Sharp Transitions in Making Squares

Ernie Croot\textsuperscript{1}
School of Mathematics
Georgia Tech
Atlanta, GA 30332-0160
ecroot@math.gatech.edu

Andrew Granville \textsuperscript{2}
Département de mathématiques et de statistique
Université de Montréal
Montréal QC H3C 3J7, Canada
andrew@dms.umontreal.ca

Prasad Tetali
School of Mathematics and College of Computing
Georgia Tech
Atlanta, GA 30332-0160
tetali@math.gatech.edu

\textsuperscript{1}Supported by an NSF grant.
\textsuperscript{2}Partiellement soutenu par une bourse de la Conseil de recherches en sciences naturelles et en génie du Canada.
Abstract

In many integer factoring algorithms, one produces a sequence of integers (created in a pseudo-random way), and wishes to determine a subsequence whose product is a square. A good model for how this sequence is generated is the following process introduced by Pomerance in his 1994 ICM lecture: Select integers \(a_1, a_2, \ldots\), at random from the interval \([1, x]\), until some subsequence has product equal to a square. Estimating the expected stopping time of this process turns out to be a central problem in developing heuristic running time estimates for integer factoring algorithms. Also, if one knows how long the other parts of the algorithm take, one can use such stopping time estimates to determine the optimal choice of algorithm parameters that minimizes the running time.

Here we determine this expected stopping time up to a constant factor, which improves previous estimates due to Pomerance (1994) and Schroeppel (1985), who showed that this stopping time lies in an interval \([z, z^{1+o(1)}]\), for an appropriate \(z = z(x)\). Our result significantly tightens this interval to

\[
((\pi/4)(e^{-\gamma} - o(1))z, (e^{-\gamma} + o(1))z),
\]

where \(\gamma = 0.577\ldots\) is Euler’s constant, which gives \(e^{-\gamma} = 0.5614594\ldots\). Note that this result comes close to proving a sharp threshold for the monotone property of having a square dependence in a random sequence of integers. Our proof uses the first and second moment methods from probabilistic combinatorics, Husimi graphs from statistical physics, and analytical estimates on smooth numbers from number theory.
1 Introduction

Several algorithms for factoring integers \( n \) (including Dixon’s random squares algorithm [2], the quadratic sieve [6], the multiple polynomial quadratic sieve [10], and the number field sieve [1] – see [9] for a nice expository article on factoring algorithms) work by generating a pseudorandom sequence of integers \( a_1, a_2, \ldots \), with each

\[ a_i \equiv b_i^2 \pmod{n}, \]

until some subsequence of the \( a_i \)'s has product equal to a square. Say we have such a subsequence

\[ a_{i_1}, \ldots, a_{i_k}, \]

where \( Y^2 = a_{i_1} \cdots a_{i_k}, \)

and set

\[ X^2 = (b_{i_1} \cdots b_{i_k})^2. \]

Then,

\[ n \mid Y^2 - X^2 = (Y - X)(Y + X), \]

and there is often a fair chance that \( \gcd(n, Y - X) \) is a non-trivial factor of \( n \). If so, then we have factored \( n \).

To address the performance of such algorithms for factoring \( n \), Pomerance [7, 8], in his 1994 ICM lecture, introduced the following problem:

**Pomerance’s Problem.** Select positive integers \( a_1, a_2, \ldots \leq x \) independently at random (using uniform distributions), until a subsequence has product equal to a square. When this occurs, we say that the sequence has a *square dependence*. What is the expected stopping time of this process?

It is clear that Pomerance’s problem is central to the running time analysis of such algorithms; and, if one knew how long the other parts of such algorithms take, one could use the stopping time estimates to determine the optimal choice of algorithm parameters to minimize the running time.

