hexagon:2611.00106v1

AN IMPROVED UPPER BOUND ON THE NUMBER OF REPRESENTATIONS OF AN INTEGER AS A BINOMIAL COEFFICIENT

BEN WHITMORE

ABSTRACT. Let $N(t)$ be the number of pairs of integers (n, m) with $\binom{n}{m} = t$. Singmaster conjectured that $N(t) = O(1)$, and the best bounds in the literature are $N(t) = O(\log t \log\log\log t/(\log\log t)^3)$ unconditionally, and $N(t) = O((\log t)^{2/3+\varepsilon})$ conditional on Cramér's conjecture. We show that $N(t) = O((\log t)^{1/2+o(1)})$ unconditionally. Solutions with $m \geq \exp((\log n)^{2/3+\varepsilon})$ are handled by a theorem of Matomäki, Radziwiłł, Shao, Tao and Teräväinen. For smaller m we show, using an interpolation determinant in the spirit of Bombieri and Pila, that the solutions with m in a short interval lie on a single algebraic curve of low degree in n, and we bound the number of solutions on such a curve by a zero estimate for exponential polynomials.

1. Introduction

For an integer $t \geq 2$ let

$$N(t) = \#\Big\{(n, m) \in \mathbb{N}^2 : \binom{n}{m} = t\Big\}.$$

Every pair counted here has $1 \leq m \leq n-1$ and $n \leq t$, so $N(t)$ is finite. In 1971 Singmaster [14] made the following conjecture.

Conjecture 1.1 (Singmaster's conjecture)**.** $N(t) = O(1)$.

The largest known value is $N(3003) = 8$, which comes from

$$3003 = \binom{3003}{1} = \binom{78}{2} = \binom{15}{5} = \binom{14}{6}$$

and the symmetry $\binom{n}{m} = \binom{n}{n-m}$. Singmaster [15] showed that $N(t) \geq 6$ for infinitely many t, using the solutions $n = F_{2j+2}F_{2j+3} - 1$, $m = F_{2j}F_{2j+3} - 1$ of $\binom{n+1}{m+1} = \binom{n}{m+2}$, where F_i is the i-th Fibonacci number. De Weger [5] conjectured that every equality $\binom{n}{m} = \binom{n'}{m'}$ arises from either $m = 0$, $m = 1$, symmetry or Singmaster's infinite family, with 7 exceptional cases. Blokhuis, Brouwer and de Weger [2] verified this for $t \leq 10^{60}$.

By symmetry it suffices to count solutions with $2m \leq n$. For these, $t \geq \binom{2m}{m} \geq 2^m$, so $m \leq \log t/\log 2$, and n is determined by m because $n \mapsto \binom{n}{m}$ is increasing. This gives Singmaster's bound $N(t) = O(\log t)$ [14]. It was improved to $O(\log t/\log\log t)$ by Abbott, Erdős and Hanson [1], to $O(\log t \log\log\log t/(\log\log t)^2)$ by Kane [8], and to $O(\log t \log\log\log t/(\log\log t)^3)$ by Kane [9]. The last bound remains the best unconditional bound in the literature. Abbott, Erdős and Hanson also observed that Cramér's conjecture [4] on gaps between consecutive primes implies $N(t) = O_\varepsilon((\log t)^{2/3+\varepsilon})$ for every $\varepsilon > 0$.

Matomäki, Radziwiłł, Shao, Tao and Teräväinen [12] proved that for fixed $\varepsilon > 0$ and t sufficiently large there are at most two solutions with $\exp((\log n)^{2/3+\varepsilon}) \le m \le n/2$ (Theorem 2.3 below). This reduces Conjecture 1.1 to solutions with $m < \exp((\log n)^{2/3+\varepsilon})$, roughly those with $m \le \log t/(\log\log t)^{3/2-\varepsilon}$ [12, Remark 1.5], but it gives no bound for the number of these solutions. Our main result is the following.

Theorem 1.2. *For all sufficiently large t,*

$$N(t) \le (\log t)^{\frac{1}{2}+51\frac{\log\log\log\log t}{\log\log\log t}}.$$

Theorem 1.2 follows from the next family of bounds by letting k grow slowly with t.

Theorem 1.3. *For an integer $k \ge 1$ put*

$$a_0(k) = 10\,(16k+12)^{2k+3/2}\,e^{10.75k+5.5}.$$

There is an absolute constant c_0 such that for every integer $k \ge 1$ and every integer t with

$$\log\log\log t \ge \max\left\{c_0, \tfrac{5}{2}\log\left(40(k+1)^2 a_0(k)\right)\right\}$$

we have

$$N(t) \le 50\,e^{2a_0(k)+1}\,(\log t)^{\frac{k+1}{2k+1}}.$$

The cases $k = 1$ and $k = 2$ give $N(t) \ll (\log t)^{2/3}$ and $N(t) \ll (\log t)^{3/5}$. The constants are very large: $\log a_0(1) > 30$, so the case $k = 1$ applies only when $\log\log\log t > 88$. All constants are explicit except c_0, which depends on the threshold in Theorem 2.3; that threshold is effective [12, Remark 1.7] but has not been computed.

1.1. **Formal verification.** The proofs of Theorems 1.2 and 1.3 have been formalised in the Lean 4 proof assistant [6] using Mathlib [11], conditional on the statement of Theorem 2.3. The latter theorem, due to Matomäki, Radziwiłł, Shao, Tao, Teräväinen, is included as a hypothesis and is not formalised here. The constants and inequalities in this paper match those of the formal proof, and Section 9 describes the correspondence between this paper and the formalisation. The source code is available at `https://github.com/benwh1/singmaster-bound`.

1.2. **Strategy of proof.** Fix $k \ge 1$, write $L = \log t$, and put $\theta = \frac{k+1}{2k+1}$ and $\alpha = \frac{k}{k+1}$. It suffices to count solutions with $2m \le n$, and each of them is determined by m. There are at most L^θ of them with $m \le L^\theta$. If $m \ge L/(\log L)^{7/5}$, then $\log n \le 2(\log L)^{7/5}$ while $\log m \ge \log L - \frac{7}{5}\log\log L$, so $m \ge \exp((\log n)^{7/10})$ once t is large, and Theorem 2.3 leaves at most two solutions. The main work concerns the range $L^\theta \le m \le L/(\log L)^{7/5}$. We cover a dyadic range $[M, 2M]$ by windows $|m - x_0| \le r_M$ with $r_M = (M/2)^\alpha e^{-2a_0-1}$ and show that each window contains at most L/M solutions (Lemma 7.2). This gives $O_k(LM^{-\alpha})$ solutions with $M \le m \le 2M$, and summing over $M = 2^i L^\theta$ gives $O_k(L^{1-\theta\alpha}) = O_k(L^\theta)$, which is Theorem 1.3. Taking $k \asymp \log\log\log t/\log\log\log\log t$ then gives Theorem 1.2.

The window bound has three ingredients. Write $\ell = L/x_0$, so that $\log n \approx \ell$ for solutions with m near x_0. From $n(n-1)\cdots(n-m+1) = t \cdot m!$ one expects $n \approx \frac{m-1}{2} + (t\,m!)^{1/m}$; compare [8, (3.1)] and [12, Remark 1.2]. In Section 3 we construct a function F, holomorphic on the disc $|z - x_0| \le x_0/2$, such that $|F(m) - n| \le e^{-L/8}$ for

every solution (n, m) with $|m - x_0| \le x_0/2$, such that $|F| \le 3x_0 e^{2\ell}$, and such that in the variable $u = 1/z$

$$F(1/u) = u^{-1}\left(\frac{1-u}{2} + \frac{e^{Lu+\kappa(u)}}{e}\right), \tag{1.1}$$

where κ is holomorphic and nearly constant near $u_0 = 1/x_0$.

