Signed Transport, Pair–Tail Reduction, and Low Layers in an Erdős Density-Doubling Problem

Przemyslaw Chojecki

20th March 2026

Abstract

Let A be a finite nonempty set of positive integers and define

$$B_A := \{n \geq 1 : a \mid n \text{ for some } a \in A\}, \qquad f_A(x) := |B_A \cap [1, x]|, \qquad D_A(x) := \frac{f_A(x)}{x}.$$

We study the Erdős problem asking whether

$$D_A(m) < 2D_A(n) \qquad (m > n \geq \max A)$$

always holds. The constant 2 is best possible for singleton sets. We prove the conjectured inequality in several substantial ranges. First, for each fixed finite A we obtain an explicit signed lcm-weight criterion implying the inequality for all sufficiently large n, together with a finite-window theorem for near-extremizers. Second, we introduce split counting functions

$$F_{U|V}(x) := \big|(B_U \setminus B_V) \cap [1, x]\big|$$

and prove a one-tail split-doubling theorem for singleton classes, with a sharp edge-case analysis, together with a pair-vs-one-tail theorem. As a consequence, the original Erdős inequality holds whenever the primitive reduction of A has size at most 3. Third, we isolate a union-bound reduction in the sparse regime and use it to prove a small-excess theorem: if $f_A(n) - |A_{\min}| \leq 5$, then $D_A(m) < 2D_A(n)$ for every $m > n$. In particular, all low layers $f_A(n) \leq 9$ are completely resolved.

We also record an exact quotient recursion for split counts, a two-block decomposition reducing the full problem to pair-vs-tail split inequalities, a derived overlap decomposition for the sparse bottleneck, and a compressed-predecessor reformulation of the order-slack problem. We then reformulate pair-tail systems through quotient tails and overlap graphs, obtaining an explicit finite-window criterion from cluster-expansion lower bounds and Janson upper bounds. Finally, we explain where the present local transport method breaks: singleton split doubling already fails for two forbidden moduli, so the first unresolved local transport problem is the pair-vs-two-tail case; we isolate a precise exact-one criterion which would settle that case.

Contents

1 Introduction

Let $A \subseteq \mathbb{N}$ be finite and nonempty. Write

$$B_A = \bigcup_{a\in A} a\mathbb{N}, \qquad f_A(x) = |B_A \cap [1,x]|, \qquad D_A(x) = \frac{f_A(x)}{x}.$$

We consider the following multiples-version of an Erdős problem:

$$D_A(m) < 2D_A(n) \qquad (m > n \ge \max A). \tag{1}$$

The formulation above appears in Erdős' 1966 paper [2] and in Guy's collection [3]. We assume throughout that $A \subseteq \{2,3,4,\dots\}$; if $1 \in A$, then $B_A = \mathbb{N}$ and the problem is trivial. The empty set is also trivial, so the only interesting case is a finite nonempty set $A \subseteq \{2,3,4,\dots\}$.

The constant 2 cannot be improved. Indeed, if $A = \{a\}$, $n = 2a-1$, and $m = 2a$, then

$$\frac{D_A(m)}{D_A(n)} = \frac{1/a}{1/(2a-1)} = 2 - \frac{1}{a} \to 2.$$

Our rigorous results are summarized below.

Theorem 1.1 (Summary of proved results)**.** *Let $A \subseteq \{2,3,\dots\}$ be finite.*

(a) *Let $A_{\min}$ denote the primitive reduction of A. If $|A_{\min}| \le 3$, then (1) holds for every $m > n \ge \max A$.*

(b) *If*
$$f_A(n) - |A_{\min}| \le 5,$$
then (1) holds for every $m > n \ge \max A$.

(c) *If $f_A(n) \le 9$, then (1) holds for every $m > n \ge \max A$.*

(d) *Let $\mathcal{L}(A)$ be the set of distinct lcm-values arising from nonempty subsets of A, and define*

$$\lambda_d := \sum_{\substack{\emptyset \ne S \subseteq A \\ \operatorname{lcm}(S)=d}} (-1)^{|S|+1}, \qquad \delta_A := \sum_{d\in\mathcal{L}(A)} \frac{\lambda_d}{d}.$$

Let

$$W_+(A) := \sum_{\lambda_d > 0} \lambda_d, \qquad W_-(A) := \sum_{\lambda_d < 0} (-\lambda_d).$$

If

$$n > \frac{2W_+(A) + W_-(A)}{\delta_A},$$

then

$$D_A(m) < 2D_A(n) \qquad (m > n).$$

(e) *More generally, if* $0 < \varepsilon < \delta_A$, $m > n$, *and*

$$\frac{D_A(m)}{D_A(n)} > 2 - \varepsilon,$$

then necessarily

$$n < \frac{2W_+(A) + W_-(A)}{\delta_A - \varepsilon}.$$

A further result, proved in section 8, gives an explicit finite-window criterion for pair-tail split inequalities in terms of quotient-tail overlap graphs, cluster-expansion lower bounds, and a Janson upper bound.

Several core ingredients below — the singleton theorem, the singleton-vs-one-tail theorem, the dense case, and the union-bound reduction — were independently checked in Lean 4. Where it clarifies the argument, we follow that proof structure.

In section 2 we record the exact signed lcm-transport identity and its large-n corollaries. In section 3 we prove the singleton and reduced two-generator cases of (1). In section 4 we introduce split counts $F_{U|V}$, prove the split doubling theorem for singleton classes against one forbidden modulus with a sharp edge-case analysis and for pairs against one forbidden modulus, and deduce (1) for primitive sets of size at most 3. In section 5 we prove a general small-excess theorem and resolve all layers $f_A(n) \leq 9$. In section 6 we isolate an exact sparse combinatorial bottleneck together with derived-overlap and compressed-predecessor reformulations. In section 7 we explain where the present local transport method stops, correct the naive singleton-tail route, and isolate the pair-vs-two-tail bottleneck. Finally, in section 8 we reformulate pair-tail split sets through quotient-tail systems and overlap graphs, and derive a cluster-style finite-window criterion for pair-tail inequalities.

2 Primitive reduction and signed lcm transport

2.1 Primitive reduction

Lemma 2.1 (Primitive reduction). *If $a, b \in A$ and $a \mid b$, then removing b from A does not change B_A. Consequently every finite A has a unique inclusion-minimal subset $A_{\min} \subseteq A$ such that*

$$B_{A_{\min}} = B_A,$$

and no two distinct elements of $A_{\min}$ divide one another.

Proof. If $a \mid b$, then $b\mathbb{N} \subseteq a\mathbb{N}$, so b is redundant in the union $B_A = \bigcup_{c \in A} c\mathbb{N}$. Repeating this deletion finitely many times leaves exactly the elements of A that are minimal under divisibility. Hence the resulting inclusion-minimal set is unique. □

We henceforth refer to $A_{\min}$ as the *primitive reduction* of A.

2.2 The lcm basis

For finite A write

$$\mathcal{L}(A) := \{\mathrm{lcm}(S) : \emptyset \neq S \subseteq A\}, \qquad \lambda_d := \sum_{\substack{\emptyset \neq S \subseteq A \\ \mathrm{lcm}(S) = d}} (-1)^{|S|+1} \qquad (d \in \mathcal{L}(A)).$$

Grouping the usual inclusion–exclusion expansion by equal lcm-values gives the following exact formula.

Lemma 2.2 (Inclusion–exclusion in the lcm basis)**.** *For every finite A and every $x \geq 1$,*

$$f_A(x) = \sum_{d \in \mathcal{L}(A)} \lambda_d \left\lfloor \frac{x}{d} \right\rfloor.$$

Consequently,

$$\delta_A := \sum_{d \in \mathcal{L}(A)} \frac{\lambda_d}{d} > 0$$

is the natural density of B_A, and

$$D_A(x) = \delta_A + O_A\left(\frac{1}{x}\right).$$

Moreover,

$$\sum_{d \in \mathcal{L}(A)} |\lambda_d| \leq 2^{|A|} - 1.$$

Proof. The first identity is inclusion–exclusion, grouped by common lcm. The density statement follows by replacing each floor by its linear part. Finally,

$$\sum_{d \in \mathcal{L}(A)} |\lambda_d| \leq \sum_{\emptyset \neq S \subseteq A} 1 = 2^{|A|} - 1,$$

since each nonempty subset contributes one sign ± 1 to exactly one coefficient λ_d. □

2.3 Exact transport and large-n corollaries

Proposition 2.3 (Exact signed transport identity)**.** *Let $m > n \geq 1$. Then*

$$D_A(m) - 2D_A(n) = -\delta_A + \sum_{d \in \mathcal{L}(A)} \lambda_d \left(\frac{2}{n}\left\{\frac{n}{d}\right\} - \frac{1}{m}\left\{\frac{m}{d}\right\} \right),$$

where $\{x\} = x - \lfloor x \rfloor$ denotes the fractional part.

