Explicit bounds on exponential sums and the scarcity of squarefree binomial coefficients

by

Andrew Granville and Olivier Ramaré *

Introduction

The distribution of squarefree binomial coefficients.

For many years, Paul Erdös has asked intriguing questions concerning the prime divisors of binomial coefficients, and the powers to which they appear. It is evident that, if k is not too small, then $\binom{n}{k}$ must be highly composite in that it contains many prime factors and often to high powers. It is therefore of interest to enquire as to how infrequently $\binom{n}{k}$ is squarefree. One well-known conjecture, due to Erdös, is that $\binom{2n}{n}$ is not squarefree once $n > 4$. Sárközy [Sz] proved this for sufficiently large n but here we return to and solve the original question:

Theorem 1. *$\binom{2n}{n}$ is not squarefree for any $n > 4$.* †

Our proof is much like Sárközy's in that we convert the problem into one about exponential sums, but we must do a lot more work to get *explicit* upper bounds on these sums. We shall succeed in proving, via such bounds, that $\binom{2n}{n}$ is divisible by the square of some prime $> \sqrt{n}$, when $n \geq 2^{1617}$. Since $\binom{2n}{n}$ is divisible by 4 if n is not a power of 2, we need only verify that $\binom{2^{k+1}}{2^k}$ is not squarefree for each k in the range $2 < k \leq 1617$ to complete the proof of Theorem 1. In fact all such binomial coefficients are divisible by 9 except $\binom{2^7}{2^6}$ which is divisible by $5^3 11^2$, and $\binom{2^9}{2^8}$ which is divisible by $7^2 13^2$. We discuss the (easy) computer verification of this in section 2.

Erdös (B33 in [Gu]) asked for the largest n for which $\binom{2n}{n}$ is not divisible by the square of an odd prime. Erdös and Graham [EG] asked whether $\binom{2n}{n}$ is divisible by the square of arbitrarily large primes once n is sufficiently large; evidently this is answered by

* Both authors have been supported, in part, by the National Science Foundation. The first author is an Alfred P. Sloan Research Fellow.

† Velammal [Ve] has also proved this result recently.

the argument above for large n, but it is desirable to state such a result for $n < 2^{1617}$. Applying the primality testing ideas of Brillhart, Lehmer and Selfridge [BLS], we shall indicate in section 2b how the following result is proved (details of the computation will be given by Cutter [C]):

Theorem 1*. $\binom{2n}{n}$ *is divisible by the square of some prime* $\geq \sqrt{n/5}$, *for all* $n \geq 2082$.

This cannot be much improved since $\binom{4160}{2080}$ is divisible by $2^2 3^4 5^2$, but not by the square of any larger prime. Another surprising one is $\binom{1572}{786}$, which is divisible by 2^4, but not by the square of any larger prime; and is, in fact, the largest $\binom{2n}{n}$ that is not divisible by the square of an odd prime.

Recently Sander [Sa1] has proved that $\binom{n}{k}$ is not squarefree if k is "close" to $n/2$, so generalizing the idea of Theorem 1. With a slightly different approach we show that $\binom{n}{k}$ cannot be squarefree unless k or $n-k$ is very small:

Theorem 2. *There exists a constant* $\tau_1 > 0$ *such that if* n *is sufficiently large and* $\binom{n}{k}$ *is squarefree then* k *or* $n-k$ *is* $< \exp\left(\tau_1 (\log n)^{2/3} (\log\log n)^{1/3}\right)$.

The primes p in our proof, for which p^2 divides $\binom{n}{k}$, are close to either $\sqrt{k}$ or $\sqrt{n}$.

In a recent preprint Wirsing ([W], Theorem 3) proved, amongst other things, a strong quantitative version of our Theorem 2: If $n^\varepsilon < k \leq n/2$ then

$$\sum_{p^2 \mid \binom{n}{k}} \frac{\log p}{p} \sim (1 - \log 2)\, \log k.$$

Wirsing also shows that if we count with $p|$ in place of $p^2|$ then we get $\log 2$ in place of $1 - \log 2$ (see also [Sa3]).

We believe that the squarefree entries in Pascal's Triangle must be much nearer still to the edge:

Conjecture 1. *There exists a constant* $\tau_2 > 0$ *such that if* n *is sufficiently large and* $\binom{n}{k}$ *is squarefree then* k *or* $n-k$ *is* $< \tau_2 (\log n \log\log n)^2$.

If true, this is more-or-less best possible since we prove, in the other direction,

Theorem 3. *There exists a constant $\tau_3 > 0$ such that there are infinitely many pairs of integers n and k for which $\binom{n}{k}$ is squarefree, with $\tau_3 \log^2 n < k < n/2$.*

There are even some rows of Pascal's Triangle which begin with lots of squarefree entries:

Theorem 4. *There exist infinitely many integers n such that $\binom{n}{k}$ is squarefree for all $k \le \frac{1}{5}\log n$.*

From Theorem 2 it is evident that there are only finitely many rows of Pascal's Triangle in which all of the entries are squarefree. In section 2 we show that this occurs only in rows $1, 2, 3, 5, 7, 11$ and 23 (a result proved by Erdös long ago).

In the other direction we show that there are no squarefree entries, other than the '1's on either end, in a positive proportion of the rows of Pascal's Triangle. Indeed that, on average there is a constant number of squarefree entries in a row; and even that there is a 'distribution function'. Specifically we prove (answering a question in [EG, p. 72]):

Theorem 5. *The sequence of integers n, for which the nth row of Pascal's Triangle has exactly $2m+2$ squarefree entries, has asymptotic density. If we denote this density by η_m then there exists a constant $\tau_4 > 0$ for which $0 < \eta_m \ll \exp\left(-\tau_4\sqrt{m}/\log(2m)\right)$ for any $m \ge 1$.*

The key ideas to gaining such a precise understanding of the distribution of the squarefree entries in Pascal's Triangle are Theorem 2 and the following result, which we prove using Brun's method:

Theorem 6. *For any positive integer k, the sequence of integers n, for which $\binom{n}{k}$ is squarefree, has asymptotic density. We denote this density by c_k, and prove that $0 < c_k = e^{-\{\alpha+o(1)\}\sqrt{k}/\log k}$ where*

$$\alpha := \sum_{j\ge 1} \binom{2j}{j} \zeta(j+1/2) \frac{1}{2^{2j-1}} \left(1 - j\sum_{i>j}\frac{1}{i^2}\right) \approx 1.825108,$$

and $\zeta(s)$ is the Riemann zeta-function. In fact, if $N > \exp(500\alpha\sqrt{k})$ then the number of integers $n \le N$ for which $\binom{n}{k}$ is squarefree is, uniformly,

$$c_k N \left(1 + O\left(\frac{1}{k \log N}\right)\right).$$

We see that Theorem 3 follows immediately from Theorem 6. Moreover Theorem 6 provides the heuristic basis upon which we make Conjecture 1.

In order to arrive at Theorem 5 (given Theorems 2 and 6) we certainly need some result that gives us an understanding of the distribution of squarefree binomial coefficients $\binom{n}{k}$ when

$$\exp\left(\tau_1 (\log n)^{2/3} (\log\log n)^{1/3}\right) > k \gg \log^2 n.$$

To do this we shall apply the large sieve to prove:

Theorem 7. *For any given $\tau_5 > 0$, there exists a constant $\tau_6 > 0$ such that if N is sufficiently large then there are $\ll N^{1-\tau_6/\log\log N}$ pairs of integers n and k satisfying $\tau_5 \log^2 N < k \le n - \tau_5 \log^2 N$ and $N/2 \le n \le N$, for which $\binom{n}{k}$ is squarefree.*

Applying Theorem 7 with $\tau_5 < 1/(500\alpha)^2$, together with Theorem 6, implies

Corollary 1. *On average, there are approximately ten-and-two-thirds squarefree entries in a row of Pascal's triangle. More precisely, there are $\sim \tau_7 N$ squarefree binomial coefficients $\binom{n}{k}$ with $0 \le k < n \le N$, where $\tau_7 = 2\sum_{k\ge 0} c_k \approx 10.66\ldots$.*

Most binomial coefficients are divisible by the squares of many small primes. However they are also usually divisible by the squares of large primes; indeed one can modify the proof of Theorem 7 to ascertain

Corollary 1*. *For any fixed prime q, there exists a constant $\kappa_q > 0$ such that there are $\sim \kappa_q N$ binomial coefficients $\binom{n}{k}$, with $0 \le k < n \le N$, which are not divisible by the square of any prime $p > q$.*

We give a related application of our methods: Erdös, Lacampagne and Selfridge [ELS] recently conjectured that if the least prime factor of $\binom{n}{k}$ is $> k$ then n is bigger than an arbitrary power of k. This follows from

Theorem 8. *If the least prime factor of* $\binom{n}{k}$ *is* $> k$ *then there exists an absolute constant* $c > 0$ *such that*

$$n > \exp\left(c\left(\log^3 k/\log\log k\right)^{1/2}\right) .$$

Bounds on exponential sums.

This paper fills what we believe to be a lacuna in the existing literature concerning upper bounds on exponential sums. Although it has always been evident that many of the known estimates can be made explicit, it is a non-trivial problem to actually do so. In particular so that the constants involved do not render the explicit estimates useless in practical applications.

We have used the practical bounds that are needed to prove Theorem 1 as motivation for our results here, though we hope that this work will be applicable to a variety of other problems which routinely apply these or related exponential sum estimates. In particular our results here can be used to say something about the questions of estimating the number of integers free of large prime factors in short intervals (see [FL]), and of the largest prime factor of an integer in an interval (see [J]).

Our key result is

Theorem 9. *If* k *is a positive integer and* $y \le \frac{1}{5}x^{3/5}$ *then*

$$\left|\sum_{y<n\le y'} \Lambda(n)\mathrm{e}(x/n)\right| \le \frac{50}{3}y\left(\frac{x}{y^{\frac{k+3}{2}}}\right)^{\frac{1}{4(2^k-1)}}(\log 16y)^{11/4},$$

for any $y \le y' \le 2y$. *(Here, as usual,* $\Lambda(n)$ *is Von Mangoldt's funtion and* $\mathrm{e}(t) = \mathrm{e}^{2i\pi t}$.*)*

The bound in Theorem 9 is minimized when k is the smallest integer satisfying

$$1 + \frac{1}{2}(k + 2^{-k}) \ge \frac{\log x}{\log y}. \tag{1}$$

For this value of k we deduce that

$$\frac{1}{y^{1/2^{k+1}}} \ge \left(\frac{x}{y^{\frac{k+3}{2}}}\right)^{\frac{1}{(2^k-1)}} \ge \frac{1}{y^{1/2^k}}.$$

We can thus deduce

Corollary 2. *If $y \le \frac{1}{5}x^{3/5}$ and k is the smallest integer satisfying (1) then*

$$\left|\sum_{y<n\le y'} \Lambda(n)\mathrm{e}(x/n)\right| \le \frac{50}{3}y^{1-1/2^{k+3}}(\log 16y)^{11/4},$$

for any $y \le y' \le 2y$.

For larger values of y we have the following result.

Theorem 9'. *If $x \ge y \ge 2x^{2/3}$ then*

$$\left|\sum_{y<n\le y'} \Lambda(n)\mathrm{e}(x/n)\right| \le 5y\left(\frac{y}{x}\right)^{\frac{1}{4}}(\log 16y)^{5/2},$$

for any $y \le y' \le 2y$.

The contents of this paper.

We begin, in section 1, by discussing Kummer's fundamental Theorem for understanding the prime power divisors of binomial coefficients. We immediately apply this to show that Theorem 1 is true for $n < 2^{100,000}$. Next we show that the nth row of Pascal's triangle contains only squarefree integers for $n = 1, 2, 3, 5, 7, 11$ or 23, and no other n values. We start section 2 by proving a strong form of Theorem 4. We then indicate how Theorem 1^* is proved, and discuss the computations necessary for that.

In section 3 we discuss more detailed ways of applying Kummer's Theorem, in particular those that we shall use later in the paper. We also prove a non-uniform version of Theorem 6, and specify the values of the constants c_k. In section 4, we prove a uniform version of Theorem 6 using Brun's method. This implies Theorem 3 also.

In section 5, we explain how our subject is related to exponential sums and discuss the relevant results in the literature. We prove Theorems 2 and 7. We then prove the estimate for $\log c_k$, given in Theorem 6, and show how the value of α is determined. We also apply such methods to prove Theorem 8. In section 6, we complete the proof of Theorem 5.

In section 7, we indicate how Theorems 1 and 1^* for $n \ge 2^{1617}$ follow from Theorem 9. Then, in section 8, we give explicit upper bounds on exponential sums of the form

$\sum \mathrm{e}(x/n)$, and in section 9 on exponential sums of the form $\sum \mathrm{e}(x/p)$, where p is prime, so giving the proof of Theorems 9 and 9'.

Details of computations are available by email request from andrew@math.uga.edu

Acknowledgements: We would like to thank Professor J.W. Sander, as well as Professors Erdös, Montgomery, Pomerance, Sárközy, Sargos, Tenenbaum, Vaughan and the anonymous referee, for their helpful comments.

1. Kummer's Theorem and some straightforward consequences.

In 1855 Kummer observed that the power to which prime p divides the binomial coefficient $\binom{n}{m}$ is given by the number of 'carries' when one adds m and $n-m$ written in base p. We shall, henceforth, refer to this as *Kummer's Theorem.* We leave the entertaining task of proving this delightful observation to our enthusiastic reader.

A useful alternate way to state Kummer's Theorem is that the power of prime p dividing $\binom{n}{m}$ is given by the number of integers $j \geq 0$ for which $\{m/p^j\} > \{n/p^j\}$, where $\{t\}$ denotes the fractional part of t (since this is equivalent to a carry occurring in the p^{j-1} column).

1a. Theorem 1 for $n \leq 10^5$.