Next let $d \ge 0$ and $D = (k+1)(d+1)$, and take D solutions (n_i, m_i) in a window. The determinant $\det[(m_i - x_0)^a n_i^b]$, with columns indexed by $0 \le a \le d$ and $0 \le b \le k$, is an integer. Replacing each n_i by $F(m_i)$ changes it by at most $\frac{1}{2}$, and the Schwarz lemma for interpolation determinants bounds the modified determinant by e^{-1}, provided the window radius is at most about $(x_0/2)^{k/(k+1)} e^{-2\ell/d}$. Hence the determinant vanishes, and all solutions in the window lie on one curve $\sum_b A_b(m) n^b = 0$ with $\deg A_b \le d$ (Section 4). This is the determinant method of Bombieri and Pila [3]. The exponent $k/(k+1)$ comes from comparing the saving $(r/R)^{D(D-1)/2}$ in the Schwarz lemma, where $R = x_0/2$, with the cost $R^{(k+1)d(d+1)/2}$ of the monomials $(m - x_0)^a$. To absorb the factor $e^{O(\ell)}$ that comes from the size of F we need $d \gg \ell/a_0$, and we take $d = \lfloor L/(a_0 M) \rfloor$.

Finally, we show that such a curve passes through few points (n, m) with $n \approx F(m)$ (Sections 5 and 6). By (1.1),

$$u^{d+k} \sum_b A_b(1/u) F(1/u)^b = \sum_{j=0}^{k} Q_j(u) E(u)^j, \qquad E(u) = e^{Lu+\kappa(u)},$$

where the Q_j are polynomials of degree at most $d + k$. After the substitution $u = u_0 + u_0 s/a_0$, the function $E(u)^j$ becomes a constant multiple of $e^{jTs} K(s)^j$, where $T = L/(a_0 x_0) \approx d$ and K is close to 1 on a disc of radius $a_0/2$. Truncating the Taylor series of K^j at degree $\lceil 2T \rceil$ produces an exponential polynomial $\sum_j R_j(s) e^{jTs}$ with $\sum_j (\deg R_j + 1) \le (4k+3)T$. Lemma 5.1, proved by Laplace inversion, bounds such an exponential polynomial and each of its components on $|s| \le 5$ in terms of its maximum on $|s| \le \frac{1}{2}$, and Jensen's formula then shows that it has at most $17(4k+3)T$ zeros near 0. The solutions only give near-zeros, so we first subtract a Lagrange interpolation polynomial, which is small because the near-zeros are well separated. Since $T \le L/(a_0 M)$ and $a_0 \ge 17(4k+3)$, each window contains at most L/M solutions.

The constant $a_0(k)$ is dictated by the last step: the truncation error $(10/a_0)^{2T}$ has to beat the factor $e^{O(kT \log k)}$ in Lemma 5.1, which forces $\log a_0 \asymp k \log k$; see (6.6). The windows then have radius $(M/2)^\alpha e^{-2a_0 - 1}$, and this is the source of the factor e^{2a_0+1} in Theorem 1.3.

Remark 1.4. Fix a real $x > 0$. The function $y \mapsto \Gamma(y+1)/\Gamma(y-x+1)$ is increasing on $[2x, \infty)$, since its logarithmic derivative $\psi(y+1) - \psi(y-x+1)$ is positive, where $\psi = \Gamma'/\Gamma$ is increasing; it also tends to infinity as $y \to \infty$. Hence, if $\Gamma(2x+1) \le t\,\Gamma(x+1)^2$, there is a unique real number $f(x) \ge 2x$ with

$$\Gamma(f(x)+1) = t\,\Gamma(x+1)\,\Gamma(f(x)-x+1),$$

that is, $\binom{f(x)}{x} = t$ with the binomial coefficient written in terms of the Gamma function. Let (n, m) be a solution with $2m \le n$. Then $m \ge 1$, since $t \ge 2$, and $\binom{2m}{m} \le \binom{n}{m} = t$, so $f(m)$ is defined and $f(m) = n$. Kane [8] continues f holomorphically to a sector. This continuation is not suitable here. Suppose that f has been continued to complex z, and write $c = f(z) - \frac{z-1}{2}$. Then $f(z) + 1 = c + \frac{z+1}{2}$ and $f(z) - z + 1 = c + \frac{1-z}{2}$,

so the defining equation becomes $\Gamma(c+\frac{z+1}{2})/\Gamma(c+\frac{1-z}{2}) = t\,\Gamma(z+1)$. Stirling's formula identifies the left side with c^z only when $\arg c$ stays away from $\pm\pi$. Taking logarithms in $c^z \approx t\,\Gamma(z+1)$ shows that $\arg c$ is approximately $\operatorname{Im}\left((L+\log\Gamma(z+1))/z\right)$, and this ranges over an interval of length $\asymp \ell$ on the discs of radius $\asymp x_0$ that the determinant method requires. The function F of Proposition 3.1 interpolates n only at the integers, and it depends on c only through the powers c^{-2j}, so no branch of $\log c$ has to be chosen.

Remark 1.5. The exponent $\frac{1}{2}$ is the limit of the method. A window cannot be longer than the range $[M, 2M]$ itself, and about $T \asymp L/(a_0 M)$ points per window is the true order for the number of zeros of exponential polynomials of the above shape in a disc of bounded radius. The bound for the range $[M, 2M]$ would therefore remain $\gg L/M$, and balancing L/M_0 against the trivial bound M_0 for $m \le M_0$ gives $M_0 = L^{1/2}$. The growth $a_0(k) = e^{\Theta(k\log k)}$ comes from the factor $(4\nu/T)^\nu$ in Lemma 5.1(b), that is, from Cauchy's estimates for the derivatives of an exponential polynomial at 0. If this factor could be removed, $a_0(k) = e^{O(k)}$ would suffice, one could take $k \asymp \log\log\log t$, and the exponent in Theorem 1.2 would become $\frac{1}{2} + O(1/\log\log\log t)$. Finally, Proposition 3.1 requires $\ell \ge 100\log x_0$, roughly $\log n \ge 100\log m$, so the argument does not reach solutions in which n is bounded by a power of m; in our proof these are covered by Theorem 2.3.

1.3. **Notation.** We write $X \ll Y$ or $X = O(Y)$ for $|X| \le CY$ with an absolute constant C, and $X \asymp Y$ for $X \ll Y \ll X$; subscripts indicate dependence of C on parameters. We use $\mathbb{N} = \{0, 1, 2, \dots\}$, write log for the natural logarithm and Log for the principal branch of the logarithm on $\mathbb{C} \setminus (-\infty, 0]$, and write $\log\Gamma(z+1)$ for the branch that is holomorphic on $\operatorname{Re} z > -1$ and real on $(-1, \infty)$. For $\operatorname{Re} z > 0$ we have $\operatorname{Re}\operatorname{Log} z = \log|z|$ and $\operatorname{Log}(1/z) = -\operatorname{Log} z$, and for $|z| \ge 1$ we have $|\operatorname{Log} z| \le \log|z| + \pi$. We write $\mathcal{D}(c, r) = \{z \in \mathbb{C} : |z-c| \le r\}$, and a function is holomorphic on a closed set if it is holomorphic on an open set containing it. The Bernoulli polynomials are defined by $\frac{we^{xw}}{e^w-1} = \sum_{p\ge 0} B_p(x)\frac{w^p}{p!}$, and $B_p = B_p(0)$.

For an integer $t \ge 2$ we write $L = \log t$ and

$$\mathcal{S}(t) = \Big\{(n, m) \in \mathbb{N}^2 : \binom{n}{m} = t,\ 2m \le n\Big\}.$$

The map $(n, m) \mapsto (n, n-m)$ sends the solutions with $2m > n$ injectively into $\mathcal{S}(t)$, so

$$N(t) \le 2|\mathcal{S}(t)|. \tag{1.2}$$

Every $(n, m) \in \mathcal{S}(t)$ has $m \ge 1$, and $n \mapsto \binom{n}{m}$ is strictly increasing on $n \ge m$, so

$$\text{an element } (n, m) \text{ of } \mathcal{S}(t) \text{ is determined by } m. \tag{1.3}$$

Finally, $n(m-i) \le m(n-i)$ for $0 \le i < m \le n$, and therefore

$$n^m \le m^m \binom{n}{m} \qquad (m \le n). \tag{1.4}$$

2. Preliminaries

Lemma 2.1 (Stirling's formula)**.** *Let* $\sigma(z) = (z+\frac{1}{2})\operatorname{Log} z - z + \frac{1}{2}\log(2\pi)$. *For* $\operatorname{Re} z > 0$ *and* $|z| \ge 1$,

$$|\log\Gamma(z+1) - \sigma(z)| \le \frac{1}{|z|}.$$

This follows from Binet's error bound [7, §5.11(ii)].