Here we determine this expected stopping time up to a constant factor, which improves previous estimates due to Pomerance and Schroeppel, who showed that this stopping time lies in an interval \([z, z^{1+o(1)}]\), for an appropriate \( z = z(x) \); in particular, we tighten this interval to

\[ [(\pi/4)(e^{-\gamma} - o(1))z, (e^{-\gamma} + o(1))z], \]

where \( \gamma \) is Euler’s constant, giving \( e^{-\gamma} = 0.5614594 \ldots \).

In Pomerance’s problem, let \( T \) be the first time \( t \) such that \( a_1, \ldots, a_t \) has a square dependence. Note that \( T \) is a random variable. We conjecture that \( T \) has a *sharp threshold*; that is, we conjecture:

\[ z \leq T \leq z^{1+o(1)}. \]
Conjecture 1 There exists a function $f(x)$ such that for every $\epsilon > 0$,
\[ \text{Prob}(T \in [(1 - \epsilon)f(x), (1 + \epsilon)f(x)]) = 1 - o(1). \] (1)

In fact, we have a candidate for such a function $f(x)$: we believe
\[ f(x) = e^{-\gamma}J_0(x), \]
for a certain function $J_0(x)$ defined below.

Remark. This $e^{-\gamma}$ is a well-known constant in number theory, and it is the ratio between the proportion of the integers without prime divisors smaller than $y$, to the proportion of primes up to $y$; that is
\[ e^{-\gamma} = \lim_{y \to \infty} \frac{\prod_{2 \leq p < y} (1 - 1/p)}{\pi(y)/y} = \lim_{y \to \infty} (\log y) \prod_{2 \leq p < y} (1 - 1/p). \]

Even so, this is not how we arrive at $e^{-\gamma}$ in our proof; it comes out of some quite complicated combinatorial identities that superficially have little to do with number theory!

We will need some notation and definitions: Let $P(n)$ be the largest prime factor of a given integer $n$. A $y$-smooth integer is an integer $n$ with $P(n) \leq y$. Define
\[ \Psi(x, y) := |\{n \leq x : P(n) \leq y\}| = \text{number of } y\text{-smooths } \leq x. \]

Let
\[ y_0 = y_0(x) := y \text{ value which maximizes } \frac{\Psi(x, y)}{y}; \text{ and let} \]
\[ J_0(x) := \frac{x}{\Psi(x, y_0)} \cdot \frac{y_0}{\log y_0}. \] (2)

In 1985, Schroeppel gave a simple argument to justify that
\[ \text{Prob}(T < (1 + \epsilon)J_0(x)) = 1 - o(1), \]
and in 1994 Pomerance showed that
\[ \text{Prob}(T > J_0(x)^{1-\epsilon}) = 1 - o(1), \]
thus showing that the transition from “unlikely to have a square product” to “almost certain to have a square product” occurs at $J = J_0(x)^{1+o(1)}$. He also asked in [3] whether there is a sharper transition, and the conjecture (1) above is a restatement of this question.

While the theorem of E. Friedgut [3], characterizing a coarse threshold for monotone graph (or more generally symmetric) properties, has been instrumental in proving the existence of a sharp threshold for several graph properties, it is not directly applicable in
the present context – the square dependence problem and other number-theoretic problems are not symmetric – symmetric here refers to being invariant under permutations of the elements (integers) involved. The follow-up strengthening of sorts to Friedgut’s theorem by J. Bourgain (see the appendix to [3]) is in principle applicable in the present context; however, despite some efforts by a few researchers, Conjecture 1 does not seem amenable to these approaches.