Proof. By lemma 2.2,

$$D_A(x) = \sum_{d \in \mathcal{L}(A)} \lambda_d \frac{\lfloor x/d \rfloor}{x} = \sum_{d \in \mathcal{L}(A)} \frac{\lambda_d}{d} - \sum_{d \in \mathcal{L}(A)} \lambda_d \frac{\{x/d\}}{x} = \delta_A - \sum_{d \in \mathcal{L}(A)} \lambda_d \frac{\{x/d\}}{x}.$$

Subtracting $2D_A(n)$ from $D_A(m)$ gives the stated formula. □

Define

$$W_+(A) := \sum_{\lambda_d > 0} \lambda_d, \qquad W_-(A) := \sum_{\lambda_d < 0} (-\lambda_d).$$

Then by lemma 2.2,

$$W_+(A) + W_-(A) = \sum_{d \in \mathcal{L}(A)} |\lambda_d| \leq 2^{|A|} - 1. \tag{2}$$

Proposition 2.4 (Signed transport upper bound)**.** *For every $m > n \geq 1$,*

$$D_A(m) - 2D_A(n) \leq -\delta_A + \frac{2W_+(A)}{n} + \frac{W_-(A)}{m} \leq -\delta_A + \frac{2W_+(A) + W_-(A)}{n}.$$

Proof. From proposition 2.3, each term with $\lambda_d > 0$ is at most $2\lambda_d/n$, while each term with $\lambda_d < 0$, say $\lambda_d = -\mu_d$ with $\mu_d > 0$, is at most μ_d/m. Summing gives the first inequality; the second follows from $m > n$. □

Corollary 2.5 (Explicit large-n criterion)**.** *If*

$$n > \frac{2W_+(A) + W_-(A)}{\delta_A},$$

then

$$D_A(m) < 2D_A(n) \qquad (m > n).$$

In particular, by (2)*, the simpler sufficient condition*

$$n > \frac{2(2^{|A|} - 1)}{\delta_A}$$

also implies (1)*.*

Proof. Under the displayed hypothesis, proposition 2.4 gives $D_A(m) - 2D_A(n) < 0$. The second claim follows from (2). □

Corollary 2.6 (Near-extremizers lie in a finite window)**.** *Fix $0 < \varepsilon < \delta_A$. If $m > n \geq \max A$ and*

$$\frac{D_A(m)}{D_A(n)} > 2 - \varepsilon,$$

then necessarily

$$n < \frac{2W_+(A) + W_-(A)}{\delta_A - \varepsilon}.$$

Proof. From the hypothesis and $D_A(n) \leq 1$ we obtain

$$D_A(m) - 2D_A(n) > -\varepsilon D_A(n) \geq -\varepsilon.$$

Together with proposition 2.4,

$$-\varepsilon < -\delta_A + \frac{2W_+(A) + W_-(A)}{n},$$

which rearranges to the stated bound. □

3 The tail-free singleton and two-generator cases

Theorem 3.1 (Singleton case)**.** *For every $d \geq 2$ and every $m > n \geq d$,*

$$\frac{\lfloor m/d \rfloor}{m} < 2\frac{\lfloor n/d \rfloor}{n}.$$

Proof. Write $u := \lfloor n/d \rfloor \geq 1$. Then $\lfloor m/d \rfloor / m \leq 1/d$. Since $n < d(u + 1)$,

$$2\frac{\lfloor n/d \rfloor}{n} = 2\frac{u}{n} > \frac{2u}{d(u + 1)}.$$

If $u \geq 2$, then $2u/(u + 1) > 1$, so the right-hand side exceeds $1/d$. If $u = 1$, then $d \leq n \leq 2d - 1$, so

$$2\frac{\lfloor n/d \rfloor}{n} = \frac{2}{n} \geq \frac{2}{2d - 1} > \frac{1}{d}.$$

Thus in all cases $\lfloor m/d \rfloor / m < 2 \lfloor n/d \rfloor / n$. □

Theorem 3.2 (Reduced two-generator case)**.** *Let $2 \le a < b$ and assume $a \nmid b$. Then for every $m > n \ge b$,*

$$\frac{|(a\mathbb{N} \cup b\mathbb{N}) \cap [1,m]|}{m} < 2\frac{|(a\mathbb{N} \cup b\mathbb{N}) \cap [1,n]|}{n}.$$

Proof. Set

$$\nu := \left\lfloor \frac{n}{a} \right\rfloor, \qquad \omega := \left\lfloor \frac{n}{b} \right\rfloor, \qquad \eta := \left\lfloor \frac{n}{\operatorname{lcm}(a,b)} \right\rfloor.$$

Then

$$|(a\mathbb{N} \cup b\mathbb{N}) \cap [1,n]| = \nu + \omega - \eta.$$

Also

$$\frac{|(a\mathbb{N} \cup b\mathbb{N}) \cap [1,m]|}{m} \le \frac{\lfloor m/a \rfloor + \lfloor m/b \rfloor}{m} \le \frac{1}{a} + \frac{1}{b}.$$

Since $\nu \le n/a < \nu + 1$ and $\omega \le n/b < \omega + 1$,

$$\frac{1}{a} + \frac{1}{b} < \frac{\nu+1}{n} + \frac{\omega+1}{n} = \frac{\nu+\omega+2}{n}.$$

So it suffices to prove

$$\nu + \omega + 2 \le 2(\nu + \omega - \eta),$$

that is,

$$2 \le \nu + \omega - 2\eta. \tag{3}$$

Let $g := \gcd(a,b)$. Because $a \nmid b$, one has $a/g \ge 2$, and therefore

$$\operatorname{lcm}(a,b) = \frac{ab}{g} = b\,\frac{a}{g} \ge 2b.$$

Hence

$$\eta \le \left\lfloor \frac{n}{2b} \right\rfloor \le \left\lfloor \frac{\omega}{2} \right\rfloor.$$

Thus

$$\nu + \omega - 2\eta \ge \nu + \omega - 2\left\lfloor \frac{\omega}{2} \right\rfloor \in \{\nu, \nu+1\}.$$

If $\nu \ge 2$, then (3) holds. If $\nu = 1$, then $n < 2a < 2b$, so $\omega = 1$ and therefore $\eta = 0$. Thus again $\nu + \omega - 2\eta = 2$. This proves the claim. □

4 Split transport, pair-vs-one-tail, and pair-tail reduction

4.1 Split counts and exact recursion

For finite sets $U, V \subseteq \mathbb{N}$ define

$$B_U := \bigcup_{u \in U} u\mathbb{N}, \qquad F_{U|V}(x) := |(B_U \setminus B_V) \cap [1,x]|.$$

When $U = \{d\}$ we write $F_{d|V}$, and when $U = \{a,b\}$ we write $F_{a,b|V}$.

We allow the value 1 to appear in quotient-recursion descendants. In particular, if $1 \in V$ then $B_V = \mathbb{N}$ and hence $F_{U|V} \equiv 0$. Overlaps between U and V are also harmless: the definition depends only on the unions B_U and B_V.

Lemma 4.1 (Split inclusion–exclusion)**.** *For all finite sets $U, V \subseteq \mathbb{N}$ and every $x \ge 1$,*

$$F_{U|V}(x) = \sum_{\emptyset \ne S \subseteq U} \sum_{T \subseteq V} (-1)^{|S|+|T|+1} \left\lfloor \frac{x}{\operatorname{lcm}(S \cup T)} \right\rfloor.$$

Proof. This is inclusion–exclusion for the event "some $u \in U$ divides" together with the exclusions "no $v \in V$ divides". $\square$

Proposition 4.2 (Exact quotient recursion)**.** *Let $U, V \subseteq \mathbb{N}$ be finite and let $h \in V$. Put $V' := V \setminus \{h\}$ and*

$$U^{(h)} := \left\{ \frac{u}{\gcd(u,h)} : u \in U \right\}, \qquad V^{(h)} := \left\{ \frac{v}{\gcd(v,h)} : v \in V' \right\}.$$

Then for every $x \geq 1$,

$$F_{U|V}(x) = F_{U|V'}(x) - F_{U^{(h)}|V^{(h)}}\left(\left\lfloor \frac{x}{h} \right\rfloor \right).$$

Proof. Numbers counted by $F_{U|V'}(x)$ but not by $F_{U|V}(x)$ are precisely those $n \leq x$ which are divisible by some $u \in U$, by no element of V', and also by h. Writing $n = ht$, the divisibility condition $u \mid ht$ is equivalent to $u/\gcd(u,h) \mid t$, and similarly $v \nmid ht$ for $v \in V'$ is equivalent to $v/\gcd(v,h) \nmid t$. This gives the second term. $\square$

4.2 Singletons against one forbidden modulus

The Lean formalization naturally isolates the following refined one-tail estimate.

Lemma 4.3 (Uniform refined one-tail estimate)**.** *Let $d, e \geq 2$ and $m > n \geq d$. Then*

$$\frac{F_{d|\{e\}}(m)}{m} < \frac{F_{d|\{e\}}(n) + 1}{n} + \frac{1}{m} < \frac{F_{d|\{e\}}(n) + 2}{n}.$$

Proof. If $e \mid d$, then $F_{d|\{e\}} \equiv 0$ and there is nothing to prove. Assume therefore that $e \nmid d$, and write

$$L := [d,e], \qquad u := \left\lfloor \frac{n}{d} \right\rfloor, \qquad U := \left\lfloor \frac{m}{d} \right\rfloor, \qquad \omega := \left\lfloor \frac{n}{L} \right\rfloor, \qquad \Omega := \left\lfloor \frac{m}{L} \right\rfloor.$$

Then $F_{d|\{e\}}(x) = \lfloor x/d \rfloor - \lfloor x/L \rfloor$ for every x.

Since $m > n \geq d$, one has

$$nU < (u+1)m,$$

because $U \leq m/d$ and hence $nU \leq nm/d < (u+1)m$ from $n < d(u+1)$. Also $m < (\Omega + 1)L$, so

$$\omega m < \omega(\Omega + 1)L = \omega\Omega L + \omega L \leq n\Omega + n,$$

using $n \geq \omega L$. Hence

$$n\Omega > \omega m - n.$$

Subtracting gives

$$n(U - \Omega) < (u - \omega + 1)m + n = \big(F_{d|\{e\}}(n) + 1\big)m + n.$$

Dividing by mn yields the first displayed inequality; the second follows from $m > n$. $\square$

Theorem 4.4 (Singleton against one forbidden modulus)**.** *Let $d, e \geq 2$.*

(i) *If $e \mid d$, then $F_{d|\{e\}} \equiv 0$.*

(ii) *If $e \nmid d$, then for every $m > n \geq d$,*

$$\frac{F_{d|\{e\}}(m)}{m} < 2\frac{F_{d|\{e\}}(n)}{n}.$$

Proof. If $e \mid d$, then every multiple of d is also a multiple of e, so $F_{d|\{e\}} \equiv 0$.