Lemma 2.2. (a) *For every integer $j \ge 1$ the polynomial*

$$P_{2j}(z) = \frac{B_{2j+1}\left(\frac{z+1}{2}\right) - B_{2j+1}\left(\frac{1-z}{2}\right)}{2j+1}$$

satisfies $P_{2j}(m) = \sum_{i=0}^{m-1} \left(i - \frac{m-1}{2}\right)^{2j}$ for every $m \in \mathbb{N}$.

(b) *If $p \in \mathbb{N}$, $y \in \mathbb{C}$ and $p \le \pi|y|$, then $|B_p(y)| \le 8|y|^p$.*

Part (a) follows from $B_p(x+1) - B_p(x) = px^{p-1}$ [7, §24.4]. Part (b) follows from $B_p(y) = \sum_i \binom{p}{i} B_i y^{p-i}$ and $|B_i| \le 4\,i!/(2\pi)^i$, which is a consequence of Euler's formula for $\zeta(2j)$ [7, §25.6].

Theorem 2.3 (Matomäki, Radziwiłł, Shao, Tao, Teräväinen [12, Theorem 1.3])**.** *Let $0 < \varepsilon < 1$. There is $t_1(\varepsilon)$ such that for every integer $t \ge t_1(\varepsilon)$ there are at most two pairs $(n, m) \in \mathbb{N}^2$ with*

$$\binom{n}{m} = t \qquad \text{and} \qquad \exp\left((\log n)^{2/3+\varepsilon}\right) \le m \le n/2.$$

3. The interpolating function

Throughout this section $t \ge 2$ and $x_0 \ge 1$ are integers. We put

$$\ell = \frac{L}{x_0}, \qquad u_0 = \frac{1}{x_0}, \qquad \Delta = \mathcal{D}\left(x_0, \frac{x_0}{2}\right),$$

$$U = \mathcal{D}\left(u_0, \frac{u_0}{2}\right), \qquad \eta_0 = \frac{2(\log x_0 + 4)}{x_0} + 2e^{-\ell/5}.$$

Proposition 3.1. *Suppose that $x_0 \ge e^{30}$ and $\ell \ge 100 \log x_0$. Then there are a function F holomorphic on Δ and a function κ holomorphic on U with the following properties.*

(a) *$|F(m) - n| \le e^{-L/8}$ for every $(n, m) \in \mathcal{S}(t)$ with $|m - x_0| \le x_0/2$.*
(b) *$|F(z)| \le 3x_0 e^{2\ell}$ for $z \in \Delta$.*
(c) *$F(1/u) = u^{-1}\left(\frac{1-u}{2} + e^{Lu+\kappa(u)}/e\right)$ for every $u \in U$ with $1/u \in \Delta$.*
(d) *$|\kappa(u) - \kappa(u_0)| \le \eta_0$ for $u \in U$, and $|\kappa(u_0)| \le 1$.*

Proof. We start from an exact identity at the integers. Let $(n, m) \in \mathcal{S}(t)$ and put $c = n - \frac{m-1}{2}$, so that $c \ge m$. The numbers $n - i$ with $0 \le i < m$ are the numbers $c - y$, where y runs over the multiset $Y = \{i - \frac{m-1}{2} : 0 \le i < m\}$, which is symmetric about 0 and satisfies $|y| < m/2 \le c/2$. Since $\binom{n}{m} = t$, Lemma 2.2(a) gives

$$L + \log m! = \sum_{y \in Y} \log(c - y) = m \log c + \frac{1}{2} \sum_{y \in Y} \log\left(1 - \frac{y^2}{c^2}\right) = m \log c - \sum_{j \ge 1} \frac{P_{2j}(m)}{2j\, c^{2j}}.$$

Let $g(z) = (L + \log \Gamma(z+1))/z$. Dividing by m and using $P_{2j}(m) \le m(m/2)^{2j}$, we obtain

(3.1)
$$\log c = g(m) + v^*, \qquad v^* = \sum_{j \ge 1} \frac{P_{2j}(m)}{2j\, m\, c^{2j}}, \qquad 0 \le \frac{P_{2j}(m)}{2j\, m\, c^{2j}} \le \frac{\rho^j}{2}, \qquad \rho = \left(\frac{m}{2c}\right)^2 \le \frac{1}{4}.$$

The series for v^* diverges when m is replaced by a non-integral complex number, so we truncate it and solve a fixed-point problem instead. Put $N_0 = \lfloor x_0/8 \rfloor$, $q = e^{-\ell/20}$,

$$\Omega = \left\{ z \in \mathbb{C} : \operatorname{Re} z > \frac{x_0}{3},\ |z| < \frac{5x_0}{2} \right\}, \qquad a_j(z) = \frac{P_{2j}(z)\, e^{-2jg(z)}}{2j\, z},$$

$$\Phi_z(w) = \sum_{j=1}^{N_0} a_j(z) e^{-2jw}.$$

The open set Ω contains Δ and the image of U under $u \mapsto 1/u$, which is the disc with diameter $[2x_0/3, 2x_0]$. Every $z \in \Omega$ satisfies $1 \le |z| \le 8 \operatorname{Re} z$. The hypotheses imply $x_0 \ge 1000$, $\ell \ge 3000$ and $q \le \min(\frac{1}{4}, \frac{1}{x_0})$.

By Lemma 2.1, for $\operatorname{Re} z > 0$ and $|z| \ge 1$,

$$\Big| \frac{\log \Gamma(z+1)}{z} - \operatorname{Log} z + 1 \Big| \le \frac{1 + \frac{1}{2}(\log|z| + \pi) + \frac{1}{2}\log(2\pi)}{|z|}, \tag{3.2}$$

and the right side is at most $\frac{1}{2}$ if $|z| \ge 20$. On Ω we have $\operatorname{Re}(L/z) = L \operatorname{Re} z/|z|^2 > \frac{4}{75}\ell \ge \ell/19$, and on Δ we have $\operatorname{Re}(L/z) \le L/|z| \le 2\ell$. Since $\operatorname{Re} \operatorname{Log} z = \log|z|$, it follows that

$$\operatorname{Re} g(z) \ge \frac{\ell}{19} + \log|z| - \frac{3}{2} \quad (z \in \Omega), \qquad \operatorname{Re} g(z) \le 2\ell + \log|z| - \frac{1}{2} \quad (z \in \Delta). \tag{3.3}$$

Let $z \in \Omega$ and $1 \le j \le N_0$. Then $|(1 \pm z)/2| > x_0/6 - \frac{1}{2}$ and $2j + 1 \le x_0/4 + 1 \le \pi(x_0/6 - \frac{1}{2})$, so Lemma 2.2(b) gives $|P_{2j}(z)| \le 16|z|^{2j+1}$. Together with (3.3) this gives $2je^{2j}|a_j(z)| \le 16e^{5-2\ell/19}$, and since $N_0 \le x_0 \le e^{\ell/100}$,

$$\sum_{j=1}^{N_0} 2j\, e^{2j} |a_j(z)| \le 16 N_0 e^{5-2\ell/19} \le q.$$

As $|\frac{d}{dw} e^{-2jw}| \le 2je^{2j}$ for $|w| \le 1$, each Φ_z maps $\mathcal{D}(0,1)$ into $\mathcal{D}(0,q)$ and is q-Lipschitz there. By the contraction mapping principle the iterates $v_0 = 0$, $v_{i+1}(z) = \Phi_z(v_i(z))$ converge uniformly on Ω. Their limit v is holomorphic on Ω and satisfies $v(z) = \Phi_z(v(z))$ and $|v| \le q$. Moreover, since $q \le \frac{1}{4}$,

$$|w - v(z)| \le 2|w - \Phi_z(w)| \qquad (z \in \Omega,\ |w| \le 1). \tag{3.4}$$

We define

$$F(z) = \frac{z-1}{2} + e^{g(z)+v(z)}, \qquad \mu(u) = u \log\Gamma(1/u+1) + \operatorname{Log} u + 1, \qquad \kappa(u) = \mu(u) + v(1/u).$$

Then F is holomorphic on $\Omega \supseteq \Delta$, and κ is holomorphic on U.

