlier mathematical proofs, examples, reductions and formalizations. The finite enumerations in Sections 7.2 and 9.1 come with executable Lean proof certificates accessible in the table below.

Appendix A. Problem sources and formal proofs

Each row corresponds to one merged submission. The section link returns to its argument in this paper. The other links open the original problem, the submitted Lean source at its fixed Git revision, and the merged pull request. A dagger marks a step in a larger problem. Theorem 2.6 is a companion extension and has its own proof file linked after the theorem.

Section	Topic	Problem source	Proof	Record
2.1†	A vanishing prime determinant	A024356	Lean	#5195
2.2†	Five rows of Pascal's triangle	A249609	Lean	#5281
2.3	Partial sums of Catalan numbers	A110199	Lean	#7776
2.4	The real part of a Möbius sequence	A292273	Lean	#8351
2.5	Repeated decimal blocks	A086766	Lean	#945
3.1	On cells under Rule 1	A265722	Lean	#6382
3.2	Off cells under Rule 1	A265723	Lean	#6419
3.3	The central column of Rule 158	A265380	Lean	#6340
3.4	Binary rows under Rule 157	A263806	Lean	#6450
4.1	Prime and composite divisors	A306345	Lean	#6584
4.2	A square divisor-sum ratio	A301517	Lean	#7662
4.3	Two square divisor sums	A318807	Lean	#9203
4.4	A denominator and a derivative	A326691	Lean	#7763
4.5	Differentiating a hyperfactorial	A260619	Lean	#9122
4.6	A family for the totient inequality	A196200	Lean	#7807
5.1†	A covering of the exponents	A108129	Lean	#5665
5.2	Shifted powers modulo a divisor	A214074	Lean	#6211
5.3	Eleventh roots of unity	A141849	Lean	#6232
5.4	Two irreducible quadratic factors	A030433	Lean	#7813
6.1	Sparse rows of Gaussian primes	A069004	Lean	#5597
7.1†	Adjoining one summand	Erdős 28	Lean	#5250
7.2†	The sharp endpoint at 70	Erdős 28	Lean	#7608
7.3†	Encoding two dyadic intervals	Erdős 1191(ii)	Lean	#9061
8.1†	Adjacent block products	Erdős 389	Lean	#5292
8.2†	A Pell equation for block products	Erdős 686	Lean	#5821
9.1†	Three red sites on a finite board	Green 14	Lean	#9049

References

[1] G. Grekos, L. Haddad, C. Helou and J. Pihko, On the Erdős–Turán conjecture, J. Number Theory 102 (2003), §6.4.
[2] OEIS Foundation. A024356. Sequence definition, conjecture and contributor history.
[3] Open Math Agents. Earlier submission for this reduction step, PR 2176.
[4] Open Math Agents. Merged proof submission 5195 and verification history.
[5] OEIS Foundation. A249609. Sequence definition, conjecture and contributor history.
[6] Open Math Agents. Earlier submission for the five specified rows, PR 2211.
[7] Open Math Agents. Merged proof submission 5281 and verification history.
[8] OEIS A000108, Catalan numbers.
[9] Paul Barry and R. J. Mathar, OEIS A110199.
[10] OEIS A008683, Möbius function.
[11] Mats Granvik, OEIS A292273, real-part formula.
[12] xor2k, earlier project submission for A292273, PR 7021.

[13] OEIS Foundation. A086766. Sequence definition, conjecture and contributor history.
[14] sweety439. The conjecture in OEIS A086766 is false. MersenneForum, 2022 (archived discussion).
[15] Maxence Faldor. Formal disproof with the same block and factorization, timestamped source.
[16] R. Price and C. Barker, OEIS A265722.
[17] R. Price and C. Barker, OEIS A265723.
[18] R. Price, M. De Vlieger and C. Barker, OEIS A265380.
[19] snoozingnewt, Lean statement of A265380, submission 963.
[20] R. Price and C. Barker, OEIS A263806.
[21] snoozingnewt, Lean statement of A263806, submission 959.
[22] Earlier project submission on A263806, 6402.
[23] OEIS Foundation. A306345. Sequence definition, conjecture and contributor history.
[24] Michel Lagneau, OEIS A301517, divisor-sum ratio conjecture.
[25] Michel Lagneau, OEIS A318807, conjecture on the two divisor sums.
[26] Antti Karttunen, OEIS A326691, denominator sequence and conjecture.
[27] OEIS A003415, arithmetic derivative.
[28] Matthew Campbell, OEIS A260619, hyperfactorial derivative.
[29] OEIS A002109, hyperfactorials.
[30] József Sándor. On the Composition of Some Arithmetic Functions, II. JIPAM 6 (2005), Article 73. See equation (1.8) and Remark 3.26 for the question and Tóth's example.
[31] Hans Riesel. Some Large Prime Numbers. Elementa 39 (1956), 258–260. English translation by Lars Blomberg.
[32] OEIS Foundation. A108129. Sequence definition, conjecture and contributor history.
[33] OEIS Foundation. A214074. Sequence definition, conjecture and contributor history.
[34] Open Math Agents. Merged proof submission 6211 and verification history.
[35] OEIS Foundation. A141849. Sequence definition, conjecture and contributor history.
[36] OEIS A030433, comments by Federico Provvedi and Jianing Song.
[37] OEIS Foundation. A069004. Gaussian-prime rows and T. D. Noe's conjecture.
[38] PastaPastaPasta. Finite sieve disproof, Open Math Agents PR 2963. The imported implication appears as finite_inert_sieve_counterexample in PR 5597.
[39] skhakim. Earlier disproof, Open Math Agents PR 2077.
[40] Burntgogi. Earlier disproof, Open Math Agents PR 3471.
[41] cjchaffin. Earlier disproof, Open Math Agents PR 4006.
[42] Erdős Problems, problem 28.
[43] TheWaler, reduction of problem 28, submission 4711.
[44] Earlier submission on the same lemma, 4900.
[45] Reduction to finite bounds, submission 5979.
[46] Erdős Problems, problem 1191.
[47] mucoucah, quantitative reduction of part (ii), PR 7467.
[48] Nicholas Wierenga, earlier related encoding, PR 3944.
[49] Erdős Problems, problem 389.
[50] Ciphore, prime-exponent reduction, submission 4822.
[51] Genbox, independent proof of the implication, submission 5104.
[52] Erdős Problems, problem 686.
[53] P. Erdős and E. G. Straus, On Products of Consecutive Integers (1977), p. 63.
[54] Formal Conjectures, formalization of Erdős problem 686.
[55] B. Landman, A. Robertson and C. Culver, Some New Exact van der Waerden Numbers (2005).
[56] Ben Green, 100 open problems, problem 14.
[57] Reduction to the red and blue branches, PR 8422.
[58] Anmol Sahu, complete three-red certificate and verifier.
[59] Earlier coloring-propagation argument, PR 4877.