Assume now that $e \nmid d$. Write $L = [d, e] = rd$ with $r \in \mathbb{N}$. Since $L = d$ would imply $e \mid d$, we have $r \geq 2$.

If $n < L$, then

$$F_{d|\{e\}}(n) = \left\lfloor \frac{n}{d} \right\rfloor, \qquad F_{d|\{e\}}(m) \leq \left\lfloor \frac{m}{d} \right\rfloor,$$

and the claim follows from theorem 3.1.

Assume now that $n \geq L$. By the second inequality in lemma 4.3,

$$\frac{F_{d|\{e\}}(m)}{m} < \frac{F_{d|\{e\}}(n) + 2}{n}.$$

If $F_{d|\{e\}}(n) \geq 2$, then $F_{d|\{e\}}(n) + 2 \leq 2F_{d|\{e\}}(n)$, and we are done.

It remains to treat the edge case $F_{d|\{e\}}(n) = 1$. Since $F_{d|\{e\}}$ is nondecreasing and

$$F_{d|\{e\}}(L) = \frac{L}{d} - 1 = r - 1,$$

we must have $r = 2$, so $L = 2d$. Moreover $n < 3d$, for otherwise

$$F_{d|\{e\}}(n) \geq F_{d|\{e\}}(3d) = 3 - 1 = 2,$$

contrary to $F_{d|\{e\}}(n) = 1$. Thus $2d \leq n < 3d$.

If $n < m < 3d$, then $F_{d|\{e\}}(m) = 1$, so

$$\frac{F_{d|\{e\}}(m)}{m} = \frac{1}{m} < \frac{2}{n} = 2\frac{F_{d|\{e\}}(n)}{n}.$$

If $m \geq 3d$, write $k := \lfloor m/d \rfloor \geq 3$. Since $L = 2d$,

$$F_{d|\{e\}}(m) = k - \left\lfloor \frac{k}{2} \right\rfloor.$$

If $k = 2s$, then

$$\frac{F_{d|\{e\}}(m)}{m} \leq \frac{s}{2sd} = \frac{1}{2d} < \frac{2}{3d} < \frac{2}{n}.$$

If $k = 2s + 1$, then

$$\frac{F_{d|\{e\}}(m)}{m} \leq \frac{s+1}{(2s+1)d} \leq \frac{2}{3d} < \frac{2}{n},$$

because $(s + 1)/(2s + 1)$ decreases for $s \geq 1$ and is maximal at $s = 1$. In all cases

$$\frac{F_{d|\{e\}}(m)}{m} < \frac{2}{n} = 2\frac{F_{d|\{e\}}(n)}{n}.$$

This proves the theorem. □

4.3 Pairs against one forbidden modulus

Theorem 4.5 (Pair against one forbidden modulus)**.** *Let $2 \leq a < b < c$ and assume that $a \nmid b$, $a \nmid c$, and $b \nmid c$. Then for every $m > n \geq c$,*

$$\frac{F_{a,b|\{c\}}(m)}{m} < 2\frac{F_{a,b|\{c\}}(n)}{n}.$$

Proof. Set

$$d_a := [a,c], \qquad d_b := [b,c],$$

and define

$$A(x) := \left\lfloor \frac{x}{a} \right\rfloor - \left\lfloor \frac{x}{d_a} \right\rfloor, \qquad B(x) := \left\lfloor \frac{x}{b} \right\rfloor - \left\lfloor \frac{x}{d_b} \right\rfloor.$$

Thus $A(x)$ counts multiples of a not divisible by c, and $B(x)$ counts multiples of b not divisible by c. One always has

$$F_{a,b|\{c\}}(x) \le A(x) + B(x), \tag{4}$$

because common multiples of a and b are counted twice on the right.

We split into two cases.

Case 1: $n < \min\{d_a, d_b\}$. Then below n no multiple of c can also be a multiple of a or b, so

$$F_{a,b|\{c\}}(n) = |(a\mathbb{N} \cup b\mathbb{N}) \cap [1,n]|.$$

Also

$$F_{a,b|\{c\}}(m) \le |(a\mathbb{N} \cup b\mathbb{N}) \cap [1,m]|.$$

Hence the claim follows from theorem 3.2.

Case 2: $n \ge \min\{d_a, d_b\}$. By symmetry we may assume $d_a \le d_b$, so $n \ge d_a$.

Applying the first inequality of lemma 4.3 to $(d,e) = (a,c)$ gives

$$\frac{A(m)}{m} < \frac{A(n)+1}{n} + \frac{1}{m} < \frac{A(n)+2}{n}. \tag{5}$$

Applying the same lemma to $(d,e) = (b,c)$ gives

$$\frac{B(m)}{m} < \frac{B(n)+1}{n} + \frac{1}{m} < \frac{B(n)+2}{n}. \tag{6}$$

Let $E_1(n)$ be the number of $x \le n$ which are divisible by exactly one of a, b and not by c, and let $E_2(n)$ be the number of $x \le n$ divisible by both a and b and not by c. Then

$$F_{a,b|\{c\}}(n) = E_1(n) + E_2(n), \qquad A(n) + B(n) = E_1(n) + 2E_2(n).$$

Combining (4), (5), and (6), we obtain

$$\frac{F_{a,b|\{c\}}(m)}{m} < \frac{A(n)+B(n)+4}{n} = 2\frac{F_{a,b|\{c\}}(n)}{n} - \frac{E_1(n)-4}{n}.$$

Thus it remains to prove

$$E_1(n) \ge 4. \tag{7}$$

Since $a \nmid c$, one has

$$d_a = [a,c] = c\,\frac{a}{\gcd(a,c)} \ge 2c > 2b.$$

Because $n \ge d_a$, the points $a, b, 2a, 2b$ all lie in $[1,n]$. Now a and b are exact-one points. Also $2a$ is exact-one: if $b \mid 2a$, then $b = 2a$, contradicting $a \nmid b$; if $c \mid 2a$, then $c = 2a$, contradicting $a \mid c$. Finally, $2b$ is not divisible by c, because $c > b$ and $c = 2b$ would imply $b \mid c$.

If $a \nmid 2b$, then $2b$ is also exact-one, and (7) is proved.

So assume $a \mid 2b$. Write

$$2b = ka, \qquad k \ge 3.$$

Then $\operatorname{lcm}(a,b) = 2b$.

If $k \ge 4$, then $3a < 2b = \operatorname{lcm}(a,b)$, so $b \nmid 3a$. Also $b \ge 2a$, hence $c > b \ge 2a$; if $c \mid 3a$, then necessarily $c = 3a$, contradicting $a \mid c$. Thus $3a$ is exact-one and c-free. Together with $a, b, 2a$ this gives four exact-one points.

If $k = 3$, then $2b = 3a$. We claim that $4a \le n$. Indeed, if $d_a = 3a$, then $c \mid d_a = 3a = 2b$; since $c > b$, this forces $c = 2b$, contradicting $b \mid c$. So $d_a \ge 4a$, hence $4a \le n$. Moreover, $b \nmid 4a$ because $b = 3a/2$, and if $c \mid 4a$, then $c > b = 3a/2$ forces $c = 2a$ or $c = 4a$, both impossible because then $a \mid c$. Thus $4a$ is exact-one and c-free. Again $a, b, 2a, 4a$ yield four exact-one points.

This proves (7), and hence the theorem. □

4.4 Block decomposition and pair-tail reduction

Proposition 4.6 (Two-block decomposition)**.** *Let $G = \{g_1 < g_2 < \cdots < g_r\}$ be primitive. Put $s := \lfloor r/2 \rfloor$ and define*

$$R_j := \{g_{2j-1}, g_{2j}\} \qquad (1 \le j \le s),$$

and, for $1 \le j \le s$,

$$T_j := \{g_{2j+1}, g_{2j+2}, \ldots, g_r\}.$$

If r is even, then

$$B_G = \bigsqcup_{j=1}^{s} (B_{R_j} \setminus B_{T_j}).$$

If r is odd, then

$$B_G = \left(\bigsqcup_{j=1}^{s} (B_{R_j} \setminus B_{T_j}) \right) \sqcup g_r\mathbb{N}.$$

In both cases the unions are disjoint.

Proof. Every integer $x \in B_G$ is divisible by at least one generator. Let j be the unique largest block index such that some element of R_j divides x; if r is odd and $g_r \mid x$, then x belongs to the final singleton block. By construction, x lies in the corresponding displayed piece and in no later piece. Conversely every displayed piece is contained in B_G. Hence the displayed unions are exact and disjoint. □

Corollary 4.7 (Primitive sets of size at most three)**.** *If the primitive reduction $A_{\min}$ of A has size at most* 3*, then* (1) *holds for every $m > n \ge \max A$.*

Proof. Replace A by $G = A_{\min}$. If $|G| = 1$, this is theorem 3.1. If $|G| = 2$, this is theorem 3.2. If $|G| = 3$, write

$$B_G = (B_{\{g_1,g_2\}} \setminus B_{\{g_3\}}) \sqcup g_3\mathbb{N}$$

using proposition 4.6. Apply theorem 4.5 to the first piece and theorem 3.1 to the second. Summing the two strict inequalities gives (1). □

Conjecture 4.8 (Pair-vs-tail split doubling)**.** Let $2 \le a < b$ and let $V \subseteq \{2, 3, 4, \ldots\}$ be finite with $b < \min V$. Then for every $m > n \ge \max(V \cup \{a, b\})$,

$$\frac{F_{a,b|V}(m)}{m} < 2\frac{F_{a,b|V}(n)}{n}.$$

Proposition 4.9 (Reduction to the pair-tail conjecture)**.** *If conjecture 4.8 holds, then the original Erdős inequality* (1) *holds for every finite A.*

Proof. Let $G = A_{\min} = \{g_1 < \cdots < g_r\}$, and write $s := \lfloor r/2 \rfloor$. By proposition 4.6, B_G is a disjoint union of the two-generator blocks $B_{R_j} \setminus B_{T_j}$, together with the final singleton block $g_r\mathbb{N}$ when r is odd.