To prove (a), let $(n, m) \in \mathcal{S}(t)$ with $|m - x_0| \le x_0/2$, and let c, ρ and v^* be as in (3.1). Then $m \in \Omega$ and $0 \le v^* \le \rho \le \frac{1}{4}$. Since $e^{-2j(g(m)+v^*)} = c^{-2j}$, the number $\Phi_m(v^*)$ is the sum of the first N_0 terms of the series for v^*, so $0 \le v^* - \Phi_m(v^*) \le \rho^{N_0+1}$, and (3.4) gives $|v(m) - v^*| \le 2\rho^{N_0+1}$. As $n - F(m) = c(1 - e^{v(m)-v^*})$ and $\log c \ge g(m) \ge L/m$, we get

$$|n - F(m)| \le 4c\rho^{N_0+1} \le m^{2N_0+2} c^{-(2N_0+1)} \le m^{2N_0+2} e^{-(2N_0+1)L/m}.$$

Since $L/m \ge 2\ell/3$, $2N_0 + 1 \ge x_0/4 - 1$, $\log m \le \log x_0 + \frac{1}{2}$ and $x_0 \log x_0 \le L/100$, the right side is at most $e^{-L/8}$.

For (b), let $z \in \Delta$. By (3.3) and $|v| \le \frac{1}{4}$,

$$|F(z)| \le \frac{|z|+1}{2} + e^{2\ell + \log|z| - 1/4} \le 3x_0 e^{2\ell}.$$

For (c), let $u \in U$ with $z = 1/u \in \Delta$. Since $\operatorname{Log} u = -\operatorname{Log} z$ we have $g(z) = Lu + \mu(u) - \operatorname{Log} u - 1$, so $e^{g(z)+v(z)} = u^{-1} e^{Lu+\kappa(u)}/e$, while $(z-1)/2 = (1-u)/(2u)$.

For (d), let $u \in U$ and $z = 1/u$. Then $\mu(u) = \log\Gamma(z+1)/z - \operatorname{Log} z + 1$ and $2x_0/3 \le |z| \le 2x_0$, so (3.2) gives $|\mu(u)| \le (9 + \frac{3}{4}\log x_0)/x_0$. Since $|v| \le q \le 1/x_0$ and $\log x_0 \ge 30$, it follows that

$$|\kappa(u) - \kappa(u_0)| \le |\mu(u)| + |\mu(u_0)| + 2q \le \frac{20 + 1.5\log x_0}{x_0} \le \eta_0$$

and $|\kappa(u_0)| \le |\mu(u_0)| + q \le 1$. □

4. Solutions in a short window lie on a curve

Proposition 4.1. *Let $k \ge 1$, $d \ge 0$ and $x_0 \ge 2$ be integers, and let $\ell \ge 0$. Put $D = (k+1)(d+1)$, $R = x_0/2$, $B_F = 3x_0 e^{2\ell}$,*

$$r = R^{k/(k+1)} \exp\left(-\frac{2\ell + \log(3x_0) + 2\log D + 1}{d+1}\right),$$

$$\delta_3 = \left(2 \cdot D! \cdot D \cdot k \cdot R^{dD}(B_F+1)^{kD}\right)^{-1}.$$

Let F be holomorphic on $\mathcal{D}(x_0, R)$ with $|F| \le B_F$ there, let $0 \le \delta \le \delta_3$, and let $\mathcal{T}$ be a finite set of pairs of integers (n, m) with $|m - x_0| \le r$ and $|F(m) - n| \le \delta$. Then there are polynomials $A_0, \dots, A_k \in \mathbb{C}[X]$ of degree at most d, not all zero, such that

$$\sum_{b=0}^{k} A_b(m)\, n^b = 0 \qquad \text{for every } (n, m) \in \mathcal{T}.$$

Note that the relation holds for the integers n themselves and not only for their approximations $F(m)$.

Proof. Consider the integer matrix with rows indexed by $\mathcal{T}$, columns indexed by the pairs (a, b) with $0 \le a \le d$ and $0 \le b \le k$, and entries $(m - x_0)^a n^b$. It suffices to show that every $D \times D$ matrix formed from D of its rows, repetitions allowed, has determinant 0. The matrix then has rank less than D, so it has a nonzero kernel vector $(c_{a,b})$, and $A_b(X) = \sum_a c_{a,b}(X - x_0)^a$ has the required properties.

Fix rows $(n_1, m_1), \dots, (n_D, m_D)$, and let $\Delta_0 = \det[(m_i - x_0)^a n_i^b]$ and $\widetilde{\Delta}_0 = \det[(m_i - x_0)^a F(m_i)^b]$. We apply the Schwarz lemma for interpolation determinants (see, e.g., [10, 13]) on $\mathcal{D}(x_0, R)$ to the functions $(z - x_0)^a F(z)^b$, which are bounded by $R^a B_F^b$ there; note that $r \le R$. With

$$W = \frac{D(D-1)}{2}, \qquad S_a = \sum_{(a,b)} a = \frac{(k+1)d(d+1)}{2},$$

$$S_b = \sum_{(a,b)} b = \frac{(d+1)k(k+1)}{2}, \qquad \Lambda = \log B_F,$$

this gives, since $\log D! \le D\log D$,

$$\log|\widetilde{\Delta}_0| \le D\log D - W\log\frac{R}{r} + S_a \log R + S_b\Lambda.$$

Write $\log r = \frac{k}{k+1}\log R - \xi$ with $\xi = (\Lambda + 2\log D + 1)/(d+1)$. A direct computation gives $\frac{W}{k+1} - S_a = \frac{k(d+1)}{2} \ge 0$, so the terms involving $\log R$ contribute at most 0. Moreover

$W\xi = \frac{(k+1)(D-1)}{2}(\Lambda + 2\log D + 1)$, and $D - 1 \geq k(d+1)$, so $W\xi \geq S_b\Lambda + (D-1)(2\log D + 1)$. Hence

$$\log|\widetilde{\Delta}_0| \leq D\log D - (D-1)(2\log D + 1) \leq -1,$$

as $D \geq 2$.

Since $\delta \leq 1$ we have $|n_i| \leq B_F + 1$. The entries of both matrices are therefore bounded by $R^d(B_F+1)^k$, and corresponding entries differ by at most $R^d k(B_F+1)^{k-1}\delta$. Expanding both determinants by the Leibniz formula gives

$$|\Delta_0 - \widetilde{\Delta}_0| \leq D! \cdot D \cdot k \cdot R^{dD}(B_F+1)^{kD}\delta = \frac{\delta}{2\delta_3} \leq \frac{1}{2}.$$

Hence $|\Delta_0| \leq e^{-1} + \frac{1}{2} < 1$, and since $\Delta_0 \in \mathbb{Z}$, it vanishes. □

5. Growth of exponential polynomials

Lemma 5.1. *Let $k \geq 1$ be an integer and $T > 0$. Let $\nu_0, \dots, \nu_k \geq 1$ be integers, and put $\nu = \sum_j \nu_j$ and $\nu_{\max} = \max_j \nu_j$. Let $G_0(s) = \sum_{j=0}^k R_j(s)e^{jTs}$ with $R_j \in \mathbb{C}[s]$ and $\deg R_j < \nu_j$, and let $m_0 \geq \max_{|s|\leq 1/2}|G_0(s)|$. Put*

$$\varrho = \max(2kT, 2\nu), \qquad C_a = \nu\, 3^\nu e^{5\varrho},$$

$$C_b = 2\nu\left((k+1)2^{4k+2}\right)^{\nu_{\max}} e^{(4k+2)T}\max\left(1, (4\nu/T)^\nu\right).$$

Then

(a) $|G_0(s)| \leq C_a\, m_0$ *for* $|s| \leq 5$;
(b) $|R_j(s)e^{jTs}| \leq C_b\, m_0$ *for* $|s| \leq 4$ *and* $0 \leq j \leq k$.