Any integer n in base 2 is of the form $\sum_{i=1}^{k} 2^{a_i}$ where the a_i's are distinct. Adding n to itself in base 2 we get exactly k carries, and so 2^k divides $\binom{2n}{n}$ by Kummer's Theorem. Therefore

Proposition 1.1. *If $n \geq 1$ then 4 divides $\binom{2n}{n}$, unless n is a power of 2.*

Thus we need only verify Theorem 1 where n is a power of 2, and it seems likely that 9 will divide $\binom{2^{k+1}}{2^k}$ once k is sufficiently large. We tested this for $k \leq 100,000$ by examining $2^k \pmod{3^{40}}$ with Kummer's Theorem: Write

$$2^k \equiv a[0] + a[1] * 3^1 + a[2] * 3^2 + \ldots + a[39] * 3^{39} \pmod{3^{40}}.$$

If $a[i] = a[j] = 2$ or $a[i] = 2,\ a[i+1] = 1$ then we shall have two carries when we add 2^k to itself in base 3, and so, by Kummer's theorem, 9 divides $\binom{2^{k+1}}{2^k}$, Here is a Maple program to test this:

$a[0] := 1 :$ for i from 1 to 39 do $a[i] := 0$ od :

for k from 0 to 100000 do $c := 0 :\ t := 0 :$

for i from 0 to 39 do $a[i] := 2 * a[i] + c :\ c := 0 :$

if $a[i] > 2$ then $a[i] := a[i] - 3 :\ c := 1 :\ t := t + 1 :$ fi : od :

if $t < 2$ then print(k) fi;

od :

This program ran in slightly under $75\frac{1}{2}$ minutes cpu time on a Sun 3-260. The print out was just $k = 0, 1, 2, 6$ and 8. In the latter two cases one has $5^3 11^2$ divides $\binom{2^7}{2^6}$, and $7^2 13^2$ divides $\binom{2^9}{2^8}$. This gives

Corollary 1.2. *Either 4 or 9 divides* $\binom{2n}{n}$ *for* $4 < n < 2^{100,000}$, *except in the following two cases where* 5^2 *divides* $\binom{128}{64}$, *and* 7^2 *divides* $\binom{512}{256}$.

Remarks: Goetgheluck [Go] proved this for $4 < n < 2^{42205184}$ with an almost identical algorithm. Sander [Sa4] has conjectured that 4 or 9 divides $\binom{2n}{n}$ for all n except $1, 2, 4, 64$ and 256.

1b. Rows whose entries are all squarefree.

We shall next return to the problem, raised in the introduction, of finding all those rows of Pascal's triangle whose entries are squarefree. We start by proving

Lemma 1.3. *Suppose that* p *is a prime for which* p^2 *does not divide* $\binom{n}{m}$ *for all* $0 \le m \le n$. *Then* p^{r-1} *divides* $n+1$, *where* $p^{r+1} > n \ge p^r$.

Proof: Write n in base p so that

$$n = a_r p^r + a_{r-1} p^{r-1} + \ldots + a_0,$$

where $0 \le a_i \le p-1$ and $a_r \ge 1$. If p^{r-1} does not divide $n+1$, then there exists some integer $i \le r-2$ such that $a_i \ne p-1$. Let I be the smallest such integer. Taking $m = p^r - 1$ we get carries in columns p^I and p^{I+1} when we add m and $n-m$ in base p, which gives that p^2 divides $\binom{n}{m}$, a contradiction.

Corollary 1.4. *If $\binom{n}{m}$ is squarefree for all $0 \le m \le n$ then $n = 1, 2, 3, 5, 7, 11$ or 23.*

Proof: If $n \ge 25$ then $n \ge 5^2$ and 2^r, where $2^{r+1} > n$. Therefore, by Lemma 1.3, $2^{r-1}5$ divides $n+1$. Thus $n+1 \ge 2^{r-1}5 > 5n/4$, which is impossible. If $24 \ge n \ge 9$ then 12 divides $n+1$ by Lemma 1.3, giving only the possibilities $n = 11$ and $n = 23$. By considering the power of 2 that must divide $n+1$ (because of Lemma 1.3) when $n \le 8$, we are left with the possibilities $n = 1, 2, 3, 5, 7$. From explicit computations we then get the result.

2. Further elementary consequences of Kummer's Theorem.

2a. Lots of successive squarefree binomial coefficients.

Theorem 4 follows from

Theorem 2.1. *There exist infinitely many integers n such that $\binom{n}{k}$ is squarefree for every positive integer $k \le \left(\frac{1}{4} - o(1)\right) \log n$.*

Proof: Fix integer y and let $m = \prod_{p \le y} p^{\left[\frac{\log y}{\log p}\right]+1}$; then $m = e^{\{2+o(1)\}y}$ by the prime number theorem. We shall consider the powers of primes that divide $\binom{n}{k}$, where $k \le y$ and $n \equiv -1 \pmod m$.

If p is a prime $\le y$ then $n \equiv -1 \pmod{p^{\left[\frac{\log y}{\log p}\right]+1}}$, (by definition), so that the p^jth digit of n is $p-1$, for $0 \le j \le \left[\frac{\log y}{\log p}\right]$. Therefore there cannot be any carries when we add k to $n-k$ in base p (since $k \le y \le p^{\left[\frac{\log y}{\log p}\right]+1}$); and so p does not divide $\binom{n}{k}$ by Kummer's Theorem.

If p is a prime $> y(\geq k)$ then p can divide at most one of the integers $n, n-1, \dots, n-(k-1)$. So if p^2 divides $\binom{n}{k}$ then p^2 divides $n-j$ for some non-negative integer $j \leq k-1$.

Combining the remarks in the two paragraphs immediately above we have that $\binom{n}{k}$ is squarefree for all $0 \leq k \leq y$ when $n \equiv -1 \pmod m$, provided that p^2 does not divide any of the integers $n, n-1, \dots, n-(y-1)$, for any prime $y < p \leq \sqrt{n}$.

Now, the number of integers $n \leq x$ with $n \equiv -1 \pmod m$, for which one of $n, n-1, \dots, n-(y-1)$ is divisible by the square of a prime in $(y, \sqrt{x}]$ is

$$\leq \sum_{y<p\leq\sqrt{x}} y\left(\frac{x}{mp^2}+1\right) \leq \frac{yx}{m}\sum_{p>y}\frac{1}{p^2} + y\sqrt{x} \ll \frac{x}{m\log y} + y\sqrt{x}.$$

This is less than $[(x+1)/m]$, the number of integers $n \leq x$ with $n \equiv -1 \pmod m$, once $x \gg m^2y^2$. Therefore there exists $n \leq e^{\{4+o(1)\}y}$ for which $\binom{n}{k}$ is squarefree for every positive integer $k \leq y$.

2b. Theorem 1* for $n \leq 2^{1617}$.

Using Theorem 9 we shall prove, later in this paper, that $\binom{2n}{n}$ is divisible by the square of some prime $> \sqrt{n}$, for every $n \geq 2^{1617}$.

Theorem 1* may be verified for $n \leq 2081$ by factoring each $\binom{2n}{n}$. This is most easily achieved, by induction on $n = 1, 2, \dots$, by multiplying $\binom{2(n-1)}{n-1}$ already factored, through by $2(2n-1)/n$ already factored. The values of n for which p^2 does not divide $\binom{2n}{n}$ for any prime $p \geq \sqrt{n/5}$ are 1, 2, 4, 21-22, 28-31, 36-37, 50-60, 77, 80, 110, 133-136, 143, 156-161, 170-171, 210-212, 330-331, 345-346, 368-379, 391-402, 414-420, 442-445, 529-535, 651-652, 754-756, 783-786, 902, 1045, 1653-1655, 2024-2035, 2074-2081.

To verify Theorem 1* for $2081 < n < 2^{1617}$ we shall use the following immediate consequence of Kummer's Theorem, since it guarantees a carry in the p^0 and p^1 digits when we add n to n in base p.

Lemma 2.2. *If p is a prime for which $\{n/p\}, \{n/p^2\} > 1/2$ then p^2 divides $\binom{2n}{n}$.*

Corollary 2.3. *If p is a prime for which $\{N/p\}, \{N/p^2\} > 1/2$ then p^2 divides $\binom{2n}{n}$ for each integer n in the range $N \le n \le p(1 + [N/p]) - 1$.*

Proof: If $N \le n \le p(1 + [N/p]) - 1$ then $\{n/p\} = (n - N)/p + \{N/p\} > 1/2$ and $\{n/p^2\} = \{N/p^2\} > 1/2$, so the result follows from Lemma 2.2.

We verified Theorem 1^* for $2082 \le n \le 10^{10}$ by directly using Corollary 2.3, as follows: Suppose we have already verified Theorem 1^* for $2082 \le n \le N - 1$. Let p be the largest prime $< \sqrt{2N}$. We check whether $\{N/p^2\} > 1/2$ and $\{N/p\} > 1/2$. If so then Corollary 2.3 implies that Theorem 1^* holds in a longer interval. If not then we try the next smallest prime p. We keep checking whether smaller and smaller primes p can satisfy the hypothesis of Corollary 2.3; and in each case we did find such a prime p. Once we have found such a prime, and thus a new (and longer) interval in which Theorem 1^* holds, we apply the algorithm to this new interval.

At each step this algorithm gives an interval around N of length $\ll N^{1/2}$. This is too small to allow us, in practice, to get out as far as 2^{1617}. Instead we use the following, somewhat different, consequence of Lemma 2.2 to do that:

Proposition 2.4. *If m is a positive integer for which $p = 6m + 1$, $q = 12m - 1$ and $r = 12m + 1$ are all primes then at least one of p^2, q^2 or r^2 divides $\binom{2n}{n}$, for each $n \in [96m^2 - 2m, 108m^2 + 3m - 2]$, with the one exception, namely $m = 1, n = 104$.*

Proof: We verify this directly for $m = 1$. The next value of m for which p, q and r are all prime is $m = 5$, so henceforth assume that $m \ge 5$.

For $-1 \le i \le m - 1$, consider those integers n in the interval

$$Q_i = [(9m - 1 - i)q + 6m, (9m - i)q - 1].$$

Evidently both $\{n/q\}$ and $\{n/q^2\} > 1/2$ so that q^2 divides $\binom{2n}{n}$ by Lemma 2.2.

Similarly, for $0 \le i \le m - 1$, we consider those integers n in the intervals

$$R_i = [(9m - 2 - i)r + 6m + 1, (9m - 1 - i)r - 1]$$
$$\text{and} \quad P_i = [(18m - 4 - 2i)p + 3m + 1, (18m - 3 - 2i)p - 1].$$

Since $\{n/r\}$ and $\{n/r^2\} > 1/2$ for $n \in R_i$, and $\{n/p\}$ and $\{n/p^2\} > 1/2$ for $n \in P_i$, we have that r^2 and p^2 divide $\binom{2n}{n}$, respectively, by Lemma 2.2.

The result then follows since the (consecutive) intervals

$$Q_{m-1}, R_{m-1}, P_{m-1}, Q_{m-2}, R_{m-2}, P_{m-2}, \dots, R_1, P_1, Q_0, R_0, P_0, Q_{-1}$$

cover all of the integers in $[96m^2 - 2m, 108m^2 + 3m - 2]$.

Each time we find an integer m as in Proposition 2.4, it will give us an interval around N of length $\gg N$. Thus, by using Proposition 2.4, it is now a practical computational problem to establish Theorem 1^* for all n satisfying $10^{10} < n < 2^{1617}$.

The biggest difficulty in applying Proposition 2.4 arises when the integers involved are large since then it is difficult to **prove** that p, q and r are all prime in a reasonable amount of time (in general, this is a difficult task for primes larger than 2^{1000}). However, there are relatively quick techniques to verify the primality of prime numbers of certain special forms. D.H. Lehmer, in 1928, realized that if $p-1 = FR$, where $F > p^{1/2}$ is factored, then there is, in practice, a quick way to show that p is prime. In 1975 Lehmer, with Brillhart and Selfridge [BLS], extended this so that one needs only have the factored $F > p^{1/3}$ to get a quick test; and very recently Konyagin and Pomerance [KP] have shown how to extend this to $F > p^{3/10}$.

In order to be able to apply the primality testing method of [BLS] to finding primes p, q and r as in Proposition 2.4 we need only have the factorizations of part of m and $6m - 1$. To do this we proceed as follows: For given odd integer ℓ, let k be the smallest integer for which $2^k > 5^\ell$. We select m_0 to be the least positive integer satisfying the two congruences $m_0 \equiv 0 \pmod{2^k}$ and $m_0 \equiv (5^\ell + 1)/6 \pmod{5^\ell}$. For any $m \equiv m_0 \pmod{2^k 5^\ell}$ we have that 2^k divides m and 5^ℓ divides $6m-1$. If $m \le 5^{3\ell-1}$ then the factored part of $p-1, q-1$ and $r-1$ are $> p^{1/3}, q^{1/3}$ and $r^{1/3}$, respectively; and so we can use the Brillhart-Lehmer-Selfridge test to determine whether each of p, q and r is prime. These computations, as well as various generalizations, have been performed by Pam Cutter, and will be described in detail in her paper [C].

Remark: One can prove other results that are similar to Proposition 2.4. For example, if $q, p = q+2$ and $r = q+4k$ are all prime, where k is a positive integer and $q \geq 12k+20$, then at least one of p^2, q^2 or r^2 divides $\binom{2n}{n}$, for each $n \in \left[\frac{3}{4}q^2 + (3k+2)q, \left(\frac{3}{4} + \frac{1}{8k}\right)q^2 + (3k+1)q\right]$.

3. Further applications of Kummer's Theorem.

3a. Primes near to $\sqrt{n}$.

Proposition 3.1. *Suppose that $\binom{n}{k}$ is squarefree with $1 \leq k \leq n/2$. For any prime p in the range $n-k < p^2 \leq n$ we must have*

$$\left\{\frac{n}{p}\right\} = \left\{\frac{k}{p}\right\} + \left\{\frac{n-k}{p}\right\}, \tag{3.1}$$

where $\{t\}$ is the fractional part of t.

Proof: Since $k \leq n-k < p^2$ we have $k = ap+b$ and $n-k = cp+d$ when writing in base p. However $n \geq p^2$, and so there must be a carry in the p^1 column when we add k and $n-k$. Since $\binom{n}{k}$ is squarefree, we know, by Kummer's Theorem that there can be no more than this one 'carry' when we add k and $n-k$ in base p. Thus there is no carry in the p^0 column, which implies that (3.1) holds.