For $j < s$, and also for $j = s$ when r is odd, the tail T_j is nonempty and satisfies

$$\max R_j = g_{2j} < g_{2j+1} = \min T_j.$$

Hence the corresponding contribution is exactly $F_{R_j|T_j}(x)$, and conjecture 4.8 gives its strict density-doubling inequality.

If r is even, then $T_s = \varnothing$, so the last block is

$$F_{R_s|\varnothing}(x) = |B_{R_s} \cap [1, x]|.$$

This block is handled by theorem 3.2. If r is odd, the final singleton block is handled by theorem 3.1.

Summing the resulting strict inequalities over the disjoint blocks yields (1). ☐

5 Small excess and the layers $f_A(n) \le 9$

The Lean formalization isolates the following general reduction, which packages the union bound together with a weak singleton estimate.

Proposition 5.1 (Union-bound reduction). *Let $G = A_{\min} = \{g_1, \dots, g_r\}$. If*

$$\sum_{i=1}^{r} \left\lfloor \frac{n}{g_i} \right\rfloor + r \le 2f_A(n),$$

then for every $m > n$,

$$D_A(m) < 2D_A(n).$$

Proof. Since $B_A = B_G$,

$$f_A(m) \le \sum_{i=1}^{r} \left\lfloor \frac{m}{g_i} \right\rfloor.$$

For each i,

$$\frac{\lfloor m/g_i \rfloor}{m} \le \frac{1}{g_i} < \frac{\lfloor n/g_i \rfloor + 1}{n},$$

because $\lfloor m/g_i \rfloor \le m/g_i$ and $n < g_i(\lfloor n/g_i \rfloor + 1)$. Hence

$$n \left\lfloor \frac{m}{g_i} \right\rfloor < \left(\left\lfloor \frac{n}{g_i} \right\rfloor + 1 \right) m.$$

Summing over i gives

$$nf_A(m) \le n \sum_{i=1}^{r} \left\lfloor \frac{m}{g_i} \right\rfloor < m \sum_{i=1}^{r} \left(\left\lfloor \frac{n}{g_i} \right\rfloor + 1 \right) = m \left(\sum_{i=1}^{r} \left\lfloor \frac{n}{g_i} \right\rfloor + r \right).$$

Under the displayed hypothesis, this is at most $2mf_A(n)$, proving

$$nf_A(m) < 2mf_A(n),$$

which is equivalent to (1). ☐

Theorem 5.2 (Small excess theorem). *Let $G = A_{\min}$ and write $r := |G|$. If*

$$f_A(n) - r \le 5,$$

then for every $m > n \ge \max A$,

$$D_A(m) < 2D_A(n).$$

Proof. Write $G = \{g_1 < \cdots < g_r\}$, set

$$E := f_A(n) - r, \qquad u_i := \left\lfloor \frac{n}{g_i} \right\rfloor \quad (1 \le i \le r),$$

and let

$$k := \#\{i : u_i \ge 2\}.$$

By proposition 5.1, it is enough to prove

$$\sum_{i=1}^{r} u_i \le r + 2E. \tag{8}$$

We use three elementary facts.

(F1) Each index with $u_i \ge 2$ contributes the distinct nongenerator point $2g_i \le n$. Distinctness is clear, and $2g_i$ cannot itself be a generator because that would make g_i divide another element of G. Hence $k \le E$.

(F2) If $u_i \ge E + 2$, then $2g_i, 3g_i, \ldots, (E+2)g_i$ give $E + 1$ distinct nongenerator points up to n; none can be a generator because each is a proper multiple of g_i. Thus $u_i \le E + 1$ for every i.

(F3) No even multiple $2tg$ with $t \ge 2$ can equal a double $2h$ of another generator: from $2tg = 2h$ one gets $h = tg$, contradicting primitivity. In particular, numbers such as $4g, 6g, \ldots$ are never doubles of generators.

If $E = 0$, then $k = 0$ by (F1), so every $u_i = 1$, and (8) is immediate.
If $E = 1$, then by (F1) at most one u_i exceeds 1, and by (F2) all $u_i \le 2$. Hence

$$\sum_{i=1}^{r} u_i \le (r-1) \cdot 1 + 2 = r + 1 < r + 2 = r + 2E.$$

If $E = 2$, then at most two indices satisfy $u_i \ge 2$, and all $u_i \le 3$. Therefore

$$\sum_{i=1}^{r} u_i \le (r-2) \cdot 1 + 3 + 3 = r + 4 = r + 2E.$$

Assume now that $E = 3$. If $k \le 2$, then

$$\sum_{i=1}^{r} u_i \le (r-2) \cdot 1 + 4 + 4 = r + 6 = r + 2E.$$

So only $k = 3$ needs attention. Let $a < b < c$ be the three generators with $u \ge 2$. Then the three doubles $2a, 2b, 2c$ already account for all three nongenerator points up to n. Hence $u_c \le 2$. Also $u_a, u_b \le 3$: if some $u_g \ge 4$, then $4g \le n$ is a nongenerator point which is not a double, by (F3), contradiction. Thus

$$\sum_{i=1}^{r} u_i \le (r-3) \cdot 1 + 3 + 3 + 2 = r + 5 < r + 6.$$

Assume next that $E = 4$. If $k \le 2$, then

$$\sum_{i=1}^{r} u_i \le (r-2) \cdot 1 + 5 + 5 = r + 8 = r + 2E.$$

If $k = 4$, let $a < b < c < d$ be the four generators with $u \geq 2$. Their doubles exhaust all four nongenerator points, so $u_d \leq 2$, while $u_a, u_b, u_c \leq 3$ by the same $4g$-argument as above. Hence

$$\sum_{i=1}^{r} u_i \leq (r-4) \cdot 1 + 3 + 3 + 3 + 2 = r + 7 < r + 8.$$

It remains to treat $E = 4$ and $k = 3$. Let $a < b < c$ be the three generators with $u \geq 2$. Then the nongenerator points are

$$2a,\ 2b,\ 2c,\ z$$

for one further point z.

We claim that $u_a \leq 4$. Otherwise $u_a \geq 5$, so $3a, 4a, 5a \leq n$ are three nongenerator points. Since $4a$ is not a double, and only one nongenerator point lies outside $\{2a, 2b, 2c\}$, both $3a$ and $5a$ must be doubles. Necessarily

$$3a = 2b, \qquad 5a = 2c.$$

Then $n \geq 2c = 5a$, so

$$3b = \frac{9a}{2} \leq n.$$

Moreover

$$2a < 2b = 3a < 4a < 3b < 2c = 5a,$$

so $3b$ is a second nongenerator point outside the doubles, contradiction. Hence $u_a \leq 4$.

Next, $u_b \leq 3$. Indeed, if $u_b \geq 4$, then also $u_a \geq 4$, so $4a$ and $4b$ are two distinct nongenerator points outside the doubles, contradicting the uniqueness of z. Therefore $u_c \leq u_b \leq 3$, and

$$\sum_{i=1}^{r} u_i \leq (r-3) \cdot 1 + 4 + 3 + 3 = r + 7 < r + 8.$$

Assume finally that $E = 5$. If $k \leq 2$, then

$$\sum_{i=1}^{r} u_i \leq (r-2) \cdot 1 + 6 + 6 = r + 10 = r + 2E.$$

If $k = 5$, let $a_1 < \cdots < a_5$ be the five generators with $u \geq 2$. Their doubles already account for all five nongenerator points. Hence $u_{a_5} \leq 2$, and $u_{a_1}, \ldots, u_{a_4} \leq 3$ by (F3). Therefore

$$\sum_{i=1}^{r} u_i \leq (r-5) \cdot 1 + 3 + 3 + 3 + 3 + 2 = r + 9 < r + 10.$$

If $k = 4$, write the four relevant generators as $a < b < c < d$. Then the nongenerator set has the form

$$2a,\ 2b,\ 2c,\ 2d,\ z.$$

We first claim that $u_a \leq 5$. Otherwise $u_a \geq 6$, so $3a, 4a, 5a, 6a \leq n$ are four nongenerator points. Since only one point lies outside the doubles, at least three of them would have to be doubles. But $4a$ and $6a$ are not doubles, by (F3), contradiction. Hence $u_a \leq 5$.

Next, $u_b \leq 3$: if $u_b \geq 4$, then also $u_a \geq 4$, so $4a$ and $4b$ are two distinct nongenerator points outside the doubles, again contradicting the uniqueness of z. Therefore

$$\sum_{i=1}^{r} u_i \leq (r-4) \cdot 1 + 5 + 3 + 3 + 3 = r + 10.$$

The last remaining case is $E = 5$ and $k = 3$. Let $a < b < c$ be the three generators with $u \geq 2$. Then the nongenerator set is

$$2a,\ 2b,\ 2c,\ z_1,\ z_2.$$

If $u_c \geq 4$, then $4a, 4b, 4c \leq n$ are three distinct nongenerator points outside the doubles, impossible. So $u_c \leq 3$.

We also claim $u_b \leq 4$. Suppose for contradiction that $u_b \geq 5$. Then also $u_a \geq 5$, so $4a, 4b \leq n$. By (F3) these are two distinct nongenerator points outside the doubles $2a, 2b, 2c$; since there are only two such points, they must be exactly z_1, z_2. Hence every other nongenerator point up to n must lie in

$$\{2a, 2b, 2c, 4a, 4b\}.$$

In particular, the proper multiples $3b$ and $5b$ belong to this set. Since $3b, 5b > 2b$, neither can equal $2a$ or $2b$, and clearly $5b \neq 4a, 4b$; therefore $5b = 2c$. Also $3b \neq 4b$, and $3b \neq 2c$ because $2c = 5b > 3b$. Thus $3b = 4a$, so $b = \frac{4}{3}a$. Consequently

$$n \geq 5b = \frac{20}{3}a > 6a,$$

and hence $6a \leq n$. Now $6a$ is a nongenerator point, and by (F3) it is not a double; moreover $6a \neq 4a, 4b$ because $4b = \frac{16}{3}a$. This gives a third nongenerator point outside the doubles, contradiction. Thus $u_b \leq 4$.