We will prove something a little weaker than the above conjecture using methods from combinatorics, analytic and probabilistic number theory:

**Theorem 1** We have that

\[
\text{Prob}(T \in [(\pi/4)(e^{-\gamma} - o(1))J_0(x), (e^{-\gamma} + \epsilon)J_0(x)]) = 1 - o(1).
\]

**Remarks.** Roughly speaking, obtaining the lower bound of the interval involves estimating from above, the expected number of non-trivial subsets whose product is a square, and as such uses the so-called “first moment method”. As to the upper bound: Schroeppel established \((1 + o(1))J_0(x)\) for the stopping time (with probability \(1 - o(1)\)) by looking simply for enough \(y\)-smooth integers among the \(a_i\)’s to ensure a square dependence, and then optimized the choice of \(y\), which turns out to be \(y = y_0\). This is actually how many integer factoring algorithms operate; that is, rather than working with all subsequences of the \(a_1, a_2, \ldots\) that have product equal to a square, they instead only focus on those subsequences where the \(a_i\)’s are \(y\)-smooth. Certain “large prime variations” of these algorithms, in the hopes of taking less time to factor \(n\), work with larger collections of subsequences of the \(a_i\)’s, consisting of numbers of the form \(s_iq_i\), where \(s_i\) is \(y\)-smooth, and \(q_i\) is either 1 or is a prime exceeding \(y\). One can look at further variations where two large primes are used – that is, \(a_i = s_iq_iq_i'\) – or even three or more large primes. Analysing these “large prime variations” quickly descends into a combinatorial mess, and the most difficult part of our paper comes in sorting that mess out, via the theory of Huisimi cacti graphs (see section 4). When this difficult analysis is completed, we arrive at our improved (over Schroeppel) upper bound of \((e^{-\gamma+o(1)})J_0(x)\).

When our process terminates at time \(T\), it is of interest to get some idea of what the subsequence(s) \(I\) (of the numbers \(a_1, \ldots, a_T\)) that has product equal to a square, look(s) like:

**Theorem 2**

a) Given that \(T < (\pi/4)(e^{-\gamma} - o(1))J_0(x)\), with probability \(1 - o(1)\), we have that \(I\) consists of a single number, which is therefore a square.

b) If \(T < (e^{-\gamma} + \epsilon)J_0(x)\) then, with probability \(1 - o(1)\), we have that

\[
y_0 \exp(-(c_3 + \epsilon)\sqrt{\log y_0}) \leq |I| \leq y_0 \exp((c_3 + \epsilon)\sqrt{\log y_0}),
\]

where \(c_3 = \sqrt{2 - \log 2}\). In other words, almost certainly, when the algorithm terminates the square product is composed of \(y_0^{1+o(1)} = J_0(x)^{1/2+o(1)}\) numbers \(a_i\).

c) With probability \(1 - o(1)\) all the elements of \(I\) are

\[
y_0^2 \exp(2(c_3 + \epsilon)\sqrt{\log y_0}) \text{ smooth}.
\]
**Remark.** Conclusion (c) is telling us that smooth numbers arise naturally in the square dependence problem; moreover, it tells us that in algorithms that rely on the square dependencies (such as Dixon’s algorithm) need only consider numbers that are fairly smooth.

One can ask for more precise results:

- Can one describe how the typical set $I$ looks, given that $T \sim cJ_0(x)$ with
  \[(\pi/4)(e^{-\gamma} - \epsilon) < c < e^{-\gamma} + \epsilon?\]

- Is it true that for any function $\phi(x) \to \infty$ we have that, with probability $1 - o(1)$, all the elements of $I$ are
  \[y_0\phi(x) - \text{smooth}?\]

The paper is organized as follows. In Section 2, we derive the necessary technical lemmas involving smooth numbers. In Section 3, we develop some brief intuition for the problem. In Sections 4 and 5, we derive the upper bound and the lower bound in Theorem 1, respectively. Finally, Theorem 2 is proved in Section 6.