As is usual, we define $\psi(t) = 0$ if t is an integer, and $\psi(t) = \{t\} - \frac{1}{2}$ otherwise. If p divides $k(n-k)n$ then it is straightforward to see that

$$\psi\left(\frac{n}{p}\right) = \psi\left(\frac{k}{p}\right) + \psi\left(\frac{n-k}{p}\right).$$

Otherwise

$$\psi\left(\frac{n}{p}\right) = \psi\left(\frac{k}{p}\right) + \psi\left(\frac{n-k}{p}\right) \pm \frac{1}{2},$$

depending only on whether there is a carry in the p^0 column when we add k and $n-k$ in base p. Specifically, from (3.1) we deduce that

$$\psi\left(\frac{n}{p}\right) = \psi\left(\frac{k}{p}\right) + \psi\left(\frac{n-k}{p}\right) + \frac{1}{2}; \tag{3.2}$$

and so, summing over all primes in this interval we obtain

Corollary 3.2. *Suppose that $\binom{n}{k}$ is squarefree with $1 \le k \le n/2$. If $\mathcal{P}$ is a set of primes p in the range $n-k < p^2 \le n$ then*

$$\left|\sum_{p\in\mathcal{P}} \psi\left(\frac{n}{p}\right)\log p\right| + \left|\sum_{p\in\mathcal{P}} \psi\left(\frac{k}{p}\right)\log p\right| + \left|\sum_{p\in\mathcal{P}} \psi\left(\frac{n-k}{p}\right)\log p\right| \ge \frac{1}{2}\sum_{\substack{p\in\mathcal{P}\\ p\not| n k(n-k)}} \log p. \tag{3.3}$$

3b. Primes near to $\sqrt{k}$.

Proposition 3.3. *Suppose that $\binom{n}{k}$ is squarefree with $1 \le k \le n/2$. Let $\mathcal{P}$ be a set of primes p in the range $\sqrt{k} < p \le \frac{10}{9}\sqrt{k}$, which do not divide $k(n-k)n$. If $\{n/p^2\} < 0.81$ then (3.2) holds. In particular*

$$\begin{aligned}\left|\sum_{p\in\mathcal{P}} \psi\left(\frac{n}{p}\right)\log p\right| + \left|\sum_{p\in\mathcal{P}} \psi\left(\frac{k}{p}\right)\log p\right| + \left|\sum_{p\in\mathcal{P}} \psi\left(\frac{n-k}{p}\right)\log p\right| \ge \\ \ge \quad \frac{1}{2}\sum_{p\in\mathcal{P}} \log p - \sum_{\substack{p\in\mathcal{P}\\ \{n/p^2\}\ge .81}} \log p.\end{aligned} \tag{3.4}$$

Proof: If $\{n/p^2\} < 0.81 \le \{k/p^2\}$ then there is a carry in the p^1 column when we add k and $n-k$ in base p. Since $\binom{n}{k}$ is squarefree, we know, by Kummer's Theorem that there can be no more than this one 'carry' when we add k and $n-k$ in base p. Thus there is no carry in the p^0 column, which implies that (3.1) and consequently (3.2) holds. Summing this result together for all $p \in P$, and taking into account the remarks in section 3a, we deduce (3.4).

3c. How often does p^2 divide $\binom{n}{k}$?

First consider primes $p > k$. Evidently p^2 divides $\binom{n}{k}$ if and only if p^2 divides $n-j$ for some integer j, $0 \le j \le k-1$. Therefore the proportion of integers n for p^2 does not divide $\binom{n}{k}$ is

$$c_{k,p} := 1 - k/p^2, \quad \text{for primes } p > k.$$

Now consider primes $p \le k$. Write k in base p, say as $k = a_0 + a_1 p + \ldots + a_\ell p^\ell$. If p^2 does not divide $\binom{n}{k}$ then there can be no more than one carry when we add k to $n-k$

in base p. If $n \equiv n_0 + n_1 p + \ldots + n_{\ell+1} p^{\ell+1} \pmod{p^{\ell+2}}$ then either each $n_i \geq a_i$, or there exists $n_j < a_j$ with $n_{j+1} \geq a_{j+1} + 1$ and otherwise $n_i \geq a_i$. Thus the proportion of integers n for which p^2 does not divide $\binom{n}{k}$ is

$$c_{k,p} := \prod_{i=0}^{\ell} \left(1 - \frac{a_i}{p}\right) \left\{1 + \sum_{j=0}^{\ell} \frac{a_j(p-1-a_{j+1})}{(p-a_j)(p-a_{j+1})}\right\}.$$

By an application of the combinatorial sieve we deduce the first part of Theorem 6:

Proposition 3.4. *The number of integers $n \leq N$ for which $\binom{n}{k}$ is squarefree is $\sim c_k N$ as $N \to \infty$, where $c_k = \prod_p c_{k,p}$, and the $c_{k,p}$ are as defined above. As examples, $c_1 = 6/\pi^2$ and $c_2 = \frac{3}{4} \prod_{p \geq 3} \left(1 - \frac{2}{p^2}\right)$.*

In the next section we shall use some deeper sieve theory, and develop the ideas here, to prove the first part of Theorem 6.

The c_k may be computed with any desired required accuracy. For $k = 1, 2, 3, \ldots, 50$, the values of c_k are (to three significant digits):

k:	1	2	3	4	5	6	7	8	9	0
00:	.608,	.484,	.251,	.36,	.191,	.189,	.0625,	.106,	.204,	.216
10:	.0772,	.11,	.0477,	.0255,	.0271,	.11,	.0282,	.0872,	.0219,	.0656
20:	.0592,	.0275,	.00533,	.00716,	.0153,	.00696,	.0271,	.0533,	.0226,	.0222
30:	.00649,	.0517,	.0476,	.0232,	.0185,	.0937,	.0331,	.0145,	.00694,	.026
40:	.00605,	.0213,	.00432,	.00276,	.0123,	.00441,	.000695,	.00158,	.0027,	.0102

From our computations $\sum_{0 \leq k \leq 5000} c_k \approx 5.3275$, with an error term bounded by .012, which leads to the value of τ_7 given in Corollary 1. In section 5d we shall prove the asymptotic formula for $\log(c_k)$ stated in Theorem 6.

4. Squarefree $\binom{n}{k}$, with k fixed.

Proof of the last part of Theorem 6: We shall use Brun's method: Let $z = k^2 \log x$, and let a_n be the product of those primes $p \le z$ for which p^2 divides $\binom{n}{k}$. Let D be the product of all of the primes $\le z$. From section 3b, we know that p does not divide a_n if and only if n belongs to one of $c_{k,p}p^{\ell+2}$ residue classes $\pmod{p^{\ell+2}}$. Thus, letting $w(p) = 1 - c_{k,p}$ and $W(p) = w(p)p^{\ell+2}$ be multiplicative functions, we have that the number of $n \le x$ for which a_n is divisible by d is $w(d)x + O(W(d))$, for $d|D$. So, by the inclusion-exclusion principle, we have for any $I \ge 0$,

$$\begin{aligned}
\sum_{n\le x,\ a_n=1} 1 &\ge \sum_{i=0}^{2I+1} (-1)^i \sum_{\substack{d|D\\ \Omega(d)=i}} \sum_{\substack{n\le x\\ d|a_n}} 1 \\
&\ge \sum_{i=0}^{2I+1} (-1)^i \sum_{\substack{d|D\\ \Omega(d)=i}} w(d)x - O\left(\sum_{i=0}^{2I+1} \sum_{\substack{d|D\\ \Omega(d)=i}} W(d)\right) \\
&\ge x\left(\prod_{p\le z}(1-w(p)) - \sum_{i>2I+1} \sum_{\substack{d|D\\ \Omega(d)=i}} w(d)\right) - O\left(\sum_{i=0}^{2I+1} \frac{1}{i!}\left(\sum_{p\le z} W(p)\right)^i\right) \\
&\ge x\left(e^{-C} - \sum_{i>2I+1} \frac{C^i}{i!}\right) - O\left(\sum_{i=0}^{2I+1} \frac{z^{4i}}{i!}\right)
\end{aligned}$$

since $W(p) \le kp^2 \le z^3$ for all p, and

$$\sum_{p\le z} w(p) \le C := -\sum_{p\le z} \log\left(1 - w(p)\right)$$

so that

$$\sum_{\substack{d|D\\ \Omega(d)=i}} w(d) \le \frac{1}{i!}\left(\sum_{p\le z} w(p)\right)^i.$$

Selecting $I = [\log x/16\log z]$ we thus get

$$\sum_{\substack{n\le x\\ a_n=1}} 1 \ge xe^{-C}(1 - O(e^{-I})) - O(x^{1/2}).$$

An analogous argument gives an upper bound of the same size, so we have proved

$$\sum_{n\le x,\ a_n=1} 1 = xe^{-C}(1+O(e^{-I}))+O(x^{1/2}),$$

in our range. However we have yet to take account of those primes p bigger than z whose squares divide $\binom{n}{k}$. In each case this happens for exactly k residue classes $\pmod{p^2}$. If $p\le x^{1/4}$ then we split up the values of n according to their residue class $\pmod{p^2}$, and now consider a_n as above to get an upper bound on the remaining n for which $\binom{n}{k}$ is divisible by p^2. Thus

$$\sum_{n\le x,\ a_n=1,\ p|\binom{n}{k}} 1 \ll k\frac{x}{p^2}e^{-C}+x^{1/2}.$$

If $x^{1/4}<p\le x^{1/2}$ then there are evidently $\ll k\frac{x}{p^2}$ such n. Combining all of these estimates, we find that the number of squarefree $\binom{n}{k}$ with $n\le x$ is

$$xe^{-C}\left(1+O\left(e^{-I}+\frac{k}{z\log z}\right)\right)+O(kx^{3/4}).$$

Finally note that

$$e^{-C}=c_{k,p}\prod_{p>z}(1-k/p^2)^{-1}=c_{k,p}\left(1+O\left(\frac{k}{z\log z}\right)\right),$$

and the result follows.

5. How exponential sums get involved.

5a. Estimates in the literature.

To estimate exponential sums involving primes one usually writes the characteristic function of these primes as a linear combination of suitably chosen bilinear forms. In 1974 Jutila [J] did this for exponential sums involving reciprocals of primes. His results have since been improved by Sander who used Vaughan's identity (which was discovered since Jutila's paper) instead of the more complicated technique of Vinogradov. Sander has also shown how to consider reciprocals of powers of prime p. We shall apply, from [Sa2],

Lemma 5.1. *Fix $\varepsilon > 0$ and integer $J \ge 1$. There exists a constant $c > 0$ such that for any $y \le x^{1/J}$, there are*

$$\sigma_1\sigma_2 \dots \sigma_J \pi(y) + O\left(\left(y^{1-c(\log y/\log x)^2} + y^{J/2+1+\varepsilon}x^{-1/2}\right)(\log x)^{4J}\right)$$

primes $p \le y$ for which $\{x/p^j\} < \sigma_j$ for $j = 1, 2, \dots, J$.

By partial summation one can deduce (with $J = 1$) that

$$\left|\sum_{p<y} \psi\left(\frac{x}{p}\right)\log p\right| \ll \left(y^{1-c(\log y/\log x)^2} + y^{3/2+\varepsilon}x^{-1/2}\right)\log^5 x \tag{5.1}$$

5b. No squarefree binomial coefficients near the center of Pascal's Triangle.

Proof of Theorem 2: Fix $\delta > 0$ sufficiently small.

We begin by proving Theorem 2 for $n/2 \ge k \ge n^{1-\delta}$ for sufficiently large n. First note that

$$\sum_{p\in\mathcal{P}} \log p \ge \sum_{n^{1/2}-k/3n^{1/2}\le p\le n^{1/2}} \log p - \log\left(nk(n-k)\right) \gg k/n^{1/2+\varepsilon},$$

by Hoheisel's Theorem (where $\mathcal{P}$ is as in Corollary 3.2). Inserting this and (5.1) into (3.3) we get

$$k \ll n^{\varepsilon}\left(n^{1-c/8} + n^{3/4} + n^{1-(c/8)(\log n/\log k)^2} + n^{5/4}/k^{1/2}\right),$$

which is false for $k \ge n^{1-\delta}$ and n sufficiently large. Thus by Corollary 3.2 we know that $\binom{n}{k}$ cannot be squarefree.

Next we prove Theorem 2 for $n^{1-\varepsilon} \ge k \ge \exp\left(\tau_1(\log n)^{2/3}(\log\log n)^{1/3}\right)$.

We shall use Proposition 3.3 assuming that $\binom{n}{k}$ is squarefree. Taking $J = 2, \sigma_1 = 1, \sigma_2 = .81, x = n$ and both $y = \sqrt{k}$ and $y = \frac{10}{9}\sqrt{k}$ in Lemma 5.1, we have that the right side of (3.4) is

$$\left(\frac{31}{900} + O\left(\frac{1}{\log k}\right)\right)\sqrt{k} + O\left(\left(k^{1/2-(c/8)(\log k/\log n)^2} + k^{1+\varepsilon}n^{-1/2} + 1\right)\log^8 n\right)$$

using the prime number theorem. Meanwhile, using Lemma 5.1 we get the upper bound

$$\ll \left(k^{1/2-(c/8)(\log k/\log n)^2} + k^{1/4+\varepsilon}\right) \log^5 n$$

for the left side of (3.4). Combining these two estimates we have a contradiction in the range indicated, and so $\binom{n}{k}$ is not squarefree.

By taking $\varepsilon < \delta$ we can combine the results above and deduce Theorem 2.

5c. Where there are few squarefree binomial coefficients.

In this section we look to bound the number of squarefree binomial coefficients in the range for k between those given by Theorems 2 and 6; that is we assume that

$$\tau_5 \log^2 n \le k \le \exp\left(\tau_1 (\log n)^{2/3} (\log\log n)^{1/3}\right).$$

Let $\mathcal{P}$ now be the set of primes > 5 in the interval $\left(\sqrt{k}, \frac{10}{9}\sqrt{k}\right]$ for which $\{k/p\} \ge 2/3$. By Lemma 5.1 we find that $|\mathcal{P}| \sim (2/27)\sqrt{k}/\log k$, using the prime number theorem. Note that $\{k/p^2\} \le 81/100$ for every such prime p.

Let Ω_p denote those residue classes $m \pmod{p^2}$ for which $\{m/p\} < 2/3$ and $\{m/p^2\} < 81/100$. Thus $|\Omega_p| \sim 27p^2/50$. If $n \equiv m \pmod{p^2}$ for some $m \in \Omega_p$ then $\{k/p\} > \{m/p\} = \{n/p\}$ and $\{k/p^2\} > \{m/p^2\} = \{n/p^2\}$, so that p^2 divides $\binom{n}{k}$ by Kummer's Theorem.