Finally $u_a \leq 6$ by (F2). Hence

$$\sum_{i=1}^{r} u_i \leq (r - 3) \cdot 1 + 6 + 4 + 3 = r + 10 = r + 2E.$$

Thus (8) holds in every case $E \leq 5$. By proposition 5.1, this proves the theorem. □

Corollary 5.3 (The layers $f_A(n) \leq 9$ are completely resolved)**.** *If A is finite and $f_A(n) \leq 9$, then for every $m > n \geq \max A$,*

$$D_A(m) < 2D_A(n).$$

Proof. Let $r := |A_{\min}|$. If $r \leq 3$, apply corollary 4.7. Otherwise $r \geq 4$, so

$$f_A(n) - r \leq 9 - 4 = 5,$$

and theorem 5.2 applies. □

Remark 5.4. The first unresolved low layer is $f_A(n) = 10$. Among those, every case with $|A_{\min}| \neq 4$ is already covered by the present arguments, so the only genuinely new layer-10 configurations have primitive reduction of size 4.

6 Further exact reformulations in the sparse regime

Let $G = A_{\min}$ and write $r := |G|$.

Proposition 6.1 (The dense case is automatic)**.** *If $D_A(n) \geq \frac{1}{2}$, then for every $m > n$ one has*

$$D_A(m) < 2D_A(n).$$

Proof. Since $1 \notin G$, not every positive integer belongs to B_A, so $D_A(m) < 1$ for every m. If $D_A(n) \geq 1/2$, then $2D_A(n) \geq 1$, and hence $D_A(m) < 1 \leq 2D_A(n)$. □

Thus the remaining difficulty lies entirely in the sparse regime

$$D_A(n) < \frac{1}{2}. \tag{9}$$

Set

$$\mu_G := \sum_{g \in G} \frac{1}{g}.$$

Lemma 6.2 (Strict reciprocal majorant)**.** *Let G be finite and primitive, and let*

$$\mu_G := \sum_{g \in G} \frac{1}{g}.$$

For every $x \geq 1$ one has

$$D_G(x) \leq \mu_G.$$

Moreover, if $|G| \geq 2$, then in fact

$$D_G(x) < \mu_G \qquad (x \geq 1),$$

while if $G = \{g\}$ is a singleton, equality holds if and only if $g \mid x$.

Proof. Always

$$D_G(x) = \frac{1}{x} \left| \bigcup_{g \in G} (g\mathbb{N} \cap [1, x]) \right| \leq \sum_{g \in G} \frac{\lfloor x/g \rfloor}{x} \leq \sum_{g \in G} \frac{1}{g} = \mu_G.$$

If $G = \{g\}$, then equality is exactly the condition $\lfloor x/g \rfloor = x/g$, that is, $g \mid x$.

Assume now that $|G| \geq 2$. If some $g \in G$ satisfies $g > x$, then $\lfloor x/g \rfloor = 0 < x/g$, so the second displayed inequality is strict. Otherwise every $g \in G$ is at most x. If equality held, then equality in the second displayed inequality would force $g \mid x$ for every $g \in G$. Choosing distinct $g, h \in G$, we would then have $\operatorname{lcm}(g, h) \mid x$, hence $\operatorname{lcm}(g, h) \leq x$; but the integer $\operatorname{lcm}(g, h)$ is counted at least twice in the sum $\sum_{g \in G} \lfloor x/g \rfloor$ and only once in the union, contradicting equality in the first displayed inequality. Therefore $D_G(x) < \mu_G$ for every $x \geq 1$ when $|G| \geq 2$. □

Lemma 6.3 (Single-time reduction)**.** *If*

$$D_A(n) \geq \frac{\mu_G}{2},$$

then (1) *holds for every $m > n$.*

Proof. If $|G| = 1$, write $G = \{g\}$. The hypothesis $D_A(n) \geq \mu_G/2 = 1/(2g)$ forces $\lfloor n/g \rfloor \geq 1$, hence $n \geq g$, and theorem 3.1 gives the result. Assume therefore that $|G| \geq 2$. Since $B_A = B_G$, lemma 6.2 gives

$$D_A(m) = D_G(m) < \mu_G \leq 2D_A(n).$$

This is exactly (1). □

We now isolate an exact combinatorial inequality which would settle the sparse regime.

Definition 6.4. For $x \leq n$ define the incidence degree

$$\deg_G(x) := \#\{g \in G : g \mid x\}.$$

Let

$$I_G(n) := \sum_{g \in G} \left\lfloor \frac{n}{g} \right\rfloor = \sum_{x \leq n} \deg_G(x), \qquad N_{\mathrm{nb}}(n) := f_A(n) - r,$$

and

$$\operatorname{Dup}_G(n) := \sum_{\substack{x \leq n \\ \deg_G(x) \geq 1}} (\deg_G(x) - 1) = I_G(n) - f_A(n).$$

Proposition 6.5 (Exact sparse bottleneck)**.** *The inequality*

$$I_G(n) + r \le 2f_A(n) \tag{10}$$

is equivalent to each of the following:

$$\mathrm{Dup}_G(n) \le N_{\mathrm{nb}}(n), \tag{11}$$

$$\sum_{\substack{x \le n \\ \deg_G(x) \ge 1}} \deg_G(x) \le 2f_A(n) - r. \tag{12}$$

Moreover, if (10) *holds, then* (1) *holds for every* $m > n$*.*

Proof. Since $\mathrm{Dup}_G(n) = I_G(n) - f_A(n)$ and $N_{\mathrm{nb}}(n) = f_A(n) - r$,

$$I_G(n) + r \le 2f_A(n) \iff I_G(n) - f_A(n) \le f_A(n) - r \iff \mathrm{Dup}_G(n) \le N_{\mathrm{nb}}(n).$$

This proves the equivalence of (10) and (11); the equivalence with (12) is immediate from $I_G(n) = \sum_{x \le n} \deg_G(x)$.

Finally,

$$\mu_G \le \sum_{g \in G} \frac{\lfloor n/g \rfloor + 1}{n} = \frac{I_G(n) + r}{n},$$

so (10) implies $\mu_G \le 2f_A(n)/n = 2D_A(n)$. Now lemma 6.3 applies. □

The first refinement records the overlap layers themselves.

Definition 6.6. For $k \ge 1$ define

$$\mathcal{D}_k(G) := \{\mathrm{lcm}(S) : S \subseteq G,\ |S| = k + 1\}.$$

Proposition 6.7 (Derived overlap decomposition)**.** *For every* $k \ge 1$*,*

$$f_{\mathcal{D}_k(G)}(n) = \#\{x \le n : \deg_G(x) \ge k + 1\}.$$

Consequently,

$$\mathrm{Dup}_G(n) = \sum_{k \ge 1} f_{\mathcal{D}_k(G)}(n),$$

where the sum is finite.

Proof. An integer $x \le n$ is counted by $f_{\mathcal{D}_k(G)}(n)$ if and only if it is divisible by $\mathrm{lcm}(S)$ for some $(k+1)$-subset $S \subseteq G$, equivalently if and only if every element of S divides x, that is, $\deg_G(x) \ge k + 1$. This proves the first identity.

Now for each $x \le n$ one has

$$(\deg_G(x) - 1)_+ = \sum_{k \ge 1} \mathbf{1}_{\deg_G(x) \ge k+1},$$

so summing over $x \le n$ gives

$$\mathrm{Dup}_G(n) = \sum_{x \le n} (\deg_G(x) - 1)_+ = \sum_{k \ge 1} \#\{x \le n : \deg_G(x) \ge k + 1\} = \sum_{k \ge 1} f_{\mathcal{D}_k(G)}(n).$$

The sum is finite because $\mathcal{D}_k(G) = \emptyset$ for $k \ge r$. □

Corollary 6.8 (Triple-free criterion)**.** *If every* $x \le n$ *is divisible by at most two generators from* G *— equivalently, if*

$$\mathrm{lcm}(g, h, \ell) > n \qquad \text{for all distinct } g, h, \ell \in G,$$

— then (10) *holds. Consequently* (1) *holds for every* $m > n$*.*

Proof. Under the hypothesis, $\deg_G(x) \le 2$ for every $x \le n$. Hence

$$\mathrm{Dup}_G(n) = \#\{x \le n : \deg_G(x) = 2\}.$$

Every such point is a nongenerator point, so $\mathrm{Dup}_G(n) \le N_{\mathrm{nb}}(n)$. By proposition 6.5, (10) follows, and therefore so does (1). □

The following order-slack identity is another exact reformulation of (10).

Proposition 6.9 (Order-slack identity)**.** *Fix an ordering $G = \{g_1, \dots, g_r\}$. Define*

$$\nu_i := \left\lfloor \frac{n}{g_i} \right\rfloor, \qquad \tau_i := \#\Big(g_i\mathbb{N} \cap [1,n] \setminus \bigcup_{j<i} g_j\mathbb{N}\Big).$$

Then

$$I_G(n) = \sum_{i=1}^{r} \nu_i, \qquad f_A(n) = \sum_{i=1}^{r} \tau_i,$$

and

$$2f_A(n) - I_G(n) - r = \sum_{i=1}^{r} (2\tau_i - \nu_i - 1). \tag{13}$$

In particular, if there exists an ordering with

$$\tau_i \ge \frac{\nu_i + 1}{2} \qquad (1 \le i \le r),$$

then (1) *holds.*

Proof. The first two identities are definitions. Summing them yields (13). If each summand on the right-hand side is nonnegative, then (10) holds, and proposition 6.5 finishes the proof. □

The local terms τ_i themselves admit an exact quotient-space description.