## 2 Smooth numbers

### 2.1 Classical smooth number estimates

Throughout let $\pi(y)$ be the number of primes up to $y$. From [4] we have that the estimate

\[\Psi(x, y) = x\rho(u) \left\{ 1 + O\left(\frac{\log(u + 1)}{\log y}\right) \right\} \quad \text{as} \quad x \to \infty \quad \text{where} \quad x = y^u, \quad (4)\]

holds in the range

\[\exp \left((\log \log x)^2\right) \leq y \leq x, \quad (5)\]

where $\rho(u) = 1$ for $0 \leq u \leq 1$, and where

\[\rho(u) = \frac{1}{u} \int_{u-1}^u \rho(t) \, dt \quad \text{for all} \quad u > 1.\]

This function $\rho(u)$ satisfies

\[\rho(u) = \exp(-u + o(u)) \log u;\]

and so

\[\Psi(x, y) = x \exp(-u + o(u)) \log u). \quad (6)\]

Now let

\[L := L(x) = \exp \left(\sqrt{\frac{1}{2} \log x \log \log x} \right).\]
Then, using (6) we deduce that for $\beta > 0$,
\[ \Psi(x, L(x)^{\beta+o(1)}) = xL(x)^{-1/\beta + o(1)}. \]  
(7)

From this one can easily deduce that
\[ y_0(x) = L(x)^{1+o(1)} \text{, and } J_0(x) = y_0^{2-\{1+o(1)\}/\log \log y_0} = L(x)^{2+o(1)}, \]
where $y_0$ and $J_0$ are as in the introduction (see Eqn. (2)). From this we can deduce the following basic estimate, which we will use in later proofs:

Lemma 1 For a constant $\beta > 0$ we have the following: If $y = y_0^{\beta+o(1)}$, then
\[ \frac{\Psi(x, y)}{\Psi(x, y_0)} = \frac{2-\beta-\beta^{-1}+o(1)}{y_0^{\beta+o(1)}}. \]

2.2 Hildebrand-Tenenbaum saddle point method estimates

For any $\alpha > 0$, one has
\[ \Psi(x, y) \leq \sum_{\substack{\alpha \leq \epsilon \leq \epsilon^p(n) \leq y}} (x/n)^\alpha \leq x^\alpha \xi(\alpha, y), \]  
(8)

where
\[ \xi(s, y) = \prod_{p \leq y} \left(1 - \frac{1}{p^s}\right)^{-1}. \]

Define $\alpha = \alpha(x, y)$ to be the solution to
\[ \alpha \log x = \sum_{p \leq y} \frac{\log p}{p^\alpha - 1}. \]  
(9)

By [4, Theorem 1 and (7.19)] we obtain in the range (5) with $u \to \infty$,
\[ \Psi(x, y) \sim \frac{x^\alpha \xi(\alpha, y)}{\alpha \sqrt{2\pi \log x \log y}}. \]
(10)

Let $\xi = \xi(u)$ be the solution to $e^\xi = u\xi + 1$ so that
\[ \xi(u) = \log(u \log u) + \frac{(1 + o(1)) \log \log u}{\log u}, \text{ as } u \to \infty. \]

Note also that $\xi'(u) \sim 1/u$. In the range (5) it turns out that
\[ (1 - \alpha(x, y)) \log y = \xi(u) + O(1/u) \]  
(11)

which implies that
\[ y^{1-\alpha} = e^{\xi(u)} (1 + O(1/u)) = u\xi(u) (1 + O(1/u)). \]  
(12)
So, for
\[ y = L(x)^{\beta + o(1)} = y_0^{\beta + o(1)} \]
we will have
\[ y^{1-\alpha} \sim \beta^{-2} \log y \sim \beta^{-1} \log y_0. \] (13)
By [4, Theorem 3] and (11) above, we have
\[ \Psi \left( \frac{x}{d}, y \right) = \frac{1}{d^{\alpha(x,y)}} \Psi(x, y) \left( 1 + O \left( \frac{1}{u} + \frac{\log y}{y} \right) \right), \quad \text{when } 1 \leq d \leq y \leq \frac{x}{d}. \] (14)