We wish to get a good upper bound on the number of $n \le x$ for which $\binom{n}{k}$ is squarefree. From what we have written above this means that $n \notin \Omega_p \pmod{p^2}$ for all $p \in \mathcal{P}$. We may thus apply sieve methods. For small values of k, that is $\tau_5 \log^2 x \le k \le \log^{100} x$, we shall use the following trivial method:

Fix constant $\tau_8 > 0$ sufficiently small and let D be the product of $\tau_8 \log x/\log\log x$ distinct primes from $\mathcal{P}$. Evidently τ_8 must be chosen sufficiently small so that this size subset exists, and also so that $D < x^{1/3}$. For $m = 1, 2, \dots, D^2$ consider those integers $n \le x$ which are $\equiv m \pmod{D^2}$. Evidently if $m \in \Omega_p \pmod{p^2}$ for some p dividing D then $\binom{n}{k}$ is not squarefree. By the Chinese Remainder Theorem we see that the number

of $m \pmod{D^2}$ for which $m \notin \Omega_p \pmod{p^2}$ for all p dividing D is $\prod_{p|D}(p^2 - |\Omega_p|)$. Therefore the number of squarefree $\binom{n}{k}$ with $n \le x$ is

$$\ll \prod_{p|D}(p^2 - |\Omega_p|)\left(\frac{x}{D^2} + O(1)\right) \ll \frac{D^2}{2^{\tau_8 \log x/\log\log x}}\frac{x}{D^2} \ll \frac{x}{e^{2\tau_6 \log x/\log\log x}}.$$

Therefore we have proved

Lemma 5.2. *There are $\ll x\exp(-\tau_6 \log x/\log\log x)$ squarefree binomial coefficients $\binom{n}{k}$ with $n \le x$ and $\tau_5 \log^2 x \le k \le \log^{100} x$, for sufficiently large x.*

Now consider k in the range $\log^{100} x \le k \le x^{1/5}$. We shall use the arithmetic form of the large sieve, though with squares of primes rather than with primes (the proof of Théorème 6, given in [Bo], can be modified to allow one to sieve with any set of pairwise coprime integers, rather than with just primes – see also [Ga]). Note that $|\Omega_p|/(p^2 - |\Omega_p|) \sim 27/23$ and is therefore ≥ 1 if x is large enough and $p \in \mathcal{P}$. Therefore the number of $n \le x$ for which $\binom{n}{k}$ is squarefree is $\le (x + z^2)/G(z)$ where

$$G(z) = \sum_{\substack{d \le \sqrt{z},\\ p|d \Rightarrow p \in \mathcal{P}}} \mu^2(d) \prod_{p|d} \frac{|\Omega_p|}{p^2 - |\Omega_p|} \ge \sum_{\substack{d \le \sqrt{z},\\ p|d \Rightarrow p \in \mathcal{P}}} \mu^2(d).$$

Now let $z = \sqrt{x}$ and $v = \left[\frac{\log(\sqrt{z})}{\log(\frac{10}{9}\sqrt{k})}\right] \sim \frac{\log x}{2\log k}$. Since $v \le k^{1/100}$ thus

$$G(z) \ge \binom{|\mathcal{P}|}{v} \ge \left(\frac{|\mathcal{P}|}{v} - 1\right)^v \ge k^{(49/100 + o(1))v} \ge x^{49/200 + o(1)},$$

so that the number of $n \le x$ for which $\binom{n}{k}$ is squarefree is $< x^{151/200 + o(1)} < x^{19/25}$ if x is large enough. This implies

Lemma 5.3. *There are $< x^{24/25}$ squarefree binomial coefficients $\binom{n}{k}$ with $n \le x$ and $\log^{100} x \le k \le x^{1/5}$, for sufficiently large x.*

Combining the last two lemmas with Theorem 2 gives Theorem 7.

5d. An asymptotic formula for $\log(c_k)$.

We shall study the size of the $c_{k,p}$ as defined in section 3c. Fix $\varepsilon > 0$. The first thing to note is that each one is a positive rational number with denominator $p^{\ell+2} \le kp^2$. Therefore

$$\left| \sum_{p<\varepsilon\sqrt{k}/\log k} \log c_{k,p} \right| \ll \frac{\varepsilon\sqrt{k}}{\log k} .$$

Suppose that $k \ge p > \varepsilon^{-1}\sqrt{k}$, so that if $k = ap + b$ then $1 \le a < \varepsilon^2 p$ and $0 \le b \le p-1$. Then, taking $b < p$ in the expression for $c_{k,p}$, we find that $c_{k,p} > (1 - (a+1)/p) > (1 - 2k/p^2)$. Combining this lower bound with $c_{k,p} = 1 - k/p^2$ when $p > k$, we have

$$\left| \sum_{p>\varepsilon^{-1}\sqrt{k}} \log c_{k,p} \right| \ll \frac{\varepsilon\sqrt{k}}{\log k} .$$

So all remaining primes p lie in the interval $J = [\varepsilon^{-1}\sqrt{k}, \varepsilon\sqrt{k}/\log k]$, and we can write $k = dp^2 + ap + b$ where $0 \le d \ll \log^2 k$, and $a > \varepsilon^2 p$ if $d = 0$.

Now, if $b/p > 1 - \varepsilon^2/\log^2 k$ then p divides $k+i$ for some $i \le p\varepsilon^2/\log^2 k \le \varepsilon\sqrt{k}/\log^2 k$. However as all of the primes in J are $> (k+i)^{1/3}$ we see that each such $k+i$ can have no more than two such prime factors. Thus

$$\left| \sum_{b/p>1-\varepsilon^2/\log^2 k} \log c_{k,p} \right| \ll \frac{\varepsilon\sqrt{k}}{\log k} ,$$

where the sum here is only over primes $p \in J$.

If, for a given d, we have $a/p > 1 - \varepsilon/\log^2 k$ then prime p lies in an interval of length $\ll \sqrt{k/(d+1)}(\varepsilon/(d+1)\log k)$, immediately above $\sqrt{k/(d+1)}$. Thus

$$\left| \sum_{a/p>1-\varepsilon/\log^2 k} \log c_{k,p} \right| \ll \frac{\varepsilon\sqrt{k}}{\log k} \sum_{d\ge 0} \frac{1}{(d+1)^{3/2}} \ll \frac{\varepsilon\sqrt{k}}{\log k} .$$

Thus we may also assume henceforth that $1-b/p,\ 1-a/p > \varepsilon^2/\log^2 k$. Inserting these assumptions on a and b, as well as the range for d, into the definition of $c_{k,p}$ we get that

$$\begin{aligned} c_{k,p} &= \left(1-\frac{a}{p}\right)\left(1-\frac{b}{p}\right)\left(1+\frac{a}{p-a}+\frac{b}{p-b}\right)\left(1+O\left(\frac{\varepsilon^{-2}\log^2 k}{p}\right)\right) \\ &= \left(1-\frac{ab}{p^2}\right)\left(1+O\left(\frac{\varepsilon^{-2}\log^2 k}{p}\right)\right). \end{aligned}$$

Therefore, collecting the estimates above, we have

$$\log c_k = \sum_{\substack{\varepsilon\sqrt{k}/\log k<p<\varepsilon^{-1}\sqrt{k}\\ a/p,\ b/p\le 1-\varepsilon^2/\log^2 k}} \log\left(1-\frac{ab}{p^2}\right)+O\left(\frac{\varepsilon\sqrt{k}}{\log k}\right), \tag{5.2}$$

where $k\equiv ap+b \pmod{p^2}$.

From Lemma 5.1 we know that the value of b/p is very well equi-distributed on $[0,1)$, as p runs through relatively short intervals of primes. By the prime number theorem with a reasonable error term, we know that a/p is also so distributed. Thus we may estimate (5.2) via partial summation. Without going through the straightforward though lengthy details we simply note that, for fixed integer $d\ge 0$, the sum in (5.2), restricted to those primes p with $[k/p^2]=d$, is

$$= \frac{\sqrt{k}}{\log k}\left(\int_0^1 \frac{1}{(d+t)^{3/2}}\left((1-t^{-1})\log(1-t)-1\right)\mathrm{d}t+O(\varepsilon)\right).$$

We now sum this formula over all integers $d\ge 0$, taking $x=d+t$ so that $t=\{x\}$, and noting that $(1-t^{-1})\log(1-t)-1=-\sum_{j\ge 1} t^j/j(j+1)$, to get

$$c_k = e^{-\{\alpha+o(\varepsilon)\}\sqrt{k}/\log k},$$

where

$$\alpha := \sum_{j\ge 1}\frac{1}{j(j+1)}\int_0^\infty \{x\}^j x^{-3/2}\mathrm{d}x. \tag{5.3}$$

Letting $\varepsilon\to 0$ gives the asymptotic formula for $\log(c_k)$ in Theorem 6, though we still need to find the value of α:

First we split the integral back up, and integrate by parts to get

$$\int_0^\infty \{x\}^j x^{-3/2}\mathrm{d}x = \sum_{n\geq 1}\int_{n-1}^n \{x\}^j x^{-3/2}\mathrm{d}x = \sum_{n\geq 1}\left[\sum_{i\geq 1}\frac{j!}{(j+i)!}\frac{(2i)!}{2^{2i-1}i!}\{x\}^{j+i}x^{-i-1/2}\right]_{n-1}^{n}$$
$$= \sum_{i\geq 1}\frac{j!}{(j+i)!}\frac{(2i)!}{2^{2i-1}i!}\sum_{n\geq 1} n^{-i-1/2} = \sum_{i\geq 1}\frac{j!}{(j+i)!}\frac{(2i)!}{2^{2i-1}i!}\zeta(i+1/2).$$

Therefore

$$\alpha = \sum_{i\geq 1}\frac{(2i)!}{2^{2i-1}i!}\,\zeta(i+1/2)\,\theta_i \qquad \text{where } \theta_i := \sum_{j\geq 1}\frac{1}{j(j+1)}\frac{j!}{(j+i)!}, \tag{5.4}$$

and $\zeta(s)$ is the Riemann zeta-function. Now

$$(i-1)\theta_i = \sum_{j\geq 1}((j+i)-(j+1))\frac{1}{j(j+1)}\frac{j!}{(j+i)!} = \theta_{i-1} - \sum_{j\geq 1}\frac{j-1!}{(j+i)!}$$

Moreover, integrating by parts, we get

$$\frac{t^i}{i(i!)} = \int \frac{t^i t^{-1}}{i!}\mathrm{d}t = \sum_{j\geq 1}\frac{j-1!}{(j+i)!}t^{i+j}t^{-j},$$

so that, above,

$$(i-1)!\theta_i = (i-2)!\theta_{i-1} - \frac{(i-2)!}{i(i!)} = (i-2)!\theta_{i-1} - \frac{1}{i^2(i-1)}$$
$$= (i-3)!\theta_{i-2} - \frac{1}{i^2(i-1)} - \frac{1}{(i-1)^2(i-2)} = \ldots = \theta_1 - \sum_{m=2}^{i}\frac{1}{m^2(m-1)}$$
$$= \sum_{m\geq i+1}\frac{1}{m^2(m-1)} = \sum_{m\geq i+1}\left(\frac{1}{(m-1)} - \frac{1}{m} - \frac{1}{m^2}\right) = \frac{1}{i} - \sum_{m\geq i+1}\frac{1}{m^2}\ .$$

Substituting this into (5.4), we get the formula for α given in the statement of Theorem 6.

To compute α we insert the third to last expression into (5.4) to get

$$\alpha = \sum_{i\geq 1}\binom{2i}{i}\frac{i}{2^{2i-1}}\,\zeta(i+1/2)\sum_{m\geq i+1}\frac{1}{m^2(m-1)}. \tag{5.5}$$

Of course $\zeta(i+1/2) = 1 + O(2^{-i})$ so we start by investigating

$$\alpha_1 = \sum_{i\ge 1} \binom{2i}{i} \frac{i}{2^{2i-1}} \sum_{m\ge i+1} \frac{1}{m^2(m-1)} = \sum_{m\ge 2} \frac{1}{m^2(m-1)} \sum_{i=1}^{m-1} \binom{2i}{i} \frac{i}{2^{2i-1}}.$$

Now, by the binomial theorem, $1/(1-x)^{1/2} = \sum_{i\ge 0} \binom{2i}{i} \frac{x^i}{2^{2i}}$. Therefore $\sum_{i=1}^{m-1} \binom{2i}{i} \frac{i}{2^{2i-1}}$ is the coefficient of x^m in $x/(1-x)$ times $x/(1-x)^{3/2}$, that is $x^2/(1-x)^{5/2}$, and thus equals $\frac{2}{3}\binom{2m}{m}\frac{m(m-1)}{2^{2m-1}}$. Therefore

$$\alpha_1 = \frac{4}{3} \sum_{m\ge 2} \frac{1}{m} \binom{2m}{m} \frac{1}{2^{2m}}.$$

We may integrate the above expansion of $1/(1-x)^{1/2}$ to note that $2\log\left(2(1-\sqrt{1-x})/x\right) = \sum_{m\ge 1} \binom{2m}{m} \frac{x^m}{m2^{2m}}$. Taking $x = 1$ we find that $\alpha_1 = (8\log 2 - 2)/3 \approx 1.181725815\ldots$.

Above we saw that $c_i = \sum_{m\ge i+1} \frac{1}{m^2(m-1)} < 1/i$. Also $\binom{2i}{i}/\binom{2}{1} = \prod_{2\le j\le i} (2j(2j-1)/j(j-1)) \le 2^{2(i-1)}$, so that $\binom{2i}{i}\frac{ic_i}{2^{2i-1}} < 1$. If $s > 1$ then $\zeta(s) - 1 < 2^{-s} + \int_{t\ge 2} t^{-s}\mathrm{d}t = 2^{-s}(s+1)/(s-1)$; so that $\zeta(i+1/2) - 1 < 2^{-(i-2)}$ for any $i \ge 1$. Now by (5.5)

$$\alpha - \alpha_1 = \sum_{i\le I} \binom{2i}{i} \frac{ic_i}{2^{2i-1}} \left(\zeta(i+1/2) - 1\right) + \mathrm{Error}(I),$$

where

$$\mathrm{Error}(I) = \sum_{i>I} \binom{2i}{i} \frac{ic_i}{2^{2i-1}} \left(\zeta(i+1/2) - 1\right) < \sum_{i>I} 2^{-(i-2)} = 2^{-(I-2)}.$$

To facilitate the computation of the c_i we note that $c_1 = 2 - \pi^2/6$ and $c_i = c_{i-1} - \frac{1}{i^2(i-1)}$. Using Maple we then computed α up to an error smaller than 10^{-6} by applying the above with $I = 22$ and got that $\alpha - \alpha_1 \approx .6433819$ and so $\alpha \approx 1.825108$.