Proof. Let $\chi(w) = \prod_{h=0}^k (w - hT)^{\nu_h} = \sum_{l=0}^\nu a_l w^l$, and for $0 \leq i < \nu$ let $\chi_i(w) = \sum_{l>i} a_l w^{l-i-1}$ and $y_i = G_0^{(i)}(0)$. The function G_0 satisfies the differential equation

$$\chi\left(\frac{d}{ds}\right) G_0 = 0,$$

and integration by parts shows that its Laplace transform is

$$Y(w) = \sum_{i<\nu} \frac{y_i\, \chi_i(w)}{\chi(w)}.$$

Writing $R_j(s) = \sum_p r_{jp}s^p$, the same computation identifies $Y(w)$ with $\sum_j \sum_p r_{jp}\, p!\,(w - jT)^{-p-1}$, so Cauchy's integral formula gives

$$G_0(s) = \frac{1}{2\pi i}\oint_{|w|=\varrho} e^{sw}Y(w)\,dw, \qquad R_j(s)e^{jTs} = \frac{1}{2\pi i}\oint_{|w-jT|=T/2} e^{sw}Y(w)\,dw, \tag{5.1}$$

since all poles hT of Y lie inside the first circle and only jT lies inside the second. By Cauchy's estimates on $|s| = \frac{1}{2}$ we have $|y_i| \leq i!\, 2^i m_0$.

On $|w| = \varrho$ we have $|w - hT| \geq \varrho/2$, so $|\chi(w)| \geq (\varrho/2)^\nu$. The coefficients of χ are dominated by those of $(w + kT)^\nu$, so $|\chi_i(w)| \leq \varrho^{-i-1}(3\varrho/2)^\nu$ and hence $|\chi_i/\chi| \leq 3^\nu \varrho^{-i-1}$. Using (5.1) and $|e^{sw}| \leq e^{5\varrho}$, we obtain

$$|G_0(s)| \leq 3^\nu e^{5\varrho} m_0 \sum_{i<\nu} i!\left(\frac{2}{\varrho}\right)^i \leq \nu\, 3^\nu e^{5\varrho} m_0,$$

because $i!(2/\varrho)^i \leq i!/\nu^i \leq 1$. This proves (a).

For (b), fix j and let C_j be the circle $|w - jT| = T/2$. On C_j we have $T/2 \le |w| \le (k+\frac12)T$, so $|e^{sw}| \le e^{(4k+2)T}$ for $|s| \le 4$, and $|w - hT| \ge \max(|h-j| - \frac12, \frac12)T$ for every h. Write $w^{i+1}\chi_i(w) = \chi(w) - \sum_{l\le i} a_l w^l$. The coefficients of χ are dominated by those of $\prod_h (w+hT)^{\nu_h}$, so on C_j

$$\frac{|\sum_{l\le i} a_l w^l|}{|\chi(w)|} \le \prod_{h=0}^{k} \rho_h^{\nu_h} \le \left(\prod_{h=0}^{k} \rho_h\right)^{\nu_{\max}}, \qquad \rho_h = \frac{h+j+\frac12}{\max(|h-j|-\frac12, \frac12)} \ge 1.$$

Since $\max(r-\frac12, \frac12) \ge r/2$ for every integer $r \ge 1$,

$$\prod_h \rho_h \le \frac{2^{k+1}(k+j+1)!}{j!\,j!\,(k-j)!} = 2^{k+1}(k+1)\binom{k+j+1}{j}\binom{k}{j} \le (k+1)2^{4k+2}.$$

With $\Xi = ((k+1)2^{4k+2})^{\nu_{\max}}$ this gives $|\chi_i/\chi| \le |w|^{-i-1}(1+\Xi) \le 2\Xi(2/T)^{i+1}$ on C_j, and by (5.1)

$$|R_j(s)e^{jTs}| \le 2\Xi e^{(4k+2)T} m_0 \sum_{i<\nu} i! \left(\frac{4}{T}\right)^i \le C_b m_0,$$

because $i!(4/T)^i \le (4\nu/T)^i \le \max(1, (4\nu/T)^\nu)$ for $i < \nu$. □

6. A zero estimate

Proposition 6.1. *Let $k \ge 1$ and $d \ge 0$ be integers, and let $L > 0$, $x_0 \ge 1$ and $\eta \ge 0$ be real numbers. Let $a_0 = a_0(k)$, $u_0 = 1/x_0$, $\Delta = \mathcal{D}(x_0, x_0/2)$ and $U = \mathcal{D}(u_0, u_0/2)$, and put*

$$T = \frac{L}{a_0 x_0}, \qquad \delta_5 = \exp\left(-1000(k+1)^2(T+1)\log(6x_0+6)\right).$$

Suppose that $T \ge 10(k+1)^2$, $d \le 2T$ and $k\eta \le 0.09$. Let κ be holomorphic on U with $|\kappa(u) - \kappa(u_0)| \le \eta$ for $u \in U$ and $|\kappa(u_0)| \le 1$, and let $F : \Delta \to \mathbb{C}$ satisfy $F(1/u) = u^{-1}\left(\frac{1-u}{2} + e^{Lu+\kappa(u)}/e\right)$ whenever $u \in U$ and $1/u \in \Delta$. Let $A_0, \dots, A_k \in \mathbb{C}[X]$ be polynomials of degree at most d, not all zero, and let $0 \le \delta \le \delta_5$. Let $Z \subseteq \mathcal{D}(x_0, x_0/(4a_0))$ be a finite set of points at mutual distance at least 1 such that every $z \in Z$ admits some $w_z \in \mathbb{C}$ with

$$\sum_{b=0}^{k} A_b(z) w_z^b = 0 \qquad \text{and} \qquad |F(z) - w_z| \le \delta.$$

Then $|Z| \le 17(4k+3)T$.

Proof. Put $\hat{A}_b(u) = u^d A_b(1/u)$ and

$$Q_j(u) = e^{-j} \sum_{b=j}^{k} \binom{b}{j} u^{k-b} \left(\frac{1-u}{2}\right)^{b-j} \hat{A}_b(u) \qquad (0 \le j \le k).$$

These are polynomials of degree at most $d+k$, and the binomial theorem gives, for $u \ne 0$ and $Y \in \mathbb{C}$,

$$u^{d+k} \sum_{b=0}^{k} A_b(1/u) \left(u^{-1}\left(\frac{1-u}{2} + \frac{Y}{e}\right)\right)^b = \sum_{j=0}^{k} Q_j(u) Y^j. \tag{6.1}$$

If b^* is the largest b with $A_b \ne 0$, then $Q_{b^*} = e^{-b^*} u^{k-b^*} \hat{A}_{b^*} \ne 0$.

We now rescale. Put $h = u_0/a_0$ and $u(s) = u_0 + hs$, so that $T = Lh$ and $u(s) \in U$ for $|s| \le a_0/2$. Let

$$E(u) = e^{Lu+\kappa(u)}, \qquad K(s) = e^{\kappa(u(s))-\kappa(u_0)}, \qquad q_j(s) = e^{j(Lu_0+\kappa(u_0))}Q_j(u(s)).$$

Then $\deg q_j \le d + k$, and the function

$$G(s) = \sum_{j=0}^{k} Q_j(u(s))E(u(s))^j = \sum_{j=0}^{k} q_j(s)e^{jTs}K(s)^j \tag{6.2}$$

is holomorphic on $|s| < a_0/2$. Recall that $a_0 \ge 10(16k + 12) = 40(4k + 3) \ge 280$. If $z \in \mathcal{D}(x_0, x_0/(4a_0))$, then $z \in \Delta$, $3x_0/4 \le |z| \le 2x_0$ and $|1/z - u_0| \le h/3$, so $1/z \in U$ and $s_z = (1/z - u_0)/h$ satisfies $u(s_z) = 1/z$ and $|s_z| \le \frac{1}{3}$. Moreover $\mathrm{Re}(L/z) \ge 0$ and $|\kappa(1/z)| \le |\kappa(u_0)| + \eta \le 2$, so $|E(1/z)| \ge e^{-2}$. For distinct $z, z' \in Z$ we have $s_z - s_{z'} = (z' - z)/(zz'h)$, and therefore

$$|s_z - s_{z'}| \ge \frac{1}{4x_0^2 h} = \frac{a_0}{4x_0} \ge \frac{5}{x_0}. \tag{6.3}$$