**Proposition 1** Throughout the range (5), for any \( 1 \leq d \leq x \), we have
\[ \Psi \left( \frac{x}{d}, y \right) \leq \frac{1}{d^{\alpha(x,y)}} \Psi(x, y) \{ 1 + o(1) \}, \]
where \( \alpha \) is the solution to (9). In fact,
\[ \Psi \left( \frac{x}{d}, y \right) < \Psi(x, y) \frac{d}{d^{\alpha(x,y)}}, \]
provided that
\[ \log d \gg \log u \log y + \sqrt{u} \log u \log y. \]

**Proof.** By (4), for \( d = y^r \) with \( 0 \leq r \leq u/2 \), we have
\[ \frac{\Psi \left( \frac{x}{d}, y \right) d^{\alpha}}{\Psi(x, y)} = d^{-(1-\alpha)} \rho(u - r)(1 + O(\log(u + 1)/\log y)) \frac{\rho(u)}{\rho(u)}. \]
The logarithm of the main term on the right side is
\[ -(1 - \alpha)r \log y + \log(\rho(u - r)/\rho(u)). \]
Using the fact that \( u = (\log x)/\log y \) this can be rewritten as
\[ r(\xi(u) - (1 - \alpha) \log y) + \left( -\int_{u-r}^{u} \frac{\rho'(v)}{\rho(v)} dv - r \xi(u) \right). \]
The first term is \( O(r/u) \) by (11). Corollary 8.3 of [11] gives that
\[ -\rho'(v)/\rho(v) = \xi(v)(1 + O(1/v)), \] (15)
so that the second term equals
\[ -\int_{0}^{r} (\xi(u) - \xi(u - t)) dt + O(r \log u/u). \]
Now, differentiating \( e^\xi = u\xi + 1 \) we obtain
\[ \xi + u\xi' = \xi' e^\xi = \xi'(u\xi + 1), \]
so that
\[ \xi' = \frac{1}{u - (u - 1)\xi^{-1}} = \frac{1}{u(1 + O(1/\log u))}. \]

Therefore
\[
\int_0^r (\xi(u) - \xi(u - t))dt = \int_0^r (r - v)\xi'(u - v)dv = \left(1 + O\left(\frac{1}{\log u}\right)\right) \int_0^r \frac{(r - v)}{(u - v)}dv
\]
\[
= \left(1 + O\left(\frac{1}{\log u}\right)\right) (r - (r - u)\log(1 - r/u)).
\]

Combining this with the above yields that
\[
\log \left(\frac{\Psi(\frac{u}{y}, y) dx}{\Psi(x, y)}\right) = -\left(1 + O\left(\frac{1}{\log u}\right)\right) (r - (r - u)\log(1 - r/u))
\]
\[
+ O\left(\frac{r \log u}{u}\right)
\]
\[
= -\frac{r^2}{2u} \left\{1 + O\left(\frac{r}{u} + \frac{1}{\log u} + \frac{\log u}{r}\right)\right\} + O\left(\frac{\log(u + 1)}{\log y}\right).
\]

From (16) and the first equation here we find that this is negative provided
\[
\log u + \sqrt{u \log u/\log y} \ll r \ll u/2.
\]

and is \(o(1)\) in the complementary range.

If \(d > \sqrt{x}\) we simply iterate the above result: The proposition follows by noting that \(\alpha(x, y)\) is a decreasing function in \(x\) for fixed \(y\), by definition. \(\Box\)

Finally, we will require the following lemma, which is in one sense stronger, and in another sense weaker, than Lemma 1.