5e. The Erdös-Lacampagne-Selfridge problem.

The proof of Theorem 8: Suppose that the smallest prime factor of $(n+1)(n+2)\cdots(n+k)/k!$ is $> k$. That $n \gg k^2/\log k$ has been proved in [ELS], so we may assume this. Let p be any prime in the interval $\frac{k}{2} < p < \frac{k}{2}(1+\varepsilon)$ where $\varepsilon = 1/10$ say. Then $p \mid k!$ but $p^2 \not| k!$, so that $p \mid (n+1)\ldots(n+k)$ but $p^2 \not| (n+1)\ldots(n+k)$. Consider the multiples

$p\left(\left[\frac{n}{p}\right]+1\right)$ and $p\left(\left[\frac{n}{p}\right]+2\right)$ of p. Evidently both are $> n$. However they can't both divide $(n+1)\dots(n+k)$ so

$$p\left(\left[\frac{n}{p}\right]+2\right) > n+k.$$

$$\text{Thus} \qquad \varepsilon k > 2p-k > n - p\left[\frac{n}{p}\right] = p\left\{\frac{n}{p}\right\}$$

$$\text{so that} \qquad 0 \le \left\{\frac{n}{p}\right\} < \frac{\varepsilon k}{p} < \frac{\varepsilon k}{k/2} = 2\varepsilon = \frac{1}{5},$$

$$\text{and then} \qquad \left|\sum_{\frac{k}{2}<p<\frac{k}{2}(1+\varepsilon)} \psi\left(\frac{n}{p}\right)\log p\right| \gg \left(\frac{1}{2}-\frac{1}{5}\right)\frac{\varepsilon k}{2} \gg k.$$

This contradicts (5.1) for our range of n, and thus implies the Theorem.

The reader may care to look at the interesting data collected by Scheidler and Williams [SW] where, for each $k \le 140$, they find the smallest n for which all prime factors of $\binom{n}{k}$ are greater than k. Scheidler and Williams inform us that they have been continuing their computations since then and will soon publish a sequel with many more such n.

6. The proof of Theorem 5.

Fix integer $m \ge 0$ and select another integer M, much larger than m. By Theorems 6 and 7 we see that the number of integers $n \le N$ for which there is some k, $M \le k \le n/2$ for which $\binom{n}{k}$ is squarefree is $< N\exp(-\{\alpha+o(1)\}\sqrt{M}/\log M)$.

Now suppose that the set of integers $K := \{k_1 < k_2 < \dots < k_r\}$ and prime p are given. Just as in section 3c, we can compute the proportion, $c_{K,p}$, of integers n for which p^2 does not divide any of $\binom{n}{k_1}$, $\binom{n}{k_2}, \dots \binom{n}{k_r}$. For example, if $p > k_r$ then $c_{K,p} = 1 - k_r/p^2$. In general the value of $c_{K,p}$ is well defined and is a rather complicated function of the base p digits of $k_1, k_2, \dots, k_r$. It is not, however, necessary to compute $c_{K,p}$, though we do note that it is > 0 since p^2 does not divide any of $\binom{n}{k_1}$, $\binom{n}{k_2}, \dots \binom{n}{k_r}$ when $n \equiv -1 \pmod{p^\ell}$ where $p^\ell > k_r$.

Thus, by the combinatorial sieve, the number of integers $n \le N$ for which $\binom{n}{k_1}$, $\binom{n}{k_2}, \dots \binom{n}{k_r}$ are all squarefree is $\sim c_K N$ where $c_K = \prod_p c_{K,p}$.

A rather different application of the inclusion-exclusion formula tells us that the proportion of integers $n \le N$, for which the set of integers $k \le M$ such that $\binom{n}{k}$ is squarefree is exactly a given set K of m integers, is

$$c_K + o(1) - (m+1) \sum_{\substack{K \subset L \subset \{1,\dots,M\} \\ |L|=m+1}} c_L + \binom{m+2}{2} \sum_{\substack{K \subset L \subset \{1,\dots,M\} \\ |L|=m+2}} c_L - \binom{m+3}{3} \sum_{\substack{K \subset L \subset \{1,\dots,M\} \\ |L|=m+3}} c_L + \dots,$$

which equals $c'_{K,M} + o(1)$, for some constant $c'_{K,M}$, as $N \to \infty$. Let $\eta_{m,M}$ be the sum of $c'_{K,M}$ over all m element subsets K of $\{1, \dots, M\}$. Therefore the number of integers $n \le N$ for which there are exactly $2m$ integers $1 \le k \le n-1$ with $\binom{n}{k}$ squarefree is $\sim N\left(\eta_{m,M} + o(1) + O\left(e^{-\{\alpha+o(1)\}\sqrt{M}/\log M}\right)\right)$. Letting $M \to \infty$, the first part of Theorem 5 follows with $\eta_m = \lim_{M\to\infty} \eta_{m,M}$.

For the second part of Theorem 5, note that if there are m integers k, $1 \le k \le n/2$ for which $\binom{n}{k}$ is squarefree then the largest of them is $\ge m$. Thus, from Theorems 6 and 7 we have that

$$\eta_m \le \sum_{k \ge m} c_k \ll e^{-\{\alpha+o(1)\}\sqrt{m}/\log(2m)}.$$

Remark: It is perhaps worth noting that the c_k are not multiplicatively independent, in the sense that $c_{i,j} \ne c_i c_j$. For example $c_{1,2} = (5/6)c_2$ but $c_1 \ne 5/6$ (see Proposition 3.4 for the values of c_1 and c_2). It is true in general though that if k is the largest element of K then c_K/c_k is a rational number between 0 and 1.

7. Proving Theorem 1 for (explicit) large n.

We shall assume throughout this section that $\binom{2n}{n}$ is squarefree.

Corollary 3.2 holds with 'prime' p changed to 'prime power' p in the hypothesis (the changes in the proof are straightforward). Replacing k and n in Corollary 3.2 by n and $2n$, respectively, we now have

$$\left|\sum_{d\in I}\psi(2n/d)\Lambda(d)\right| + 2\left|\sum_{d\in I}\psi(n/d)\Lambda(d)\right| \geq \frac{1}{2}\sum_{d\in I,\ (d,2n)=1}\Lambda(d), \tag{7.1}$$

where I is the set of integers d in the range $\sqrt{n} < d \leq \sqrt{2n}$.

Theorem 18 and the display following (7.17) from [Va] give the following (also using Theorem 6 and (6.5) from [Va]):

Lemma 7.1. *For any positive integer R, we have*

$$\pm\psi(t) \leq \frac{1}{2R+2} + \sum_{\substack{|r|\leq R\\ r\neq 0}} a_r^{\pm}\mathrm{e}(rt),$$

where

$$a_r^{\pm} = \frac{i}{2\pi(R+1)}\left(\pi\left(1-\frac{|r|}{R+1}\right)\cot\left(\frac{\pi r}{R+1}\right)+\frac{|r|}{r}\right) \pm \frac{1}{2R+2}\left(1-\frac{|r|}{R+1}\right).$$

Therefore we get

$$\begin{aligned}\pm\sum_{d\in I}\psi\left(\frac{X}{d}\right)\Lambda(d) &\leq \frac{1}{2R+2}\sum_{d\in I}\Lambda(d) + \sum_{0<|r|\leq R} a_r^{\pm}\left(\sum_{d\in I} e\left(\frac{rX}{d}\right)\Lambda(d)\right)\\ &\leq \frac{1}{2R+2}\sum_{d\in I}\Lambda(d) + \left(\sum_{0<|r|\leq R}|a_r^{\pm}|\right)\max_{X\leq x\leq XR}\left|\sum_{d\in I} e\left(\frac{x}{d}\right)\Lambda(d)\right|.\end{aligned}$$

Taking $R = 10$, we deduce from (7.1) (and the appropriate computations) that

$$\sum_{d\in I}\Lambda(d) \leq \frac{43}{6}\max_{n\leq x\leq 20n}\left|\sum_{d\in I} e\left(\frac{x}{d}\right)\Lambda(d)\right| + \frac{11}{8}\log n, \tag{7.2}$$

since $\sum_{d\in I,\ (d,2n)>1}\Lambda(d) \leq \log n$.

The entry for $b = 30$ in the table on page 358 of [Sch] means that $|\sum_{d\leq x}\Lambda(d) - x| < x/(4\cdot 10^5)$ for $x \geq \mathrm{e}^{30}$. Therefore

$$\begin{aligned}\sum_{d\in I}\Lambda(d) - \frac{11}{8}\log n &\geq (\sqrt{2}-1)\sqrt{n} - \frac{\sqrt{2}+1}{4\cdot 10^5}\sqrt{n} - \frac{11}{8}\log n\\ &\geq \frac{9999}{10000}\cdot(\sqrt{2}-1)\sqrt{n} \ \geq\ \frac{43}{105}\sqrt{n}\end{aligned}$$

for $n \geq \mathrm{e}^{60}$. Substituting this into (7.2) we get

Lemma 7.2. *Suppose that* $\binom{2n}{n}$ *is not divisible by the square of any prime* $> \sqrt{n}$. *If* $n \ge \mathrm{e}^{60}$ *then*

$$\max_{n\le x\le 20n}\left|\sum_{\sqrt{n}<d\le\sqrt{2n}} e\left(\frac{x}{d}\right)\Lambda(d)\right| \ge \frac{2}{35}\sqrt{n}.$$

Now, if we take $k=2$ in Theorem 9 we find that for $n > 5^{10}$ we have

$$\max_{n\le x\le 20n}\left|\sum_{\sqrt{n}<d\le\sqrt{2n}} e\left(\frac{x}{d}\right)\Lambda(d)\right| \le (3.2)n^{23/48}(\log 256n)^{11/4}.$$

Comparing this to Lemma 7.2 we find that $n \le (56)^{48}(\log 256n)^{132}$, which gives a contradiction for $n \ge 2^{1617}$.

8. Explicit bounds on exponential sums over integers.

We now give several lemmas which provide *explicit* upper bounds for the size of certain exponential sums. The main results are given in Propositions 8.1 and 8.2, where the reader may recognize the exponent pair $(1-k/(2^{k+1}-2), 1/(2^{k+1}-2))$. (For an historical account, the reader is referred to chapter V of Titchmarsh's book [Ti]). It is worth mentioning that in our work we omit the truncated Poisson summation formula from the usual theory.

Proposition 8.1.

(a) If $A \ge (2x)^{1/2}$ *then*

$$\left|\sum_{A<n\le B} e\left(\frac{x}{n}\right)\right| \le \frac{2B^2}{\pi x}.$$

(b) If k is a positive integer, x a positive real number and $1000 \le A < B \le 2A$ with $A \le 4x^{3/5}$ then

$$\left|\sum_{A<n\le B} e\left(\frac{x}{n}\right)\right| \le 32\ \left(\frac{2x}{A^{k+2}}\right)^{1/(2^{k+1}-2)} A\sqrt{\log A}. \tag{8.1}$$

Remark: Note that Proposition 8.1(b) follows from Proposition 8.1(a) for $(2x)^{1/2} \le A \le 4x^{3/5}$. If $A \le (2x)^{1/2}$ the right side of (8.1) is minimized when k is the largest integer satisfying

$$k + 2^{1-k} \le \log(2x)/\log A \tag{8.2}$$

(note that $k \ge 1$). Thus Proposition 8.1(b) follows if we just prove it for this one particular value of k.

Proposition 8.1(b) follows from the more general (and technical).

Proposition 8.2. *Let $k \ge 1$ be an integer. Suppose that, on the interval $[A, B]$, $f(t)$ is $(k+2)$-times differentiable with $f^{(k+1)}(t)$ monotonic. Let m_k and M_k be the minimum and maximum values of $f^{(k+1)}(t)$ on $[A, B]$, respectively. Let $Q = (2M_k)^{-1/(2-2^{-k+1})}$ and suppose that N is an integer such that there are $\le N$ integers in $]A, B]$. If $0 < m_k \le M_k \le 1/(2 \cdot 4^k)$ and $Q \le N$ then*

$$\left| \frac{1}{8N} \sum_{A < n \le B} \mathrm{e}(\pm f(n)) \right| \le \left\{ \frac{1}{8Q} + \frac{3}{\pi N} \frac{M_k}{m_k} \left(\frac{\log Q}{\sqrt{k}} \right)^k \right\}^{1/2^k}.$$

To prove this we shall need various lemmas. First a version of the Weyl-van der Corput lemma (the following can be proved by making suitable (minor) modifications to the proof of Lemma 2.7 in [GK] *):

Lemma 8.3. *Suppose that $\lambda_1, \lambda_2, \dots, \lambda_N$ is a sequence of complex numbers, each with $|\lambda_i| \le 1$, and define $\Delta\lambda_m = \lambda_m$, $\Delta_r \lambda_m = \lambda_{m+r}\overline{\lambda}_m$ and*

$$\Delta_{r_1,\dots,r_k,s}\lambda_m = (\Delta_{r_1,\dots r_k}\lambda_{m+s})\overline{(\Delta_{r_1,\dots,r_k}\lambda_m)}.$$

Then for any given $k \ge 1$, and real number $Q \in [1, N]$,

$$\left| \frac{1}{8N} \sum_{m=1}^{N} \lambda_m \right|^{2^k} \le \frac{1}{8Q} + \frac{1}{8Q^{2-2^{-k+1}}} \sum_{r_1=1}^{Q^{2^{-0}}} \sum_{r_2=1}^{Q^{2^{-1}}} \cdots \sum_{r_k=1}^{Q^{2^{-k+1}}} \left| \frac{1}{N} \sum_{m=1}^{N-r_1-\dots-r_k} \Delta_{r_1,\dots,r_k}\lambda_m \right|.$$

We also need a version of the Kusmin-Landau lemma (see Theorem 2.1 and the notes at the end of of chapter 2 in [GK]):

* There is a slight misprint there. One needs to change the first q to a Q.