Next we replace K^j by polynomials. Put $J = \lceil 2T \rceil$ and $\delta' = \frac{1}{5}(10/a_0)^{J+1}$, let $\Pi_0 = 1$, and for $1 \le j \le k$ let Π_j be the Taylor polynomial of degree J of K^j at 0. On $\mathcal{D}(0, a_0/2)$ we have $|j(\kappa(u(s)) - \kappa(u_0))| \le k\eta \le 0.09$, so $|K^j - 1| \le \frac{1}{10}$, and Cauchy's estimates give

$$|K^j - \Pi_j| \le \delta' \quad \text{and} \quad |\Pi_j| \ge \tfrac{1}{2} \qquad \text{on } \mathcal{D}(0, 5),\ 0 \le j \le k. \tag{6.4}$$

Let $R_j = q_j\Pi_j$, $G_0(s) = \sum_j R_j(s)e^{jTs}$ and $m_0 = \max_{|s|\le 1/2} |G_0(s)|$. We apply Lemma 5.1 with $\nu_0 = d + k + 1$ and $\nu_j = d + k + J + 1$ for $j \ge 1$. Using $d \le 2T$, $2T \le J \le 2T + 1$ and $T \ge 10(k + 1)^2$ one checks that $2kT \le \nu \le (4k + 3)T$ and $\nu_{\max} \le 5T$; in particular $\varrho = 2\nu$ and $\nu \ge 80$. The function G_0 does not vanish identically, since $R_{b^*} \ne 0$, so $m_0 > 0$. The same inequalities give $4k\nu \le e^T$, $((k + 1)2^{4k+2})^{\nu_{\max}} \le e^{(17.5k+7)T}$ and $(4\nu/T)^\nu \le (16k + 12)^{(4k+3)T}$, hence

$$C_b \le 2kC_b \le \exp\left((21.5k + 10)T + (4k + 3)T \log(16k + 12)\right). \tag{6.5}$$

By the definition of a_0,

$$\log(a_0/10) = (2k + \tfrac{3}{2}) \log(16k + 12) + 10.75k + 5.5, \tag{6.6}$$

and since $J + 1 \ge 2T$, we deduce from (6.5) that $2kC_b\delta' \le e^{-T} \le \frac{1}{2}$. By Lemma 5.1(b) and (6.4), $|q_j e^{jTs}| = |R_j e^{jTs}|/|\Pi_j| \le 2C_b m_0$ on $\mathcal{D}(0, 4)$, and therefore

$$|G - G_0| = \Big|\sum_{j=1}^{k} q_j e^{jTs}(K^j - \Pi_j)\Big| \le 2kC_b\delta' m_0 \le \frac{m_0}{2} \qquad \text{on } \mathcal{D}(0, 4). \tag{6.7}$$

The points of Z give near-zeros of G: we claim that

$$|G(s_z)| \le \gamma := 2.2\,(k + 1)\,3^k \cdot 2e^3\,C_b\,m_0\,\delta \qquad (z \in Z). \tag{6.8}$$

Let $z \in Z$, $u = 1/z$ and $E = E(u)$, and define Y' by $w_z = u^{-1}(\frac{1-u}{2} + Y'/e)$. By (6.1) with $Y = Y'$ we have $\sum_j Q_j(u)Y'^j = u^{d+k}\sum_b A_b(z)w_z^b = 0$, while $G(s_z) = \sum_j Q_j(u)E^j$ by (6.2). Hence $G(s_z) = \sum_j Q_j(u)E^j\left(1 - (Y'/E)^j\right)$. By hypothesis $F(z) = u^{-1}(\frac{1-u}{2} + E/e)$, so $E - Y' = eu(F(z) - w_z)$. Since $|u| \le 2$ and $|E| \ge e^{-2}$, this gives $|Y'/E - 1| \le 2e^3\delta \le 1$, and so $|1 - (Y'/E)^j| \le 3^j \cdot 2e^3\delta$. Finally, $Q_j(u)E^j = q_j(s_z)e^{jTs_z}K(s_z)^j$ has absolute value at most $2.2\,C_b m_0$, because $|K^j| \le 1.1$. Summing over j gives (6.8).

Suppose now that $|Z| > 17(4k+3)T$. Let $N_1 = \lfloor 17(4k+3)T \rfloor + 1$, choose $Z' \subseteq Z$ with N_1 elements, and let $\mathcal{L}$ be the Lagrange interpolation polynomial of G at the points s_z, $z \in Z'$, which are distinct by (6.3). For $|s| \le 4$ each factor $(s - s_{z'})/(s_z - s_{z'})$ in $\mathcal{L}$ has absolute value at most $\frac{13}{3} \cdot \frac{x_0}{5} \le x_0$, so $|\mathcal{L}| \le N_1 \gamma x_0^{N_1}$ on $\mathcal{D}(0,4)$ by (6.8). A short computation using (6.5), $\delta \le \delta_5$, $N_1 \le 17(4k+3)T + 1$ and $T \ge 10(k+1)^2$ shows that $4N_1\gamma x_0^{N_1} \le m_0$, so that

$$|\mathcal{L}| \le \frac{m_0}{4} \qquad \text{on } \mathcal{D}(0,4). \tag{6.9}$$

Choose s_* with $|s_*| \le \frac{1}{2}$ and $|G_0(s_*)| = m_0$, and let $f = G - \mathcal{L}$. By (6.7) and (6.9), $|f - G_0| \le \frac{3}{4}m_0$ on $\mathcal{D}(0,4)$. Hence $|f(s_*)| \ge \frac{1}{4}m_0$, and by Lemma 5.1(a), $|f| \le (C_a + \frac{3}{4})m_0$ on the circle $|s - s_*| = 3$. The function f vanishes at the N_1 points s_z with $z \in Z'$, which lie in $\mathcal{D}(0, \frac{1}{3}) \subseteq \mathcal{D}(s_*, \frac{3}{2})$. By Jensen's formula,

$$N_1 \log 2 \le \log(4C_a + 3).$$

Since $C_a = \nu\, 3^\nu e^{10\nu}$ and $\nu \ge 80$, the right side is at most $17\nu \log 2$. Hence $N_1 \le 17\nu \le 17(4k+3)T < N_1$, a contradiction. □

7. Proof of Theorem 1.3

Throughout this section $k \ge 1$ is an integer, t is an integer with $\log L > 0$, $a_0 = a_0(k)$, and

$$\theta = \frac{k+1}{2k+1}, \qquad \alpha = \frac{k}{k+1},$$
$$M_0 = L^\theta, \qquad M_1 = \frac{L}{(\log L)^{7/5}}, \qquad r_M = \left(\frac{M}{2}\right)^\alpha e^{-2a_0 - 1}.$$

The main hypothesis on k and t is

$$40(k+1)^2 a_0 \le (\log L)^{2/5}. \tag{7.1}$$

Since $(k+1)^2 \ge 4$, it implies $k+1 \le (\log L)^{1/5}$ and $160a_0 \le (\log L)^{2/5}$.

Lemma 7.1. *There is an absolute constant Λ_1 such that, if $\log L \ge \Lambda_1$, then*

$$\#\{(n,m) \in \mathcal{S}(t) : m \ge M_1\} \le 2.$$

Proof. Let $(n,m) \in \mathcal{S}(t)$ with $m \ge M_1$. By (1.4) and $n \ge 2m$ we have $m \log 2 \le m(\log n - \log m) \le L$, so $\log m \le \log L + \log 2$ and $\log n \le \log m + L/m \le 2(\log L)^{7/5}$. On the other hand $\log m \ge \log L - \frac{7}{5}\log\log L$. Hence $(\log n)^{7/10} \le 2(\log L)^{49/50} \le \log m$ once $\log L$ is large, that is, $m \ge \exp((\log n)^{2/3+1/30})$. Since also $m \le n/2$, Theorem 2.3 with $\varepsilon = 1/30$ gives the claim as soon as $t \ge t_1(1/30)$. □

Lemma 7.2. *Assume that $\log L \ge 2 \cdot 10^7$ and that (7.1) holds. Let $M \in [M_0, M_1]$, and let $x_0 \in [M, 2M]$ be an integer. Then*

$$\#\{(n,m) \in \mathcal{S}(t) : |m - x_0| \le r_M\} \le \frac{L}{M}.$$

Proof. Let $\mathcal{T}$ be the set on the left, and put

$$\ell = \frac{L}{x_0}, \qquad \beta = \frac{L}{M}, \qquad d = \Big\lfloor \frac{L}{a_0 M} \Big\rfloor, \qquad D = (k+1)(d+1), \qquad T = \frac{L}{a_0 x_0}.$$

Since $\theta \ge \frac{1}{2}$ we have $L^{1/2} \le M_0 \le M \le x_0 \le 2M \le 2L$, so $\log x_0 \le \log L + \log 2$. Also $\ell \ge L/(2M_1) = (\log L)^{7/5}/2$ and $\beta \ge (\log L)^{7/5}$.