**Lemma 2** We have
\[
\frac{\Psi(x, y)}{y} \leq \frac{(2/e^2 - \epsilon)\Psi(x, y_0)}{y_0}
\]
for all \(y\) outside of the range
\[
y_0 \exp(-c_3 + \epsilon)\sqrt{\log y_0} \leq y \leq y_0 \exp((c_3 + \epsilon)\sqrt{\log y_0}).
\]  

**Proof.** Let \(x = y_0^{u_0}\). Define \(g(u) = g(x) = \log \rho(u) - u^{-1}\log x\). By (4) we have \(\log(\Psi(x, y)/y) = g(u) + O(1/u)\), provided \(\log y \sim \log L\). Select \(u_1\) to maximize \(g(u)\). Therefore \(g(u_1) \geq g(u_0)\) by definition of \(u_1\); and \(g(u_0) \geq g(u_1) + O(1/u_0)\) by the definition of \(u_0\) and the above estimate; therefore \(g(u_0) = g(u_1) + O(1/u_0)\).

By (15), we have \(g'(v) = \rho'(v)/\rho(v) + v^{-2}\log x = -\xi(v) + v^{-2}\log x + O(\log v/v);\) so that, for \(t = O(u_1/\log u_1)\),
\[
g'(u_1 + t) = g'(u_1 + t) - g'(u_1)
\]
\[ \begin{align*}
= \xi(u_1) - \xi(u_1 + t) + \left( \frac{1}{(u_1 + t)^2} - \frac{1}{u_1^2} \right) \log x + O \left( \frac{\log u_1}{u_1} \right) \\
= O \left( \frac{t + \log u_1}{u_1} \right) - 2tu_1^{-3} \log x(1 + O(t/u_1)) \\
= -2t \frac{\xi(u_1)}{u_1} + O \left( \frac{t + \log u_1}{u_1} \right),
\end{align*} \]

since \( 0 = g'(u_1) = -\xi(u_1) + u_1^{-2} \log x + O(\log u_1/u_1) \). Therefore

\[
g(u_1) - g(u_1 + T) = -\int_0^T g'(u_1 + t) dt = \frac{T^2}{u_1}(\xi(u_1) + O(1)) + O \left( \frac{T \log u_1}{u_1} \right),
\]

for \( T = O(u_1 / \log u_1) \). We deduce that \( u_0 = u_1 + O(1) \), and

\[ g(u) < g(u_0) - \log(e^2/2 + \epsilon) \text{ for } |u - u_0| > (c_3 + \epsilon) \sqrt{u_0 / \log u_0}, \]

which is the desired result. \( \square \)

3 Some simple observations

3.1 A heuristic analysis

Let \( M = \pi(y) \) and

\[ p_1 = 2 < p_2 = 3 < \ldots < p_M \]

be the primes up to \( y \). Any \( y \)-smooth integer

\[ p_1^{e_1}p_2^{e_2} \cdots p_M^{e_M} \]

gives rise to the element \((e_1, e_2, \ldots, e_M)\) of the vector space \( \mathbb{F}_2^M \) and, for the sake of argument, we will assume each element of \( \mathbb{F}_2^M \) is equally likely to be obtained when we randomly choose integers from \([1, x]\), which is patently untrue but may give a some idea of the truth nonetheless.

With this heuristic our algorithm translates into one where we select \( v_1, v_2, \ldots \) at random in \( \mathbb{F}_2^M \), stopping at \( v_T \) once \( v_1, v_2, \ldots, v_T \) are linearly dependent.

Now if \( v_1, v_2, \ldots, v_{T-1} \) are linearly independent they generate a subspace \( S_T \) of dimension \( T - 1 \), which contains \( 2^{T-1} \) elements (if \( 1 \leq T \leq M + 1 \)). Then the probability that \( v_1, v_2, \ldots, v_T \) are linearly dependent is the same as the probability that \( v_T \) belongs to \( S_T \), which is \( 2^{T-1}/2^M \). Thus the expectation of \( T \) is

\[
\sum_{\ell=1}^{M+1} \ell \frac{2^{\ell-1}}{2^M} \prod_{i=1}^{\ell-1} \left( 1 - \frac{2^{i-1}}{2^M} \right) = \prod_{j=1}^{M+1} \left( 1 - \frac{1}{2^j} \right) \sum_{j=0}^{M} \frac{M+1-j}{2^j} \prod_{i=1}^{j} \left( 1 - \frac{1}{2^i} \right)^{-1}
\]

\[ = M - .60669515 \ldots \text{ as } M \to \infty. \]