Lemma 8.4. *Suppose that, on the interval $[A,B]$, $f(t)$ is a differentiable (real-valued) function, with $f'(t)$ monotonic and $0<m\le f'(t)\le \frac12$. Then*

$$\left|\sum_{A<n\le B} \mathrm{e}(\pm f(n))\right| \le \cot\frac{\pi m}{2} \le \frac{2}{\pi m}.$$

Proposition 8.1(a) follows by taking $f(t)=-x/t$ in Lemma 8.4 – we leave the details to the reader.

Now define $f_{\mathbf{0}}(t)=f(t)$, $f_r(t)=f(t+r)-f(t)$, and $f_{\mathbf{r},s}(t)=f_{\mathbf{r}}(t+s)-f_{\mathbf{r}}(t)$ in general. We have

Lemma 8.5. *If $f(t)$ is $(k+h)$-times differentiable in $[t,t+r_1+\ldots+r_k]$ then*

$$f^{(h)}_{r_1,r_2,\ldots,r_k}(t) = r_1r_2\cdots r_k\ f^{(h+k)}(t+\theta_1r_1+\ldots+\theta_kr_k)$$

for some θ_i, $0<\theta_i<1$.

Proof : By definition $f_{\mathbf{r},d}(t)=f_{\mathbf{r}}(t+d)-f_{\mathbf{r}}(t)$. Differentiating h times we have

$$\begin{aligned} f^{(h)}_{\mathbf{r},d}(t) &= f^{(h)}_{\mathbf{r}}(t+d)-f^{(h)}_{\mathbf{r}}(t)\\ &= d f^{(h+1)}_{\mathbf{r}}(t+\theta d)\end{aligned}$$

for some θ, $0<\theta<1$, by the Mean-Value Theorem. The Lemma follows from iterating this k times.

Proposition 8.2 follows easily from

Lemma 8.6. *Let $k\ge 1$ be an integer. Suppose that, on the interval $[A,B]$, $f(t)$ is $(k+2)$-times differentiable with $f^{(k+1)}(t)$ monotonic and*

$$0<m_k\le f^{(k+1)}(t)\le 1/2Q^{2-2^{-k+1}}$$

for some integer $Q\in[2^k,N]$. If N is an integer such that there are $\le N$ integers in the interval $]A,B]$ then

$$\left|\frac{1}{8N}\sum_{A<n\le B}\mathrm{e}(\pm f(n))\right|^{2^k} < \frac{1}{8Q}+\frac{3k^{-\frac{k}{2}}}{2\pi m_kN}\cdot\frac{\log^k Q}{Q^{2-2^{-k+1}}}.$$

Proof: Let C be the smallest integer in $]A,B]$. Let $\lambda_m = \mathrm{e}(f(C-1+m))$ for any integer m in the range $1 \le m \le [B]-C+1$, and $\lambda_m = 0$ otherwise, so that $\Delta_{\mathbf{r}}\lambda_m = \mathrm{e}(f_{\mathbf{r}}(C-1+m))$ for all $m \ge 1$. Using Lemma 8.3, we get that the left side above is

$$\le \frac{1}{8}\left(\frac{1}{Q} + \frac{1}{Q^{2-2^{-k+1}}}\sum_{r_1=1}^{Q^{2^{-0}}}\sum_{r_2=1}^{Q^{2^{-1}}}\cdots\sum_{r_k=1}^{Q^{2^{-k+1}}}\left|\frac{1}{N}\sum_{A<n\le B-r_1-r_2-\ldots-r_k}\mathrm{e}(\pm f_{\mathbf{r}}(n))\right|\right).$$

In each term of the final sum (when $\mathbf{r}$ is fixed) we find that

$$f'_{\mathbf{r}}(t) = r_1\cdots r_k f^{(k+1)}(y)$$

for some $y \in (t, t+r_1+\ldots+r_k) \subset (A,B)$, by Lemma 8.5. Therefore

$$0 \le f'_{\mathbf{r}}(t) \le r_1\cdots r_k/2Q^{2-2^{-k+1}} \le Q^{2^{-0}}Q^{2^{-1}}\ldots Q^{2^{-k+1}}/2Q^{2-2^{-k+1}} = 1/2$$

in the required interval. Moreover $f''_{\mathbf{r}}(t)$ has the same sign as $f^{(k+2)}(y)$ for some $y \in (A,B)$ by Lemma 8.5 (and the sign of $f^{(k+2)}(y)$ is fixed in this interval according to the hypothesis), and so $f'_{\mathbf{r}}(t)$ is monotonic in the required interval. Thus the hypotheses of Lemma 8.4 are satisfied for the function $f_{\mathbf{r}}(t)$, and so the last term above is

$$\begin{aligned}
&\le \frac{1}{8}\cdot\frac{1}{Q^{2-2^{-k+1}}}\sum_{r_1=1}^{Q^{2^{-0}}}\sum_{r_2=1}^{Q^{2^{-1}}}\cdots\sum_{r_k=1}^{Q^{2^{-k+1}}}\left(\frac{2}{\pi N}\frac{1}{\min\limits_{A\le t\le B-r_1-\ldots-r_k} f'_{\mathbf{r}}(t)}\right)\\
&\le \frac{1}{4\pi N}\frac{1}{Q^{2-2^{-k+1}}\min\limits_{A\le y\le B} f^{(k+1)}(y)}\cdot\left(\sum_{r_1=1}^{Q^{2^{-0}}}\frac{1}{r_1}\right)\left(\sum_{r_2=1}^{Q^{2^{-1}}}\frac{1}{r_2}\right)\cdots\left(\sum_{r_k=1}^{Q^{2^{-k+1}}}\frac{1}{r_k}\right)\\
&\le \frac{1}{4\pi m_k N Q^{2-2^{-k+1}}}(1+\log Q)\left(1+\frac{\log Q}{2}\right)\cdots\left(1+\frac{\log Q}{2^{k-1}}\right),
\end{aligned}$$

since $\sum_{r=1}^{R}\frac{1}{r} \le 1+\log R$. Now $Q \ge 2^k$ and so

$$(1+\log Q)\ldots\left(1+\frac{\log Q}{2^{k-1}}\right) \le \left(\frac{\log Q}{k\log 2}\right)^k(1+k\log 2)\ldots\left(1+\frac{k\log 2}{2^{k-1}}\right).$$

The result follows by establishing that

$$(1+k\log 2)\ldots\left(1+\frac{k\log 2}{2^{k-1}}\right) \le 6k^{-\frac{k}{2}}(k\log 2)^k$$

for all integers $k \ge 1$. This is easily checked for $k \le 7$. For $k \ge 8$ we use the upper bounds

$$1 + 2^{-j} k\log 2 \le \begin{cases} 2^{-j} k\log 2 \exp(2^j/(k\log 2)) & \text{for } j < l \\ \exp(2^{-j} k\log 2) & \text{for } j \ge l \end{cases}$$

where l is an integer, chosen so that $2^{l-1} \le k\log 2 < 2^l$, to get

$$\begin{aligned} \prod_{j\ge 0}(1 + 2^{-j} k\log 2) &\le \left(\prod_{0\le j<l} 2^{-j} k\log 2\right) \exp\left(\frac{2^l}{k\log 2} + \frac{2k\log 2}{2^l}\right) \\ &\le \left(2^{-(l-1)/2} k\log 2\right)^l \mathrm{e}^3 \le 2^{l(l+1)/2}\mathrm{e}^3 \le 6k^{-\frac{k}{2}}(k\log 2)^k \end{aligned}$$

for $k \ge 8$.

Proof of Proposition 8.1(b): As noted in the remark following the statement of the result, we need only prove (8.1) for $A \le (2x)^{1/2}$ and k the largest integer satisfying (8.2).

Take $f(t) = (-1)^{k+1}x/t$ in Proposition 8.2 so that $M_k = x(k+1)!/A^{k+2}$, $m_k = x(k+1)!/B^{k+2}$, and thus $(M_k/m_k) = (B/A)^{k+2} \le 4 \cdot 2^k$. Also let $N = [A] + 1$; we easily deduce that $Q \le N$ from (8.2). We may assume that $M_k \le 1/2 \cdot 4^k$ else $x/A^{k+2} > 1/2 \cdot 4^k(k+1)!$ and the bound given in (8.1) is worse than that given by trivially bounding every term in the sum by 1.

Using the fact that $Q \le N \le A + 1$, the upper bound given by Proposition 8.2 is

$$\begin{aligned} &\le 8N\left(\frac{1}{8Q} + \frac{12}{\pi Q}\left(\frac{2\log N}{\sqrt{k}}\right)^k\right)^{\frac{1}{2^k}} \\ &\le 8N\left(\frac{2x(k+1)!}{A^{k+2}}\right)^{\frac{1}{2^{k+1}-2}} (\log N)^{k2^{-k}} \left\{\frac{1}{8(\log N)^k} + \frac{12}{\pi}\left(\frac{2}{\sqrt{k}}\right)^k\right\}^{\frac{1}{2^k}}, \\ &\le 8(k+1)!^{\frac{1}{2^{k+1}-2}} \left\{\frac{1}{8(\log 1000)^k} + \frac{12}{\pi}\left(\frac{2}{\sqrt{k}}\right)^k\right\}^{\frac{1}{2^k}} \frac{1001\sqrt{\log 1001}}{1000\sqrt{\log 1000}} \times \\ &\qquad \times \left(\frac{2x}{A^{k+2}}\right)^{\frac{1}{2^{k+1}-2}} A\sqrt{\log A} \end{aligned}$$

since $1000 \le A < N \le A + 1$. The maximum value of the constant two lines above is 31.34 which is attained when $k = 1$; the result follows.

9. Explicit bounds on exponential sums over primes.

One can deduce bounds for exponential sums over primes from bounds for exponential sums over integers, using the celebrated idea of writing Λ as a linear combination of bilinear forms. We do so by using Vaughan's identity, and get non-trivial results for a wide range of values of y (see Theorems 9 and 9' and Corollary 2 above).

9a. The general principle.

We apply Vaughan's identity (see section 24 in [Da]) to get the following:

Lemma 9.1 (Vaughan's identity). *Let f be any function, and $N, K, M > 0$ real numbers satisying $2K \le N$. Then*

$$\sum_{N/2<n\le N} \Lambda(n)f(n) = \sum_{\substack{N/2<lm\le N\\ m\le M}} \mu(m)\log l f(lm) - \sum_{\substack{N/2<lr\le N\\ r\le MK}} b_r f(lr) - \sum_{\substack{N/2<kl\le N\\ k>K,\ l>M}} a_l\Lambda(k)f(kl),$$

where

$$a_l = \sum_{\substack{mr=l\\ m\le M}} \mu(m) \quad \text{and} \quad b_r = \sum_{\substack{mk=r\\ m\le M, k\le K}} \mu(m)\Lambda(k).$$

We shall take $N = y'$ and $K = M = y'^{\frac{1}{3}}$ later on, but for now we keep the more general notation.

By Vaughan's identity (Lemma 9.1) we have, for $K \le y < y' \le 2y \le x/2$, that

$$\sum_{y<n\le y'} \Lambda(n)e\left(\frac{x}{n}\right) = \Sigma_1^y - \Sigma_{2,1}^y - \Sigma_{2,2}^y - \Sigma_3^y$$

where

$$\begin{aligned}
\Sigma_1^y &= \sum_{m\le M} \mu(m) \sum_{\frac{y}{m}<l\le\frac{y'}{m}} \log l\ e\left(\frac{x}{lm}\right),\\
\Sigma_{2,1}^y &= \sum_{r\le M} b_r \sum_{\frac{y}{r}<l\le\frac{y'}{r}} e\left(\frac{x}{lr}\right),\\
\Sigma_{2,2}^y &= \sum_{M<r\le MK} b_r \sum_{\frac{y}{r}<l<\frac{y'}{r}} e\left(\frac{x}{lr}\right),\\
\Sigma_3^y &= \sum_{M<l<\frac{y'}{K}} a_l \sum_{K,\ \frac{y}{l}<k<\frac{y'}{l}} \Lambda(k)e\left(\frac{x}{kl}\right).
\end{aligned}$$

We shall need the following straightforward lemma:

Lemma 9.2. $\left|\sum_{n=A}^{B} e\left(\frac{x}{n}\right)\log n\right| \le \log\left(\frac{B^2}{A}\right) \max_{A\le t\le B}\left|\sum_{A\le n\le t} e\left(\frac{x}{n}\right)\right|$.

Proof: By partial summation we obtain

$$\sum_{n=A}^{B} e\left(\frac{x}{n}\right)\log n = \log B \sum_{A\le n\le B} e\left(\frac{x}{n}\right) - \int_A^B \left(\sum_{A\le n\le t} e\left(\frac{x}{n}\right)\right)\frac{dt}{t}$$

and the result follows.

By Lemma 9.2 we have

$$|\Sigma_1^y| \le \sum_{m\le M}\left|\sum_{\frac{y}{m}<l\le\frac{y'}{m}} e\left(\frac{x}{lm}\right)\log l\right| \le \sum_{m\le M}\log\left(\frac{y'^2}{my}\right)\max_{y\le z\le y'}\left|\sum_{\frac{y}{m}<n\le\frac{z}{m}} e\left(\frac{x}{mn}\right)\right|.$$

Also, since $|b_r|\le \log r$, we have

$$|\Sigma_{2,1}^y| = \left|\sum_{r\le M} b_r \sum_{\frac{y}{r}<l\le\frac{y'}{r}} e\left(\frac{x}{lr}\right)\right| \le \sum_{r\le M}\log r\left|\sum_{\frac{y}{r}<l\le\frac{y'}{r}} e\left(\frac{x}{lr}\right)\right|.$$

Adding these two expressions together we get

$$|\Sigma_1^y| + \left|\Sigma_{2,1}^y\right| \le \log\left(\frac{y'^2}{y}\right)\sum_{m\le M}\max_{y\le z\le y'}\left|\sum_{\frac{y}{m}<n\le\frac{z}{m}} e\left(\frac{x}{mn}\right)\right|. \tag{9.1}$$

Lemma 9.3. *If $M\le y^2/2x$ then*

$$|\Sigma_1^y| + |\Sigma_{2,1}^y| \le \frac{8}{\pi}\frac{y^2}{x}\log^2(4y).$$

Otherwise, for any positive integer k,

$$|\Sigma_1^y| + |\Sigma_{2,1}^y| \le 64\left(\frac{2^k-1}{k+1}\right)\ y\left(\frac{2xM^{k+1}}{y^{k+2}}\right)^{\frac{1}{2^{k+1}-2}}\log^{3/2}(4y)$$

provided $y \geq 1000M$ *and* $y \leq 4x^{3/5}$.