Proposition 6.10 (Compressed predecessor formula)**.** *In the notation of proposition 6.9, define*

$$P_i := \left\{ \frac{g_j}{\gcd(g_j, g_i)} : 1 \le j < i \right\}.$$

Then

$$\tau_i = \nu_i - f_{P_i}(\nu_i) \qquad (1 \le i \le r).$$

Consequently,

$$2f_A(n) - I_G(n) - r = \sum_{i=1}^{r} \left(\nu_i - 2f_{P_i}(\nu_i) - 1\right).$$

Proof. Every multiple of g_i up to n has the form $g_i t$ with $1 \le t \le \nu_i$. For $j < i$,

$$g_j \mid g_i t \iff \frac{g_j}{\gcd(g_j, g_i)} \mid t.$$

Thus $g_i t$ is already covered by some earlier generator if and only if $t \in B_{P_i}$. The uncovered multiples of g_i are therefore exactly the $t \le \nu_i$ outside B_{P_i}, so

$$\tau_i = \nu_i - f_{P_i}(\nu_i).$$

Substituting this into (13) yields the second identity. □

Conjecture 6.11 (Sparse order-slack conjecture)**.** For every primitive G and every $n \ge \max G$ satisfying (9), there exists an ordering of G for which the total slack in (13) is nonnegative. Equivalently, (10) should hold throughout the sparse regime.

By proposition 6.10, this can be rephrased as the existence of an ordering for which

$$\sum_{i=1}^{r} \left(\nu_i - 2f_{P_i}(\nu_i) - 1\right) \ge 0.$$

By propositions 6.1 and 6.5, conjecture 6.11 would imply the full Erdős inequality (1).

7 Barriers, corrections, and the next local bottleneck

We now record the first places where the split-transport approach fails, and explain what survives.

Proposition 7.1 (Singleton split doubling already fails for two tails)**.** *Let* $d = 5$*,* $V = \{2, 3\}$*,* $n = 24$*, and* $m = 35$*. Then*

$$\frac{F_{5|V}(35)}{35} > 2\frac{F_{5|V}(24)}{24}.$$

In particular, the singleton split inequality fails already for tail-size 2*.*

Proof. By inclusion–exclusion,

$$F_{5|\{2,3\}}(24) = \left\lfloor \frac{24}{5} \right\rfloor - \left\lfloor \frac{24}{10} \right\rfloor - \left\lfloor \frac{24}{15} \right\rfloor + \left\lfloor \frac{24}{30} \right\rfloor = 1,$$

while

$$F_{5|\{2,3\}}(35) = \left\lfloor \frac{35}{5} \right\rfloor - \left\lfloor \frac{35}{10} \right\rfloor - \left\lfloor \frac{35}{15} \right\rfloor + \left\lfloor \frac{35}{30} \right\rfloor = 3.$$

Therefore

$$\frac{F_{5|\{2,3\}}(35)}{35} = \frac{3}{35} > \frac{2}{24} = 2\frac{F_{5|\{2,3\}}(24)}{24}.$$

□

Thus the naive idea of proving pair-vs-two-tail inequalities by decomposing into singleton-vs-two-tail terms is invalid.

Proposition 7.2 (Largest-divisor singleton classes can attain the factor 2)**.** *Let*

$$d := 2, \qquad V := \{p \leq 97 : p \textit{ odd prime}\}.$$

Then with $n = 198$ *and* $m = 396$ *one has*

$$\frac{F_{2|V}(396)}{396} = 2\frac{F_{2|V}(198)}{198}.$$

Proof. If $2u \leq 198$ and no odd prime ≤ 97 divides $2u$, then the odd part u has no odd prime factor at most 97. Since $u \leq 99$, this forces $u = 1$. Thus the surviving even integers up to 198 are precisely

$$2, 4, 8, 16, 32, 64, 128,$$

so $F_{2|V}(198) = 7$.

Similarly, if $2u \leq 396$ and no odd prime ≤ 97 divides $2u$, then every odd prime factor of u exceeds 97. Because $u \leq 198$, the odd part u is either 1 or an odd prime between 101 and 197; a composite odd u with all prime factors > 97 would satisfy $u > 97^2 > 198$. Therefore the surviving even integers up to 396 are the eight powers

$$2, 4, 8, 16, 32, 64, 128, 256$$

together with the numbers $2p$ for odd primes

$$101 \leq p \leq 197.$$

There are 20 such primes, so $F_{2|V}(396) = 8 + 20 = 28$. Hence

$$\frac{F_{2|V}(396)}{396} = \frac{28}{396} = \frac{14}{198} = 2 \cdot \frac{7}{198} = 2\frac{F_{2|V}(198)}{198}.$$

□

The preceding example shows that summing largest-divisor singleton classes cannot by itself prove the strict inequality (1).

The correct surviving transport target is therefore still the pair-tail conjecture conjecture 4.8. Its first unresolved case is the pair-vs-two-tail problem.

Proposition 7.3 (Exact pair-vs-two-tail recursion)**.** *Let $2 \le a < b < c < h$ be primitive. Define*

$$\begin{aligned} P(x) &:= F_{a,b|\{c\}}(x), \\ a' &:= \frac{a}{\gcd(a,h)}, \qquad b' := \frac{b}{\gcd(b,h)}, \qquad c' := \frac{c}{\gcd(c,h)}, \\ Q(y) &:= F_{a',b'|\{c'\}}(y). \end{aligned}$$

Then for every $x \ge 1$,

$$F_{a,b|\{c,h\}}(x) = P(x) - Q\left(\left\lfloor \frac{x}{h} \right\rfloor\right).$$

Consequently, if

$$s := \left\lfloor \frac{n}{h} \right\rfloor, \qquad t := \left\lfloor \frac{m}{h} \right\rfloor,$$

then

$$\frac{F_{a,b|\{c,h\}}(m)}{m} < 2\frac{F_{a,b|\{c,h\}}(n)}{n}$$

is equivalent to

$$2\frac{P(n)}{n} - \frac{P(m)}{m} > 2\frac{Q(s)}{n} - \frac{Q(t)}{m}.$$

Proof. The first statement is Proposition 4.2 applied with $U = \{a, b\}$ and $V = \{c, h\}$. The displayed equivalence is immediate after dividing the identity by m and n at $x = m$ and $x = n$. □

Lemma 7.4 (A parent one-tail margin)**.** *Let $2 \le a < b < c$ be primitive, and put $P(x) := F_{a,b|\{c\}}(x)$. Let $E_1(n)$ be the number of $x \le n$ which are divisible by exactly one of a, b and not by c, and let $E_2(n)$ be the number of $x \le n$ divisible by both a and b and not by c. Then for every $m > n \ge c$,*

$$2\frac{P(n)}{n} - \frac{P(m)}{m} > \frac{E_1(n) - 2}{n} + \frac{E_2(n) - 2}{m}.$$

Proof. Define

$$A(x) := \left\lfloor \frac{x}{a} \right\rfloor - \left\lfloor \frac{x}{[a,c]} \right\rfloor, \qquad B(x) := \left\lfloor \frac{x}{b} \right\rfloor - \left\lfloor \frac{x}{[b,c]} \right\rfloor,$$

and

$$C(x) := \left\lfloor \frac{x}{[a,b]} \right\rfloor - \left\lfloor \frac{x}{[a,b,c]} \right\rfloor.$$

Then

$$P(x) = A(x) + B(x) - C(x).$$

Applying the first inequality of lemma 4.3 to $(d, e) = (a, c)$ and $(d, e) = (b, c)$ gives

$$\frac{A(m)}{m} < \frac{A(n) + 1}{n} + \frac{1}{m}, \qquad \frac{B(m)}{m} < \frac{B(n) + 1}{n} + \frac{1}{m}.$$

Also $C(m) \ge C(n)$, so

$$\frac{P(m)}{m} < \frac{A(n) + B(n) + 2}{n} + \frac{2 - C(n)}{m}.$$

Now $A(n) + B(n) = E_1(n) + 2E_2(n)$ and $C(n) = E_2(n)$, while $P(n) = E_1(n) + E_2(n)$. Therefore

$$\begin{aligned} 2\frac{P(n)}{n} - \frac{P(m)}{m} &> 2\frac{E_1(n) + E_2(n)}{n} - \frac{E_1(n) + 2E_2(n) + 2}{n} - \frac{2 - E_2(n)}{m} \\ &= \frac{E_1(n) - 2}{n} + \frac{E_2(n) - 2}{m}, \end{aligned}$$

as claimed. □

Proposition 7.5 (A sufficient exact-one criterion for pair-vs-two-tail)**.** *Let* $2 \le a < b < c < h$ *be primitive, and keep the notation of Proposition 7.3. Write*

$$q_h(n) := Q\left(\left\lfloor \frac{n}{h} \right\rfloor\right) = \#(h\mathbb{N} \cap ((a\mathbb{N} \cup b\mathbb{N}) \setminus c\mathbb{N}) \cap [1, n]).$$

If

$$E_1(n) \ge 2q_h(n) + 3, \tag{14}$$

then

$$\frac{F_{a,b|\{c,h\}}(m)}{m} < 2\frac{F_{a,b|\{c,h\}}(n)}{n} \qquad (m > n \ge h).$$

Proof. Set $q := q_h(n) = Q(s)$, where $s = \lfloor n/h \rfloor$ and $t = \lfloor m/h \rfloor$. If $q = 0$, then $Q(s) = 0$ and hence

$$2\frac{Q(s)}{n} - \frac{Q(t)}{m} \le 0.$$

By theorem 4.5,

$$2\frac{P(n)}{n} - \frac{P(m)}{m} > 0,$$

so Proposition 7.3 gives the conclusion.