We first check that Proposition 3.1 applies. We have $x_0 \geq L^{1/2} \geq e^{30}$, and since $(\log L)^{2/5} \geq 400$,

$$\ell \geq \tfrac{1}{2}(\log L)^{7/5} \geq 200 \log L \geq 100 \log x_0.$$

Let F and κ be the functions given by Proposition 3.1.

Next we check the hypotheses of Proposition 6.1 on the parameters. Since $x_0 \in [M, 2M]$ we have $T \leq L/(a_0 M) \leq 2T$, so $d \leq 2T$ and $d + 1 > L/(a_0 M)$. By (7.1), $L/(a_0 M) \geq (\log L)^{7/5}/a_0 \geq 40(k+1)^2 \log L$, so $T \geq 10(k+1)^2$ and $d + 1 > 40(k+1)^2 \log L$. Since $x_0 \geq L^{1/2}$ and $\ell \geq \frac{1}{2} \log L$, we have $\eta_0 \leq 12L^{-1/4} + 2L^{-1/10}$, and with $k \leq (\log L)^{1/5}$ this gives $k\eta_0 \leq 0.09$.

We now compare radii. Since $(M/2)^\alpha \leq x_0/2$ and $2a_0 e^{-2a_0 - 1} \leq 1$, we have $r_M \leq x_0/(4a_0) \leq x_0/2$. Let r be the radius of Proposition 4.1 for (x_0, k, d, ℓ). Since $(M/2)^\alpha \leq (x_0/2)^\alpha$, the inequality $r_M \leq r$ follows from

$$2\ell + \log(3x_0) + 2\log D + 1 \leq (2a_0 + 1)(d + 1),$$

which holds because $\ell \leq \beta < a_0(d+1)$, while $\log(3x_0) \leq \log L + 2$, $\log D \leq k + 2 + \log L$ and $d + 1 > 40(k+1)^2 \log L$.

Next we show that the tolerances of Propositions 4.1 and 6.1 exceed $e^{-L/8}$. We first note that

$$16000(k+1)^2(19 + 4\log M) \leq \frac{M^2}{L}. \tag{7.2}$$

Indeed, the left side is at most $368000(\log L)^{7/5}$, which is at most $\exp(\frac{1}{2}(\log L)^{4/5})$ since $\log L \geq 2 \cdot 10^7$, while $M^2/L \geq L^{2\theta - 1} = L^{1/(2k+1)} \geq \exp(\frac{1}{2}(\log L)^{4/5})$ since $2k + 1 < 2(\log L)^{1/5}$. Let δ_3 be as in Proposition 4.1 for (x_0, k, d, ℓ) with $B_F = 3x_0 e^{2\ell}$. Using $D! \leq e^{D^2}$, $B_F + 1 \leq 8Me^{2\beta}$, $d \leq \beta$ and $D \leq 2(k+1)\beta$, we find

$$\log \frac{1}{\delta_3} \leq (k+1)^2 \beta^2 (19 + 4\log M) \leq \frac{\beta^2 M^2}{16000L} \leq \frac{L}{8}$$

by (7.2). Let δ_5 be as in Proposition 6.1. Since $T + 1 \leq 2\beta$ and $\log(6x_0 + 6) \leq 3 + \log M$, and since $16000(k+1)^2(3 + \log M) \leq M^2/L \leq M$ by (7.2), we have

$$\log \frac{1}{\delta_5} \leq 2000(k+1)^2 \beta (3 + \log M) \leq \frac{\beta M}{8} = \frac{L}{8}.$$

Every $(n, m) \in \mathcal{T}$ satisfies $|m - x_0| \leq r_M \leq x_0/2$, so $|F(m) - n| \leq e^{-L/8}$ by Proposition 3.1(a). If $|\mathcal{T}| < D$, then

$$|\mathcal{T}| \leq (k+1)(d+1) - 1 \leq \frac{(k+1)L}{a_0 M} + k \leq \frac{L}{2M} + \frac{L}{2M} = \frac{L}{M},$$

since $2(k+1) \leq a_0$ and $2k \leq (\log L)^{7/5} \leq L/M$. Suppose instead that $|\mathcal{T}| \geq D$. Proposition 4.1, applied with $\delta = e^{-L/8} \leq \delta_3$ and the bound $|F| \leq 3x_0 e^{2\ell}$ from Proposition 3.1(b), gives polynomials A_b of degree at most d, not all zero, with $\sum_b A_b(m) n^b = 0$ for all $(n, m) \in \mathcal{T}$. We apply Proposition 6.1 with $Z = \{m : (n, m) \in \mathcal{T}\}$, $w_m = n$, $\eta = \eta_0$ and $\delta = e^{-L/8}$. Its hypotheses hold by the above and by Proposition 3.1(c) and (d), and distinct integers are at distance at least 1. By (1.3) we have $|Z| = |\mathcal{T}|$, and since $a_0 \geq 40(4k+3)$ we conclude that

$$|\mathcal{T}| \leq 17(4k+3)T \leq 17(4k+3)\frac{L}{a_0 M} \leq \frac{L}{M}. \qquad \square$$

Proof of Theorem 1.3. Let Λ_1 be the constant of Lemma 7.1, and put $c_0 = \log\max(2\cdot 10^7, \Lambda_1)$. Suppose that $\log\log\log t \ge \max\{c_0, \frac{5}{2}\log(40(k+1)^2 a_0)\}$. Then $\log L \ge \max(2\cdot 10^7, \Lambda_1)$ and (7.1) holds, so Lemmas 7.1 and 7.2 apply. We split $\mathcal{S}(t)$ according to whether $m \le M_0$, $M_0 < m < M_1$ or $m \ge M_1$. By (1.3) there are at most M_0 elements of the first kind, and by Lemma 7.1 at most 2 of the third kind.

Let $M \in [M_0, M_1]$. We claim that

$$\#\{(n,m) \in \mathcal{S}(t) : M \le m \le 2M\} \le 6e^{2a_0+1} L M^{-\alpha}. \tag{7.3}$$

By (7.1) we have $8e^{4a_0+2} \le L^{1/2} \le M$, so $(2e^{2a_0+1})^2 \le M/2$, and since $\alpha \ge \frac{1}{2}$ this gives $r_M \ge 2$; also $r_M \le M$. Put $p = \lfloor r_M \rfloor$, so that $p \ge r_M/2$. We assign each $(n,m) \in \mathcal{S}(t)$ with $M \le m \le 2M$ to the window centred at $\lceil M \rceil + ip$ with $i = \lfloor (m - \lceil M \rceil)/p \rfloor$. This centre is an integer in $[M, 2M]$ at distance less than $p \le r_M$ from m, and at most $M/p + 1 \le 3M/r_M$ centres occur. By Lemma 7.2 the left side of (7.3) is therefore at most

$$\frac{3M}{r_M} \cdot \frac{L}{M} = 3 \cdot 2^{\alpha} e^{2a_0+1} L M^{-\alpha} \le 6e^{2a_0+1} L M^{-\alpha}.$$

Every m with $M_0 < m < M_1$ lies in $[M, 2M]$ for some $M = 2^i M_0 < M_1$ with $i \ge 0$. Since $2^{-\alpha} \le 2^{-1/2} \le \frac{5}{7}$ and $L M_0^{-\alpha} = L^{1-\theta\alpha} = L^{\theta}$, the bound (7.3) shows that there are at most

$$6e^{2a_0+1} L^{\theta} \sum_{i \ge 0} 2^{-\alpha i} \le 21 e^{2a_0+1} L^{\theta}$$

elements of the second kind. By (1.2) we conclude that

$$N(t) \le 2\left(M_0 + 21 e^{2a_0+1} L^{\theta} + 2\right) \le 50\, e^{2a_0+1} L^{\theta}. \qquad \square$$

8. Proof of Theorem 1.2

Proof of Theorem 1.2. Let c_0 be the constant of Theorem 1.3, write $L = \log t$ and $s = \log\log L = \log\log\log t$, so that $\log s = \log\log\log\log t$. We may assume that t is so large that $s \ge c_0$, $s \ge 3$, $1 \le \log s \le s/100$, $(\log L)^{2/5} \ge 3$ and $\log\log L \le \frac{1}{4}(\log L)^{3/5}$. Put $y = s/(100\log s) \ge 1$ and $k = \lfloor y \rfloor \ge 1$.