Proof: If $M \leq y^2/2x$ then $(y/m) \geq (2x/m)^{1/2}$ for all $m \leq M$ and so by Proposition 8.1(a) we have

$$\sum_{m\leq M} \max_{y\leq z\leq y'} \left| \sum_{\frac{y}{m}<n\leq \frac{z}{m}} e\left(\frac{x}{mn}\right)\right| \leq \sum_{m\leq M} \frac{2y'^2}{\pi m x} \leq \frac{2y'^2}{\pi x}(1+\log M)$$
$$\leq \frac{8y^2}{\pi x}\ \log\ \left(\frac{\mathrm{e}y^2}{2x}\right),$$

since $M \leq y^2/2x$. The first estimate of the Lemma follows once we insert this bound into (9.1), and note that $\mathrm{e}y^2/2x \leq 4y$.

If $M > y^2/2x$ then, using Proposition 8.1(b), we get

$$\sum_{m\leq M} \max_{\frac{y}{m}\leq \frac{z}{m}\leq \frac{y'}{m}} \left| \sum_{\frac{y}{m}\leq n\leq \frac{z}{m}} e\left(\frac{x}{mn}\right)\right| \leq$$
$$\leq 32 \sum_{m\leq M} \left(\frac{y}{m}\right)^{1-\frac{k+2}{2^{k+1}-2}} \left(\frac{2x}{m}\right)^{\frac{1}{2^{k+1}-2}} \sqrt{\log y}$$
$$\leq 32 y^{1-\frac{k+2}{2^{k+1}-2}} x^{\frac{1}{2^{k+1}-2}} \sqrt{\log y} \sum_{m\leq M} \frac{1}{m^{1-\frac{k+1}{2^{k+1}-2}}}$$
$$\leq 64 \left(\frac{2^k-1}{k+1}\right)\ y^{1-\frac{k+2}{2^{k+1}-2}} x^{\frac{1}{2^{k+1}-2}} M^{\frac{k+1}{2^{k+1}-2}} \sqrt{\log y}.$$

The second estimate in the Lemma now follows from inserting this bound into (9.1).

In order to get an upper bound for $|\Sigma_{2,2}^y|$, we split the range of summation for r into ranges $R < r \leq \min(2R, MK)$ with $R = M, 2M, 2^2M, \ldots$; we also split the relevant range for l into two parts $y/2R < l \leq y'/2R$ and $y'/2R < l \leq y'/R$. Thus we get

$$\left|\Sigma_{2,2}^y\right| \leq 2\left(\frac{\log K}{\log 2}+1\right) \max_{\substack{M<R\leq MK \\ y/2R<L\leq y'/2R}} \left| \sum_{R<r\leq R'} \sum_{\substack{L<l\leq L' \\ y<lr<y'}} b_r e\left(\frac{x}{lr}\right)\right|$$

$$\leq \frac{2\log\,(8y')}{\log 8} \max_{\substack{y'^{\frac{1}{3}}<R\leq y'^{\frac{2}{3}} \\ y/2R<L\leq y/R}} \left| \sum_{R<r\leq R'} \sum_{\substack{L<l\leq L' \\ y<lr<y'}} b_r e\left(\frac{x}{lr}\right)\right| \tag{9.2a}$$

when we take $M = K = y'^{\frac{1}{3}}$; where R' and L' denote real numbers such that $R < R' \le \min(2R, y'^{\frac{2}{3}})$ and $L < L' \le \min(2L, y'^{\frac{2}{3}})$. Similarly

$$|\Sigma_3^y| \le 2\left(\frac{\log\left(\frac{y'}{MK}\right)}{\log 2}+1\right) \max_{\substack{M<L\le y'/K\\ y/2L<R\le y'/2L}} \left|\sum_{R<r\le R'} \sum_{\substack{L<l\le L'\\ y<lr<y'}} a_l\Lambda(r)e\left(\frac{x}{lr}\right)\right|$$

$$\le \frac{2\log(8y')}{\log 8} \max_{\substack{y'^{\frac{1}{3}}<L\le y'^{\frac{2}{3}}\\ y/2L<R\le y/L}} \left|\sum_{R<r\le R'} \sum_{\substack{L<l\le L'\\ y<lr<y'}} a_l\Lambda(r)e\left(\frac{x}{lr}\right)\right|. \tag{9.2b}$$

In order to bound such exponential sums we prove the following result.

Proposition 9.4. *Suppose that we are given sequences of complex numbers, α_u, supported on $]U, 2U]$, and β_v, supported on $]V, 2V]$, where $U, V \ge 10$ are integers.*

(a) If $U^2V \ge 2x$ then, for any interval I,

$$\left|\sum_{uv\in I} \alpha_u\beta_v e\left(\frac{x}{uv}\right)\right|^2 \le 2\,\|\alpha\|_2^2\|\beta\|_2^2\left(U + 8\left(\frac{(UV)^3}{\pi x}\right)^{1/2}\right). \tag{9.3}$$

(b) If $U \ge 1000$ and if k is a positive integer for which

$$V\left(\frac{x}{VU^{k+2}}\right)^{\frac{1}{2^{k+1}-2}} \ge 1 \quad\text{and}\quad V\left(\frac{x}{VU^{k+2}}\right)^{\frac{1}{2^{k+1}-2}} \ge \frac{(U/2)^{5/3}V^2}{x}$$

then, for any interval I,

$$\left|\sum_{uv\in I} \alpha_u\beta_v e\left(\frac{x}{uv}\right)\right|^2 \le 68\,\|\alpha\|_2^2\|\beta\|_2^2 UV\,\left(\frac{x}{VU^{k+2}}\right)^{\frac{1}{2^{k+1}-2}}\sqrt{\log 2U}. \tag{9.4}$$

Remarks: The Cauchy-Schwarz inequality gives the 'trivial' upper bound $\|\alpha\|_2^2\|\beta\|_2^2 UV$. Corollary 9.7 will provide an easy way to apply Proposition 9.4 to the equations (9.2).

To prove Proposition 9.4 we shall use the following lemma, which is easily deduced from Theorem 19 of [Va]. Let $\chi_{u,v}$ be the characteristic function of $[u, v] + \mathbf{Z}$; that is, $\chi_{u,v}(t) = 1$ if there exists an integer in the interval $[u-t, v-t]$, and $\chi_{u,v}(t) = 0$ otherwise.

Lemma 9.5. *Fix u and v. For any positive integer L, there exist complex numbers $\{c_\ell^\pm(L)\}_{|\ell|\le L}$, such that*

$$\sum_{|\ell|\le L} c_\ell^-(L)\mathrm{e}(\ell t) \le \chi_{u,v}(t) \le \sum_{|\ell|\le L} c_\ell^+(L)\mathrm{e}(\ell t),$$

where, for $c = c^+$ or c^-,

$$|c_\ell(L)| \le |c_0(L)| = v - u + \frac{1}{L+1}.$$

Corollary 9.6. *We have*

$$\sum_{|v_1^{-1}-v_2^{-1}|\le \Delta/V^2} |\beta_{v_1}\beta_{v_2}| \le (8\Delta+2)\|\beta\|_2^2.$$

Proof: Since β_v is only supported on $]V, 2V]$, we have

$$\sum_{|v_1^{-1}-v_2^{-1}|\le \Delta/V^2} |\beta_{v_1}\beta_{v_2}| = \sum_{|v_1-v_2|\le 4\Delta} |\beta_{v_1}\beta_{v_2}| \le \sum_{\delta|v_1-v_2|/(4\Delta)\le\delta} |\beta_{v_1}\beta_{v_2}|,$$

where $\delta > 0$ is a parameter to be chosen later. We apply Lemma 9.5 to $\chi_{-\delta,\delta}$, for some positive integer L, so that

$$\begin{aligned}
\sum_{\delta|v_1-v_2|/(4\Delta)\le\delta} |\beta_{v_1}\beta_{v_2}| &\le \sum_{v_1,\ v_2} |\beta_{v_1}\beta_{v_2}| \sum_{|\ell|\le L} c_\ell^+(L) e\left(\frac{\ell\delta(v_1-v_2)}{4\Delta}\right) \\
&\le \sum_{|\ell|\le L} |c_\ell^+(L)| \left|\sum_v |\beta_v| e\left(\frac{v\ell\delta}{4\Delta}\right)\right|^2 \\
&\le \left(2\delta + \frac{1}{L+1}\right) \sum_{|\ell|\le L} \left|\sum_v |\beta_v| e\left(\frac{v\ell\delta}{4\Delta}\right)\right|^2 \\
&\le \left(2\delta + \frac{1}{L+1}\right) \|\beta\|_2^2 \ \left(V + \frac{4\Delta}{\delta}\right),
\end{aligned}$$

by the large sieve inequality (as in Théorème 4 of [Bo]) provided $(2L+1)\delta \le 4\Delta$. Now choose $\delta = 4\Delta/(2L+1)$, while letting $L \to \infty$ (running through integer values only), and the result follows.

Proof of Proposition 9.4 : In (a) we shall let $\Delta = (UV^3/\pi x)^{1/2}$; in (b) we shall let $\Delta = V(x/VU^{k+2})^{\frac{1}{2^{k+1}-2}}$, which is ≥ 1 by the hypothesis. Applying the Cauchy-Schwarz inequality to our exponential sum we get the upper bound

$$\le \left(\sum_u |\alpha_u|^2\right)\left(\sum_{v_1,v_2} |\beta_{v_1}\beta_{v_2}| \left| \sum_{u\in]U,2U]\cap \frac{1}{v_1}I\cap\frac{1}{v_2}I} e\left(\frac{x}{u}\left(\frac{1}{v_1}-\frac{1}{v_2}\right)\right)\right|\right)$$

$$\le \|\alpha\|_2^2 \left((8\Delta+2)\|\beta\|_2^2 U + V\|\beta\|_2^2 \max_{\substack{U\le U_1\le U_2\le 2U \\ \frac{x}{V}\ge x^*\ge \frac{\Delta}{V^2}x}} \left| \sum_{U_1<u\le U_2} e\left(\frac{x^*}{u}\right)\right|\right)$$

where $x^* = x(1/v_1 - 1/v_2)$, and the first term is obtained by bounding the exponential sum by U if $x^* < \Delta x/V^2$, then applying Corollary 9.6; and the second term is obtained from the remaining terms by applying the Cauchy-Schwarz inequality to $\sum |\beta_{v_1}\beta_{v_2}|$.

For (a) we apply Proposition 8.1(a) to the remaining exponential sum, and the result follows.

For (b) we apply Proposition 8.1(b) to the remaining exponential sum and, since $U \ge 1000$ and $2U \le 4(\Delta x/V^2)^{3/5}$ by hypothesis, we deduce that the above is

$$\le \|\alpha\|_2^2\|\beta\|_2^2 UV\sqrt{\log 2U}\ \left(\frac{x}{VU^{k+2}}\right)^{\frac{1}{2^{k+1}-2}}\left\{\frac{10}{\sqrt{\log 2000}}+32\cdot 2\right\},$$

and the result follows.

Corollary 9.7. *Suppose that we are given sequences of complex numbers, α_u, supported on $]U,2U]$, and β_v, supported on $]V,2V]$, where U and V are integers satisying $(2y)^{2/3} \ge U \ge V \ge 10$ with $y/2 \le UV \le y$.*

(a) Suppose that $x \ge y \ge x^{3/5}/5$ are given real numbers. If $U^2V \ge 2x$ then, for any interval I,

$$\left|\sum_{uv\in I} \alpha_u\beta_v e\left(\frac{x}{uv}\right)\right| \le 4.61\ \|\alpha\|_2\|\beta\|_2\left(\frac{y^3}{x}\right)^{1/4}. \tag{9.5}$$

(b) Suppose that $x^{3/5}/5 \ge y \ge 5000$ are given real numbers, and k is a given positive integer. If $U \ge 1000$ then, for any interval I,

$$\left|\sum_{uv\in I}\alpha_u\beta_v e\left(\frac{x}{uv}\right)\right| \le 10.54\ \|\alpha\|_2\|\beta\|_2 y^{1/2}\left(\frac{x}{y^{\frac{k+3}{2}}}\right)^{\frac{1}{4(2^k-1)}}\log^{1/4}(6y). \tag{9.6}$$

Proof : To prove (a) we apply Proposition 9.4(a) and note, by hypothesis, that the upper bound there is $\le 2\|\alpha\|_2^2\|\beta\|_2^2\times$

$$(2y)^{2/3}+8(y^3/\pi x)^{1/2} \le (y^3/x)^{1/2}\cdot\{(16x^3/y^5)^{1/6}+8/\pi^{1/2}\},$$

and the result follows since $x^3/y^5 \le 5^5$.

To prove Corollary 9.7(b) we shall apply Proposition 9.4(b), so we must verify the hypotheses there. The first inequality is $V^{2^{k+1}-2}x \ge VU^{k+2}$. This may be re-arranged as $(UV)^{2^{k+1}-3}x \ge U^{2^{k+1}+k-1}$. Now since $UV \ge y/2$, $x \ge (5y)^{5/3}$ and $U \le (2y)^{2/3}$, the above follows since $(y/2)^{2^{k+1}-3}(5y)^{5/3} \ge (2y)^{\frac{2}{3}(2^{k+1}+k-1)}$ for $y \ge 5000$. The second inequality there is $(xV)^{2^{k+1}-2}x \ge ((U/2)^{5/3}V^2)^{2^{k+1}-2}VU^{k+1}$. This may be re-arranged as $(x/UV)^{2^{k+1}-1} \ge (U^2/32)^{\frac{1}{3}(2^{k+1}-2)}U^k$. Using the inequalities above, and the fact that $UV \le y$, this follows since $(2\cdot 5^{5/3}(2y)^{2/9})^{2^{k+1}-1} \ge 2^{5/3}(2y)^{2k/3-4/9}$.