Assume now that $q \ge 1$. Since Q is nondecreasing, $Q(t) \ge q$, and therefore

$$2\frac{Q(s)}{n} - \frac{Q(t)}{m} \le q\left(\frac{2}{n} - \frac{1}{m}\right).$$

On the other hand, by lemma 7.4 and (14),

$$2\frac{P(n)}{n} - \frac{P(m)}{m} > \frac{2q+1}{n} + \frac{E_2(n) - 2}{m}.$$

Subtracting the previous upper bound yields

$$\left(2\frac{P(n)}{n} - \frac{P(m)}{m}\right) - \left(2\frac{Q(s)}{n} - \frac{Q(t)}{m}\right) > \frac{1}{n} + \frac{E_2(n) + q - 2}{m}.$$

Since $q \ge 1$ and $E_2(n) \ge 0$, the right-hand side is positive. Now Proposition 7.3 finishes the proof. ☐

Remark 7.6 (Computational evidence for the two-tail bottleneck)**.** We found no instance, in exhaustive searches over primitive quadruples $2 \le a < b < c < h \le 60$ and all tested $n \le 250$ with $q_h(n) > 0$, in which (14) held but the conclusion of Proposition 7.5 failed. We also found no counterexample to the full pair-vs-two-tail inequality in exhaustive searches over primitive quadruples with $h \le 24$ and all $m > n \ge h$ with $m \le 100$. These computations are only experimental, but they strongly suggest that the pair-vs-two-tail problem is the correct next local target.

Remark 7.7 (The corrected frontier)**.** The singleton split route fails already for two tails by Proposition 7.1, while the pair-vs-one-tail theorem (Theorem 4.5) is fully proved. Thus the first unresolved local transport layer is the pair-vs-two-tail problem, not the pair-vs-three-tail problem. In view of Proposition 4.9, any proof of pair-vs-tail split doubling for all tails would settle the full Erdős inequality. A complementary asymptotic route for general pair-tail systems is developed in section 8.

8 Quotient tails, overlap graphs, and a finite-window criterion

We now recast the pair-tail problem in terms of quotient tails. This allows one to import local-lemma and Janson-type technology into the density problem.

8.1 Q-free counting functions and quotient tails

For a finite set $Q \subseteq \{1, 2, 3, \dots\}$ define

$$A_Q(y) := \#\{k \le y : q \nmid k \text{ for every } q \in Q\}.$$

By convention, if $1 \in Q$ then $A_Q \equiv 0$. If $1 \notin Q$, inclusion–exclusion gives

$$A_Q(y) = \sum_{S \subseteq Q} (-1)^{|S|} \left\lfloor \frac{y}{\operatorname{lcm}(S)} \right\rfloor, \qquad \operatorname{lcm}(\emptyset) := 1,$$

and hence the natural density

$$\delta(Q) := \lim_{y\to\infty} \frac{A_Q(y)}{y} = \sum_{S \subseteq Q} \frac{(-1)^{|S|}}{\operatorname{lcm}(S)}$$

exists.

For a finite tail $T \subseteq \{2, 3, \dots\}$ and $d \ge 2$, define the raw quotient system

$$\widetilde{Q}_d(T) := \left\{ \frac{t}{\gcd(t, d)} : t \in T \right\}.$$

If $1 \in \widetilde{Q}_d(T)$, set $Q_d(T) := \{1\}$. Otherwise let $Q_d(T)$ be the primitive reduction of $\widetilde{Q}_d(T)$.

Lemma 8.1 (Quotient-tail formula). *Let $d \ge 2$ and let $T \subseteq \{2, 3, \dots\}$ be finite. Then for every $x \ge 1$,*

$$F_{d|T}(x) = A_{Q_d(T)}\left(\left\lfloor \frac{x}{d} \right\rfloor \right).$$

If moreover $a < b < \min T$ and $\ell := [a, b]$, then

$$F_{a,b|T}(x) = A_{Q_a(T)}\left(\left\lfloor \frac{x}{a} \right\rfloor \right) + A_{Q_b(T)}\left(\left\lfloor \frac{x}{b} \right\rfloor \right) - A_{Q_\ell(T)}\left(\left\lfloor \frac{x}{\ell} \right\rfloor \right).$$

Consequently the pair-tail density

$$\delta_{a,b|T} := \lim_{x\to\infty} \frac{F_{a,b|T}(x)}{x}$$

exists and satisfies

$$\delta_{a,b|T} = \frac{\delta(Q_a(T))}{a} + \frac{\delta(Q_b(T))}{b} - \frac{\delta(Q_\ell(T))}{\ell}.$$

Proof. An integer counted by $F_{d|T}(x)$ has the form $dk \le x$ and satisfies

$$t \nmid dk \quad \text{for every } t \in T.$$

Equivalently,

$$\frac{t}{\gcd(t, d)} \nmid k \quad \text{for every } t \in T.$$

If $1 \in \widetilde{Q}_d(T)$, then some $t \in T$ divides d, and hence every multiple of d is excluded; both sides are identically 0. Otherwise primitive reduction does not change the forbidden-multiples set, and the first formula follows.

For the pair formula, use

$$F_{a,b|T}(x) = F_{a|T}(x) + F_{b|T}(x) - F_{\ell|T}(x),$$

where the last term removes the double count of integers divisible by both a and b. The density identity follows by dividing by x and letting $x \to \infty$. ☐

8.2 Overlap graphs and the polymer viewpoint

For finite $Q \subseteq \{2, 3, \dots\}$ define the *overlap graph* $G(Q)$ with vertex set Q, joining $q, r \in Q$ whenever $\gcd(q, r) > 1$. If $S \subseteq Q$, we write $G(Q)[S]$ for the induced subgraph.

Lemma 8.2 (Connected components factorize the lcm)**.** *Let* $1 \notin Q$ *and let* $S \subseteq Q$*. If* $S_1, \dots, S_m$ *are the vertex sets of the connected components of* $G(Q)[S]$*, then the integers* $\mathrm{lcm}(S_1), \dots, \mathrm{lcm}(S_m)$ *are pairwise coprime, and*

$$\mathrm{lcm}(S) = \prod_{i=1}^{m} \mathrm{lcm}(S_i).$$

Proof. If $i \neq j$ and some prime p divided both $\mathrm{lcm}(S_i)$ and $\mathrm{lcm}(S_j)$, then p would divide some $q \in S_i$ and some $r \in S_j$, forcing $\gcd(q, r) > 1$ and therefore an edge between S_i and S_j, a contradiction. The displayed product formula is then immediate. □

Remark 8.3 (Polymer interpretation)**.** By inclusion–exclusion and lemma 8.2, $\delta(Q)$ is the partition function of an abstract polymer gas on the connected subsets of $G(Q)$, with polymer activity

$$w(\gamma) := \frac{(-1)^{|\gamma|}}{\mathrm{lcm}(\gamma)} \qquad (\emptyset \neq \gamma \subseteq Q,\ G(Q)[\gamma] \text{ connected}).$$

This is the natural quotient-tail polymer model in the sense of Kotecký and Preiss [7]. The next two propositions extract rigorous lower and upper bounds on $\delta(Q)$ from this viewpoint.

8.3 Cluster-expansion lower bounds and Janson upper bounds

For a graph G and a weight vector $y = (y_v)_{v \in V(G)}$ with positive entries, write

$$Y_G(S; y) := \sum_{\substack{I \subseteq S \\ I \text{ independent in } G}} \prod_{v \in I} y_v \qquad (S \subseteq V(G)).$$

Also let $\Gamma_G(v)$ and $\Gamma_G^+(v) := \Gamma_G(v) \cup \{v\}$ denote the open and closed neighborhoods of v in G.

Proposition 8.4 (Cluster-expansion lower bound for Q-free density)**.** *Let* $1 \notin Q$*, and let* $G := G(Q)$ *be the overlap graph. Suppose there are positive weights* $(y_q)_{q \in Q}$ *such that*

$$\frac{1}{q} \leq \frac{y_q}{Y_G(\Gamma_G^+(q); y)} \qquad (q \in Q). \tag{15}$$

Then

$$\delta(Q) \geq \frac{1}{Y_G(Q; y)} \geq \prod_{q \in Q} (1 + y_q)^{-1}.$$

Proof. Let $L := \mathrm{lcm}(Q)$, and choose N uniformly from the residue classes modulo L. For each $q \in Q$, let E_q be the event that $q \mid N$. Then

$$\mathbb{P}(E_q) = \frac{1}{q}, \qquad \mathbb{P}\Big(\bigcap_{q \in Q} \overline{E_q}\Big) = \delta(Q).$$

Moreover, $G(Q)$ is a dependency graph for the events $\{E_q : q \in Q\}$: if q is nonadjacent to every element of $R \subseteq Q$, then $\gcd(q, \mathrm{lcm}(R)) = 1$, and the Chinese remainder theorem implies that E_q is independent of the sigma-algebra generated by $\{E_r : r \in R\}$.