(By convention, empty products have value 1.)
Thus, $|T - M|$ has expected value $O(1)$; furthermore, for $n \geq 1$ we have
\[
\operatorname{Prob}(|T - M| > n) = \sum_{\ell \geq n+1} \operatorname{Prob}(T = M - \ell) < \sum_{\ell \geq n+1} 2^{-\ell-1} = 2^{-n-1}.
\]
Thus if $\phi(t) \to \infty$ as $t \to \infty$, then
\[
\operatorname{Prob}(T \in [M - \phi(M), M]) = 1 - o(1),
\]
which means that this simplified problem has a very sharp transition function indeed, suggesting it might be so in the original problem.

3.2 No large primes, I

Suppose that we have selected integers $a_1, a_2, \ldots, a_T$ at random from $[1, x]$, stopping at $T$ since there is a non-empty subset of these integers whose product is a square. Let $q$ be the largest prime that divides this square. Then either $q^2$ divides one of $a_1, a_2, \ldots, a_T$, or $q$ divides at least two of them. The probability that $p^2$ divides at least one of $a_1, a_2, \ldots, a_T$, for a given prime $p$, is $\leq T/p^2$; and the probability that $p$ divides at least two of $a_1, a_2, \ldots, a_T$ is $\leq (T^2)/p^2$. Thus
\[
\operatorname{Prob}(q > T^2) \ll T^2 \sum_{p > T^2} \frac{1}{p^2} \ll \frac{1}{\log T},
\]
by the Prime Number Theorem.

By Pomerance’s result we know that $T \to \infty$ with probability $1 + o(1)$; and so the largest prime that divides the square is $\leq T^2$ with probability $1 - o(1)$. We will improve this result later by more involved arguments.

4 Proof of the upper bound on $T$ in Theorem 1

In this section we will prove that for all $\epsilon > 0$,
\[
\operatorname{Prob}(T < (e^{-\gamma} + \epsilon)J_0(x)) = 1 - o_\epsilon(1)
\]

4.1 Schroeppel’s argument

Given a sequence
\[
a_1, \ldots, a_J \leq x\]
of randomly chosen positive integers, let
\[
p_1 = 2 < p_2 = 3 < \ldots < p_{\pi(x)}
\]
denote the primes up to $x$ and in a similar way as was done in section 3.1, construct the $J$-by-$\pi(x)$ matrix $A$ where
\[
a_{ij} = \prod_{1 \leq j \leq \pi(x)} p_{ij}^{A_{ij}}.
\]
Then, a given subsequence of the $a_i$ has product a square if their corresponding vectors sums to the 0 vector modulo 2; and, this happens if and only if $\text{rank}(A) < J$ (where here the rank is the $\mathbb{F}_2$ rank of $A$).

Schroppel’s idea was to focus only on those rows having no 1’s beyond the $\pi(y_0)$th column, which corresponds to $y_0$-smooth integers. If we let $S(y_0)$ denote the set of all such rows, then Schroppel’s approach amounts to showing that

$$|S(y_0)| > \pi(y_0)$$

holds with probability $1 - o(1)$ for a certain $J = (1 + \epsilon)J_0$, where $J_0$ and $y_0$ are as defined in (2). If this inequality holds, then the $|S(y_0)|$ rows are linearly dependent mod 2, and therefore some subset of them sums to the 0 vector mod 2.