To deduce (9.6) from (9.4) we use the inequalities in our hypothesis, as well as $VU^{k+2} \ge (UV)^{\frac{k+3}{2}} \ge (y/2)^{\frac{k+3}{2}}$ since $U \ge V$. The upper bound that we get contains the main term above times the constant $(68)^{1/2}(2/3)^{1/4}2^{(k+3)/8(2^k-1)} \le 10.54$.

9b. Putting everything together: the proof of Theorem 9.

We now assume that k is a given integer ≥ 1 and

$$\frac{1}{5}x^{3/5} \ge y \ge 2\cdot 10^6.$$

By Lemma 9.3 we have

$$|\Sigma_1^y| + |\Sigma_{2,1}^y| \le 64\left(\frac{2^k-1}{k+1}\right)\ y\left(\frac{2xM^{k+1}}{y^{k+2}}\right)^{\frac{1}{2^{k+1}-2}}\log^{3/2}(4y). \tag{9.7}$$

To estimate the exponential sums in (9.2) we can apply Corollary 9.7 with $U \ge V$ equal to the numbers R and L: it is easy to check that the hypotheses of Corollary 9.7(b) are

satisfied for each sum. Note that $V \ge y/2(2y)^{2/3} \ge 40$. We also need bounds on $\|\alpha\|_2\|\beta\|_2$ for each sum:

For $\left|\Sigma_{2,2}^y\right|$, we have the bound

$$\|\alpha\|_2^2\|\beta\|_2^2 \le \left(\sum_{L<l\le 2L} 1^2\right)\left(\sum_{R<r\le 2R} b_r^2\right) \le UV\log^2(2U)$$
$$\le y\log^2(2(2y)^{2/3}) \le (0.023)y\log^3(16y),$$

since each $|b_r| \le \log r$, and $y \ge 2\times 10^6$.

For $|\Sigma_3^y|$ we have, using Proposition 10.1 (see section 10),

$$\|\alpha\|_2^2\|\beta\|_2^2 = \left(\sum_{L<l\le 2L} a_l^2\right)\left(\sum_{R<r\le 2R} \Lambda(r)^2\right) \le \frac{4}{3}L(\log M+3)^2\left(\sum_{R<r\le 2R}\Lambda(r)^2\right).$$

Now, by (3.35) and (3.16) of [RS] for $R \ge 125$, and by direct computation for $125 \ge R \ge 40$, we obtain

$$\sum_{R<r\le 2R}\Lambda(r)^2 \le \log(2R)(\psi(2R)-\psi(R)) \le 1.285R\log(2R),$$

Therefore, since $y \ge 2\times 10^6$,

$$\|\alpha\|_2^2\|\beta\|_2^2 \le \frac{4}{3}1.285y(\log M+3)^2\log(2R) \le (0.62)y\log^3(16y).$$

Inserting these bounds into Corollary 9.7(b), we deduce from (9.2) that

$$\left|\Sigma_{2,2}^y\right| + |\Sigma_3^y| \le (10.54)\frac{2}{3\log 2}\{(0.023)^{1/2}+(0.62)^{1/2}\}y\left(\frac{x}{y^{\frac{k+3}{2}}}\right)^{\frac{1}{4(2^k-1)}}\log^{11/4}(16y)$$

$$\le (9.52)y\left(\frac{x}{y^{\frac{k+3}{2}}}\right)^{\frac{1}{4(2^k-1)}}\log^{11/4}(16y). \tag{9.8}$$

If $x \le y^{\frac{k+3}{2}}$ and $y \ge 2^{2^k}$ then we can combine (9.7) and (9.8) to deduce the bound in Theorem 9. The bound

$$\left|\sum_{y<n\le y'}\Lambda(n)\mathrm{e}(x/n)\right| \le \left|\sum_{y<n\le 2y}\Lambda(n)\right| \le \log\binom{[2y]}{[y]} + \sum_{p\le\sqrt{2y}}\log p$$
$$\le \log(2^{2y}) + \sqrt{2y}\log(\sqrt{2y}) \le 2y\log 2 + \sqrt{y/2}\,\log(2y). \tag{9.9}$$

is evidently better than that given in Theorem 9 when $x \geq y^{\frac{k+3}{2}}$. It is also better when $1 \leq x \leq y^{\frac{k+3}{2}}$ and $y \leq \max\{2^{2^k}, 2 \cdot 10^6\}$, as may be shown from just taking $x = 1$ and performing the pertinent computations.

9c. Applying Corollary 9.7(a).

If we try to apply Corollary 9.7(a) we find that it is only for large y that the hypotheses are satisfied: in order that $U^2V \geq 2x$ we need that $y \geq 2x^{2/3}$. If we assume this then Lemma 9.3 gives

$$|\Sigma_1^y| + |\Sigma_{2,1}^y| \leq \frac{8}{\pi}\, y \left(\frac{y}{x}\right) \log^2(4y).$$

Proceeding as in section 9b, we deduce from Corollary 9.7(a) and (9.2) that

$$\begin{aligned} \left|\Sigma_{2,2}^y\right| + |\Sigma_3^y| &\leq (4.61)\frac{2}{3\log 2}\{(0.023)^{1/2} + (0.62)^{1/2}\} y \left(\frac{y}{x}\right)^{\frac{1}{4}} \log^{5/2}(16y) \\ &\leq 4.17\; y \left(\frac{y}{x}\right)^{\frac{1}{4}} \log^{5/2}(16y). \end{aligned}$$

These bounds combine to give Theorem 9' for $y \geq 2 \cdot 10^6$.

In the remaining range $y \leq 2 \cdot 10^6$, the upper bound given is (as before) bigger than that in (9.9) since $x \leq (y/2)^{3/2} \leq 10^9$.

10. Some explicit estimates.

In this section we shall prove

Proposition 10.1. *For any $N, z \geq 1$ we have*

$$\sum_{N < n \leq 2N} \left(\sum_{\substack{d \mid n \\ d \leq z}} \mu(d) \right)^2 \leq \frac{4}{3} N (\log z + 3)^2.$$

Remark : This sum was also investigated in [DIT].

Before starting on the proof we need some preparatory lemmas:

Lemma 10.2. *For any integer d and any $N \geq 1$ we have*

$$\left| \sum_{\substack{n \leq N \\ (n,d)=1}} \frac{\mu(n)}{n} \right| \leq 1.$$

Proof: We may assume that N is an integer. Let d' be the product of all of the primes up to N which do not divide d. Then

$$\begin{aligned} \sum_{\substack{m \leq N \\ (m,d')=1}} 1 = \sum_{m \leq N} \sum_{\substack{n \mid m \\ (n,d)=1}} \mu(n) &= \sum_{\substack{n \leq N \\ (n,d)=1}} \mu(n) \left[\frac{N}{n} \right] \\ &= N \sum_{\substack{n \leq N \\ (n,d)=1}} \frac{\mu(n)}{n} - \sum_{\substack{n \leq N \\ (n,d)=1}} \mu(n) \left\{ \frac{N}{n} \right\}. \end{aligned}$$

Now, since $\{N/1\} = 0$ thus

$$\left| \sum_{\substack{n \leq N \\ (n,d)=1}} \frac{\mu(n)}{n} \right| \leq \frac{1}{N} \left(\sum_{\substack{n \leq N \\ (n,d)=1 \text{ or } (n,d')=1}} 1 \right) \leq 1.$$

Lemma 10.3. *For any $N \geq 1$ we have*

$$\sum_{n \leq N} \frac{\mu^2(n)}{n} \leq \frac{2}{3}(\log N + 3) \quad \text{and} \quad \sum_{n \leq N} \frac{\mu^2(n)}{n} \tau(n) \leq \frac{4}{9}(\log N + 3)^2,$$

where $\tau(n)$ denotes the number of divisors of n.

Proof: We can put an upper bound on the number of squarefree integers up to N by just counting those that are not divisible by either 4 or 9; this gives us $\leq (1-1/4)(1-1/9)N + 2^{2-1} = (2/3)N + 2$. If $N \geq 49$ then we can remove the numbers 25 and 49 from our count, leaving $\leq (2/3)N$. By explicit computations up to $N = 49$ we thus get

$$\sum_{n \leq N} \mu^2(n) \leq (2/3)(N+2);$$

equality being attained when $N = 7$. The first result above is then deduced through partial summation. The second result from writing out each factorization $n = ab$ to get

$$\sum_{n\le N}\frac{\mu^2(n)}{n}\tau(n) \le \sum_{ab\le N}\frac{\mu^2(a)\mu^2(b)}{ab} \le \left(\sum_{a\le N}\frac{\mu^2(a)}{a}\right)^2;$$

and then substituting in the previous estimate.

We now complete the proof of Proposition 10.1:

Proof : The left side above is

$$\le N\left|\sum_{\substack{d_1,d_2\le z\\ [d_1,d_2]\le 2N}}\frac{\mu(d_1)\mu(d_2)}{[d_1,d_2]}\right| + \sum_{\substack{d_1,d_2\le z\\ [d_1,d_2]\le 2N}}\mu^2(d_1)\mu^2(d_2).$$

Let $d = (d_1, d_2)$ and $d_1 = da,\ d_2 = db$. Then $(b, ad) = 1,\ b \le z/d$ and $b \le 2N/da$ so that the first sum is

$$\le N\sum_{d_1\le z}\frac{\mu^2(d_1)}{d_1}\sum_{d|d_1}\left|\sum_{\substack{b\le \min(z/d,2N/d_1)\\ (b,d_1)=1}}\frac{\mu(b)}{b}\right| \le N\sum_{d_1\le z}\frac{\mu^2(d_1)}{d_1}\tau(d_1) \le \frac{4}{9}N(\log z+3)^2,$$

using Lemma 10.1 and then the second part of Lemma 10.2.

The second sum above is

$$\le \sum_{\substack{a,b\le z\\ (a,b)=1}}\mu^2(a)\mu^2(b)\sum_{d\le 2N/ab}\mu^2(d) \le 2N\left(\sum_{a\le z}\frac{\mu^2(a)}{a}\right)\left(\sum_{b\le z}\frac{\mu^2(b)}{b}\right) \le \frac{8}{9}N(\log z+3)^2$$

using the first part of Lemma 10.2. The result follows from adding the two bounds above together.

References

[BLS] J. Brillhart, D.H. Lehmer and J.L. Selfridge, *New primality criteria and factorizations of* $2^m \pm 1$, Math. Comp. **29** (1975), 620–647.

[Bo] E. Bombieri, *Le grand crible dans la théorie analytique des nombres*, Astérisque **18** (1987/1974) 103 pp.

[C] P. Cutter, *Finding big prime k-tuplets* (to appear).

[Da] H. Davenport, *Multiplicative Number Theory* 2nd ed. (Springer-Verlag, New York, 1980).

[DIT] F. Dress, H. Iwaniec and G. Tenenbaum, *Sur une somme liée à la fonction de Möbius*, J. reine angew. Math. 340 (1983) 53–58.

[EG] P. Erdös and R.L. Graham, *Old and new problems and results in combinatorial number theory*, Enseign. Math. Geneva (1980).

[ELS] P. Erdös, C.B. Lacampagne and J.L. Selfridge, *Estimates of the least prime factor of a binomial coefficient*, Math. Comp. **61** (1993) 215–224.

[FL] J.B. Friedlander and J.C. Lagarias, *On the distribution in short intervals of integers having no large prime factor*, J. Number Theory **25** (1987) 249–273

[Ga] P.X.Gallagher, *Sieving by prime powers*, Acta Arith. **24** (1974) 491–497

[Go] P. Goetgheluck, *On prime divisors of binomial coefficients*, Math. Comp. **51** (1988), 325–329.

[GK] S.W. Graham and G. Kolesnik, *Van der Corput's Method of Exponential Sums*, (Cambridge University Press, Cambridge, 1991)

[Gu] R.K. Guy, *Unsolved Problems in Number Theory*, 2nd ed. (Springer-Verlag, New York, 1994).

[HR] H. Halberstam and H.-E. Richert, *Sieve Methods*, (Academic Press, London, 1974).

[J] M. Jutila, *On numbers with a large prime factor*, II, J. Indian Math. Soc. **38** (1974) 125–130

[KP] S. Konyagin and C. Pomerance, *Primes Recognizable in Deterministic Polynomial Time* (to appear).

[RS] J. B. Rosser and L. Schoenfeld, *Approximate formulae for some functions of prime numbers*, Ill. J. Math. **6** (1962) 64–94.

[Sa1] J.W. Sander, *On prime power divisors of binomial coefficients*, Bull. London Math. Soc. **24** (1992) 140–142.

[Sa2] J.W. Sander, *Prime power divisors of binomial coefficients*, J. reine angew Math **430** (1992), 1–20.

[Sa3] J.W. Sander, *On primes not dividing binomial coefficients*, Proc. Camb. Phil, Soc. **113** (1993), 225–232.

[Sa4] J.W. Sander, email correspondance (25th November 1994).

[Sar] A. Sárközy, *On divisors of binomial coefficients* I, J. Number Theory **20** (1985) 70–80.

[Sch] L. Schoenfeld, *Sharper Bounds for the Chebyshev Functions $\theta(x)$ and $\psi(x)$, II*, Math. Comp. **30** (1976) 337–360.

[SW] R. Scheidler and H.C. Williams, *A method of tabulating the number-theoretic function $g(k)$*, Math. Comp. **59** (1992) 251–257.

[Ti] E.C. Titchmarsh, *The Theory of the Riemann Zeta-function* 2nd ed. (revised by Heath-Brown, D.R.), (Oxford U. Press, New York, 1988).

[Va] J.D. Vaaler, *Some extremal functions in Fourier analysis*, Bull. Amer. Math. Soc. **12** (1985) 183–216.

[Ve] G. Velammal, *Is the binomial coefficient* $\binom{2n}{n}$ *squarefree ?* (to appear).

[W] E.A. Wirsing, *Multiple prime divisors of binomial coefficients* (to appear).

Andrew Granville, U. Of Georgia, Athens, Georgia 30602, USA (andrew@math.uga.edu)

Olivier Ramaré, U. Nancy I, 54 506 Vandoeuvre-les-Nancy, France (ramare@iecn.u-nancy.fr)