Under (15), the cluster-expansion lemma in the form of Bissacot, Fernández, Procacci, and Scoppola [4] (see also Harvey and Vondrák [5]) gives

$$\mathbb{P}\Big(\bigcap_{q \in Q} \overline{E_q}\Big) \geq \frac{1}{Y_G(Q; y)}.$$

Since

$$Y_G(Q;y) \le \sum_{I\subseteq Q}\prod_{q\in I} y_q = \prod_{q\in Q}(1+y_q),$$

the second inequality follows. □

For the upper bound we use Janson's inequality. Define

$$M(Q) := \sum_{q\in Q}\frac{1}{q}, \qquad \Delta_J(Q) := \frac{1}{2}\sum_{\substack{q,r\in Q\\ q\neq r,\ \gcd(q,r)>1}}\frac{1}{\operatorname{lcm}(q,r)}.$$

Proposition 8.5 (Janson upper bound for Q-free density)**.** *Let* $Q \subseteq \{2,3,\dots\}$ *be finite. Then*

$$\delta(Q) \le \exp(-M(Q)+\Delta_J(Q)).$$

Proof. Let

$$L := \operatorname{lcm}(Q) = \prod_{i=1}^{t} p_i^{\alpha_i}.$$

Choosing N uniformly modulo L is, by the Chinese remainder theorem, equivalent to choosing independent uniform coordinates

$$(N_1,\dots,N_t) \in \prod_{i=1}^{t}\mathbb{Z}/p_i^{\alpha_i}\mathbb{Z}.$$

For each $q \in Q$, write $q = \prod_{i=1}^{t} p_i^{\beta_i(q)}$ with $0 \le \beta_i(q) \le \alpha_i$. Then the event

$$E_q := \{q \mid N\}$$

is exactly the cylinder event that

$$N_i \equiv 0 \pmod{p_i^{\beta_i(q)}} \qquad \text{for every } i \text{ with } \beta_i(q) > 0.$$

Thus the family $(E_q)_{q\in Q}$ is realized on a product probability space with independent base variables, which is exactly the setting of the standard product-space form of Janson's inequality [6]. Moreover, if $\gcd(q,r)=1$, then E_q and E_r depend on disjoint prime-power coordinates and are independent; if $\gcd(q,r)>1$, then

$$E_q \cap E_r = \{\operatorname{lcm}(q,r) \mid N\},$$

so

$$\mathbb{P}(E_q\cap E_r) = \frac{1}{\operatorname{lcm}(q,r)}.$$

Janson's inequality therefore applies and yields

$$\mathbb{P}\Big(\bigcap_{q\in Q}\overline{E_q}\Big) \le \exp(-\mu+\Delta),$$

where

$$\mu = \sum_{q\in Q}\mathbb{P}(E_q) = M(Q)$$

and

$$\Delta = \frac{1}{2}\sum_{\substack{q,r\in Q\\ q\neq r,\ \gcd(q,r)>1}}\mathbb{P}(E_q\cap E_r) = \frac{1}{2}\sum_{\substack{q,r\in Q\\ q\neq r,\ \gcd(q,r)>1}}\frac{1}{\operatorname{lcm}(q,r)} = \Delta_J(Q).$$

Since the left-hand side is exactly $\delta(Q)$, the result follows. □

8.4 A split signed transport bound and the pair-tail finite-window criterion

For finite sets $U, V \subseteq \mathbb{N}$ define

$$\mathcal{L}(U \mid V) := \{\operatorname{lcm}(S \cup T) : \emptyset \neq S \subseteq U,\ T \subseteq V\}$$

and

$$\lambda_d^{U|V} := \sum_{\substack{\emptyset \neq S \subseteq U,\ T \subseteq V \\ \operatorname{lcm}(S \cup T) = d}} (-1)^{|S|+|T|+1} \qquad (d \in \mathcal{L}(U \mid V)).$$

Set

$$\delta_{U|V} := \sum_{d \in \mathcal{L}(U|V)} \frac{\lambda_d^{U|V}}{d}, \qquad W_+(U \mid V) := \sum_{\lambda_d^{U|V} > 0} \lambda_d^{U|V}, \qquad W_-(U \mid V) := \sum_{\lambda_d^{U|V} < 0} (-\lambda_d^{U|V}).$$

Proposition 8.6 (Signed split transport)**.** *Let $U, V \subseteq \mathbb{N}$ be finite and let $m > n \geq 1$. Then*

$$\begin{aligned} \frac{F_{U|V}(m)}{m} - 2\frac{F_{U|V}(n)}{n} &\leq -\delta_{U|V} + \frac{2W_+(U \mid V)}{n} + \frac{W_-(U \mid V)}{m} \\ &\leq -\delta_{U|V} + \frac{2W_+(U \mid V) + W_-(U \mid V)}{n}. \end{aligned}$$

In particular, if $\eta > 0$ and $\delta_{U|V} \geq \eta$, then

$$n > \frac{2W_+(U \mid V) + W_-(U \mid V)}{\eta} \quad \Longrightarrow \quad \frac{F_{U|V}(m)}{m} < 2\frac{F_{U|V}(n)}{n} \qquad (m > n).$$

Proof. By Lemma 4.1,

$$F_{U|V}(x) = \sum_{d \in \mathcal{L}(U|V)} \lambda_d^{U|V} \left\lfloor \frac{x}{d} \right\rfloor = \delta_{U|V}\, x - \sum_{d \in \mathcal{L}(U|V)} \lambda_d^{U|V} \left\{ \frac{x}{d} \right\}.$$

Dividing by x and subtracting twice the n-expression from the m-expression gives

$$\frac{F_{U|V}(m)}{m} - 2\frac{F_{U|V}(n)}{n} = -\delta_{U|V} + \sum_{d \in \mathcal{L}(U|V)} \lambda_d^{U|V} \left(\frac{2}{n}\left\{\frac{n}{d}\right\} - \frac{1}{m}\left\{\frac{m}{d}\right\} \right).$$

If $\lambda_d^{U|V} > 0$, then the d-term is at most $2\lambda_d^{U|V}/n$. If $\lambda_d^{U|V} < 0$, write $\lambda_d^{U|V} = -\mu_d$ with $\mu_d > 0$; then the d-term is

$$\mu_d \left(\frac{1}{m}\left\{\frac{m}{d}\right\} - \frac{2}{n}\left\{\frac{n}{d}\right\} \right) \leq \frac{\mu_d}{m}.$$

Summing over positive and negative coefficients gives the first inequality, and the second follows from $m > n$. The final implication is immediate. ☐

Theorem 8.7 (Pair-tail finite-window criterion from quotient tails)**.** *Let $a < b < \min T$, with $T \subseteq \{2, 3, \dots\}$ finite, and set $\ell := [a, b]$. Let*

$$Q_a := Q_a(T), \qquad Q_b := Q_b(T), \qquad Q_\ell := Q_\ell(T).$$

Assume that $1 \notin Q_a \cup Q_b$, and choose positive weight vectors $y^{(a)} = (y_q^{(a)})_{q \in Q_a}$ and $y^{(b)} = (y_q^{(b)})_{q \in Q_b}$ such that

$$\frac{1}{q} \leq \frac{y_q^{(a)}}{Y_{G(Q_a)}(\Gamma^+_{G(Q_a)}(q); y^{(a)})} \qquad (q \in Q_a),$$

$$\frac{1}{q} \le \frac{y_q^{(b)}}{Y_{G(Q_b)}(\Gamma^+_{G(Q_b)}(q); y^{(b)})} \qquad (q \in Q_b).$$

Define

$$\Theta(Q_\ell) := \begin{cases} 0, & 1 \in Q_\ell, \\ \exp(-M(Q_\ell) + \Delta_J(Q_\ell)), & 1 \notin Q_\ell, \end{cases}$$

and

$$\eta_{a,b|T} := \frac{1}{a\, Y_{G(Q_a)}(Q_a; y^{(a)})} + \frac{1}{b\, Y_{G(Q_b)}(Q_b; y^{(b)})} - \frac{\Theta(Q_\ell)}{\ell}.$$

If $\eta_{a,b|T} > 0$*, then*

$$n > \frac{2W_+(\{a,b\} \mid T) + W_-(\{a,b\} \mid T)}{\eta_{a,b|T}} \quad \Longrightarrow \quad \frac{F_{a,b|T}(m)}{m} < 2\frac{F_{a,b|T}(n)}{n} \qquad (m > n).$$

Proof. By Lemma 8.1,

$$\delta_{a,b|T} = \frac{\delta(Q_a)}{a} + \frac{\delta(Q_b)}{b} - \frac{\delta(Q_\ell)}{\ell}.$$

Applying Proposition 8.4 to Q_a and Q_b, and Proposition 8.5 to Q_ℓ when $1 \notin Q_\ell$ (while $\delta(Q_\ell) = 0$ when $1 \in Q_\ell$), we obtain

$$\delta_{a,b|T} \ge \eta_{a,b|T} > 0.$$

Now apply Proposition 8.6 with $U = \{a, b\}$ and $V = T$. □

Remark 8.8 (Interpretation of the criterion)**.** Theorem 8.7 isolates a precise asymptotic obstruction. For a fixed pair-tail configuration, every failure of split doubling must lie in a finite initial window unless the positive quotient systems $Q_a(T)$ and $Q_b(T)$ evade the cluster-expansion corridor and the negative quotient system $Q_\ell(T)$ simultaneously carries unusually large density. In that sense, any infinite family of unresolved pair-tail configurations must eventually evade the cluster-expansion corridor in at least one positive quotient tail while keeping the negative quotient density large; heuristically, the remaining obstruction is a dense low-prime quotient core.

Acknowledgements

This note consolidates a sequence of exploratory arguments, reductions, counterexamples to stronger false routes, and supporting computations around Erdős problem EP-488. Several core ingredients in the present version were also formalized in Lean 4, including the singleton theorem, the singleton-vs-one-tail theorem, the dense case, and the union-bound reduction.

References

[1] P. Erdős, *Some unsolved problems*, Magyar Tud. Akad. Mat. Kutató Int. Közl. **6** (1961), 221–254.

[2] P. Erdős, *Remarks on number theory. V. Extremal problems in number theory. II*, Mat. Lapok **17** (1966), 135–155.

[3] R. K. Guy, *Unsolved Problems in Number Theory*, 3rd ed., Springer, New York, 2004.

[4] R. Bissacot, R. Fernández, A. Procacci, and B. Scoppola, *An Improvement of the Lovász Local Lemma via Cluster Expansion*, Combin. Probab. Comput. **20** (2011), no. 5, 709–719.

[5] N. J. A. Harvey and J. Vondrák, *Short proofs for generalizations of the Lovász local lemma: Shearer's condition and cluster expansion*, arXiv:1711.06797, 2017.

[6] S. Janson, *Poisson approximation for large deviations*, Random Structures Algorithms **1** (1990), no. 2, 221–229.

[7] R. Kotecký and D. Preiss, *Cluster expansion for abstract polymer models*, Comm. Math. Phys. **103** (1986), no. 3, 491–498.