For every integer $k \ge 1$ we have

$$\log\left(40(k+1)^2 a_0(k)\right) \le 37.5\, k \log(k+2). \tag{8.1}$$

Indeed, $\log(16k+12) \le 4\log 2 + \log(k+2)$ shows that the left side is at most $16\log 2 + 5.5 + (8\log 2 + 10.75)k + (2k+3.5)\log(k+2)$, and this is at most $37.5\, k\log(k+2)$ because $\log(k+2) \ge \log 3 > 1.0974$.

Since $k \le y \le s/100$ we have $k+2 \le s$, so $k\log(k+2) \le y \log s = s/100$. By (8.1),

$$\tfrac{5}{2}\log\left(40(k+1)^2 a_0\right) \le \tfrac{5}{2} \cdot 37.5 \cdot \frac{s}{100} \le s,$$

and Theorem 1.3 gives $N(t) \le 50 e^{2a_0+1} L^{\theta}$ with $\theta = \frac{k+1}{2k+1}$. Since $y < k+1 \le 2k$ we have $s \le 200 k \log s$, and therefore

$$\theta = \frac{1}{2} + \frac{1}{2(2k+1)} \le \frac{1}{2} + \frac{1}{4k} \le \frac{1}{2} + \frac{50\log s}{s}.$$

The same estimate gives $\log(40(k+1)^2 a_0) \le \frac{2}{5} s$, that is, $a_0 \le (\log L)^{2/5}$. Since $\log 50 + 1 \le 2(\log L)^{2/5}$ and $\log\log L \le \frac{1}{4}(\log L)^{3/5}$, we have $(\log 50 + 1 + 2(\log L)^{2/5})\, s \le \log L$, and

hence

$$50e^{2a_0+1} \le \exp\left(\log 50 + 1 + 2(\log L)^{2/5}\right) \le \exp\left(\frac{\log L}{s}\right) \le \exp\left(\frac{\log L \cdot \log s}{s}\right) = L^{\log s/s}.$$

Combining these bounds,

$$N(t) \le L^{\log s/s} \cdot L^{\frac{1}{2}+50\log s/s} = (\log t)^{\frac{1}{2}+51\frac{\log\log\log\log t}{\log\log\log t}}. \qquad \square$$

9. Correspondence with the formal proof

The main theorem of the formalisation is

`Singmaster.statement_of_mrstt : MRSTT → Statement`.

Here `MRSTT` is the statement of Theorem 2.3, and `Statement` asserts that there is a real number c with

$$N(t) = O\left((\log t)^{\frac{1}{2}+c\frac{\log\log\log\log t}{\log\log\log t}}\right) \qquad (t \to \infty).$$

The formal proof derives this from Theorem 1.2, so it gives $c = 51$. Lemma 2.1 is proved for the function $\sigma + J$, where $J(z) = \sum_{n\ge 0} j(z+n)$ is Gudermann's series with $j(w) = (w+\frac{1}{2})(\mathrm{Log}(w+1) - \mathrm{Log}\, w) - 1$. This function equals $\log\Gamma(z+1)$, but the formal proof never uses the Gamma function: it shows directly that $\sigma + J$ is holomorphic on $\mathrm{Re}\, z > 0$, takes the value $\log m!$ at every $m \in \mathbb{N}$, and satisfies the bound of Lemma 2.1. The table lists the Lean declarations corresponding to the results of this paper. All names are in the namespace `Singmaster`.

Result	Lean declaration
Theorem 1.2	`MainBound`, `mainBound_of_fixedKBound`, `main_pointwise`
Theorem 1.3	`FixedKBound`, `fixedKBound_of_mrstt`
(1.2)	`N_le_two_mul_card_halfSolutions`
Lemma 2.1	`StirlingFormula`, `stirlingFormula`
Lemma 2.2	`powSumPoly_eval_nat`, `bernC_eval_norm_le`
Theorem 2.3	`MRSTT` (a hypothesis)
Proposition 3.1	`InterpolatingFunction`, `interpolatingFunction`
Contraction mapping step in Proposition 3.1	`analytic_fixed_point`
Proposition 4.1	`solutions_on_curve`
Schwarz lemma in Proposition 4.1	`interpolation_det_le`
Lemma 5.1	`expPoly_bound`
Proposition 6.1	`zero_estimate`
(6.4)	`taylor_trunc`
Jensen's formula in Proposition 6.1	`card_zeros_le_of_jensen`
Lemma 7.1	`card_halfSolutions_large_le`
Lemma 7.2	`ct_window`, `ct_delta3`, `ct_delta5`
Proof of Theorem 1.3	`ct_block`, `ct_middle`, `ct_final`
(8.1)	`log_a₀_bound`

10. Acknowledgements

All of the proofs presented in this paper, their Lean 4 formalisations, and most of the text content of this paper were generated by Claude Opus 5.5.

References

[1] H. L. Abbott, P. Erdős, and D. Hanson. On the number of times an integer occurs as a binomial coefficient. *Amer. Math. Monthly*, 81:256–261, 1974.

[2] Aart Blokhuis, Andries Brouwer, and Benne de Weger. Binomial collisions and near collisions. *Integers*, 17:Paper No. A64, 8, 2017.

[3] E. Bombieri and J. Pila. The number of integral points on arcs and ovals. *Duke Math. J.*, 59(2):337–357, 1989.

[4] Harald Cramér. On the order of magnitude of the difference between consecutive prime numbers. *Acta Arith.*, 2:23–46, 1936.

[5] Benjamin M. M. de Weger. Equal binomial coefficients: some elementary considerations. *J. Number Theory*, 63(2):373–386, 1997.

[6] Leonardo de Moura and Sebastian Ullrich. The Lean 4 theorem prover and programming language. In *Automated Deduction – CADE 28*, volume 12699 of *Lecture Notes in Comput. Sci.*, pages 625–635. Springer, Cham, 2021.

[7] F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, editors. *NIST Digital Library of Mathematical Functions*. `https://dlmf.nist.gov/`.

[8] Daniel Kane. New bounds on the number of representations of T as a binomial coefficient. *Integers*, 4:A7, 10, 2004.

[9] Daniel M. Kane. Improved bounds on the number of ways of expressing t as a binomial coefficient. *Integers*, 7:A53, 7, 2007.

[10] Michel Laurent. Sur quelques résultats récents de transcendance. *Astérisque*, 198-200:209–230, 1991.

[11] The mathlib Community. The Lean mathematical library. In *Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP 2020)*, pages 367–381. ACM, New York, 2020.

[12] Kaisa Matomäki, Maksym Radziwiłł, Xuancheng Shao, Terence Tao, and Joni Teräväinen. Singmaster's conjecture in the interior of Pascal's triangle. *Q. J. Math.*, 73(3):1137–1177, 2022.

[13] Michel Waldschmidt. *Diophantine approximation on linear algebraic groups*, volume 326 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 2000.

[14] David Singmaster. Research Problems: How Often Does an Integer Occur as a Binomial Coefficient? *Amer. Math. Monthly*, 78(4):385–386, 1971.

[15] David Singmaster. Repeated binomial coefficients and Fibonacci numbers. *Fibonacci Quart.*, 13(4):295–298, 1975.