To implement Schroppel’s idea, define the random variable $Y_j = 1$ if $a_j$ is $y$-smooth, and $Y_j = 0$ otherwise, where $y$ will be chosen later. Let

$$N = Y_1 + \cdots + Y_J,$$

which is the number of $y$-smooths among $a_1, \ldots, a_J$. The probability that any such integer is $y$-smooth, that is that $Y_j = 1$, is $\psi(x, y)/x$; and so,

$$\mathbb{E}(N) = \frac{J\psi(x, y)}{x}.$$

Since the $Y_i$ are independent, we also have

$$V(N) = \sum_i (\mathbb{E}(Y_i^2) - \mathbb{E}(Y_i)^2) = \sum_i (\mathbb{E}(Y_i) - \mathbb{E}(Y_i)^2) \leq \frac{J\psi(x, y)}{x}.$$ 

Thus, selecting $J = (1 + \epsilon)x\pi(y)/\psi(x, y)$, we have, with probability $1 + o_\epsilon(1)$, that

$$N = (1 + \epsilon + o(1))\pi(y) > \pi(y).$$

Therefore, there must be some non-empty subset of the $a_i$’s whose product is a square. Taking $y = y_0$ we deduce that

$$\text{Prob}(T < (1 + \epsilon)J_0(x)) = 1 - o_\epsilon(1).$$

### 4.2 Improvement

Our idea is to work with more rows than just those in $S(y_0)$; in fact, we work with all the rows of $A$. We let $A_{y_0}$ be the matrix gotten by deleting the first $\pi(y_0)$ columns of $A$. Note that the rows corresponding to smooth numbers will be 0 rows in this new matrix.

If

$$\text{rank}(A_{y_0}) < J - \pi(y_0),$$

then

$$\text{rank}(A) \leq \text{rank}(A_{y_0}) + \pi(y_0) < J,$$
which therefore means that the rows of $A$ are dependent over $\mathbb{F}_2$, and thus $a_1, ..., a_J$ has a square dependence.

To establish (18) we will show that with probability $1 - o(1)$,

$$J - \text{rank}(A_{y_0}) > \pi(y_0),$$

for $J = (\eta + o(1))J_0$, for a certain $\eta > 0$. The way we will achieve this is to break these $J$ rows down into subcollections, where the primes $p > y_0$ involved in each subcollection are disjoint from the primes involved in the other subcollections. Say that there are $k$ subcollections, and say that the $i$th subcollection consists of $r_i$ rows; further, suppose that these rows have rank $\ell_i$. Then,

$$J - \text{rank}(A_{y_0}) \geq \sum_{i=1}^{k} (r_i - \ell_i).$$

So, we will aim to show that for $J = (\eta + o(1))J_0$,

$$\sum_{i=1}^{k} (r_i - \ell_i) > \pi(y_0)$$

holds with probability $1 - o(1)$.

Our job is thus to find just the right value of $0 < \eta < 1$ which guarantees that this last inequality holds with high probability; it will turn out that the value of $\eta$ which does this is

$$\eta_0 := e^{-\gamma},$$

where $\gamma = 0.57721566...$ is Euler’s constant. The way we will arrive at this is that there is a certain function $f(\eta)$ given in (35) below (well, actually we mean the function $f^+(\eta)$), for which we will show that if

$$\lim_{\eta \to \eta_0} f(\eta) = 1,$$

then (19) holds for $J = (\eta_0 + o(1))J_0$; and then we will show that $\eta_0 = e^{-\gamma}$ satisfies (20).

### 4.3 Appropriate submatrices

It turns out that not all submatrices of $A_{y_0}$ with dependent rows are equally likely to occur; and, in fact, some of these matrices will occur only very rarely, meaning that they can be safely ignored. In this subsection we introduce some notation for these submatrices $M$, and in the next subsection we work out which submatrices $M$ occur often enough to be worthy of consideration.

Let $M$ be an $r$-by-$\ell$ matrix, with $(i, j)$th entry $e_{i,j} \in \mathbb{F}_2$ for $1 \leq i \leq r$, $1 \leq j \leq \ell$, with no zero rows. Let

$$N(M) = \sum_{i,j} e_{i,j}$$
denote the number of 1’s in $M$, and
\[ m_j = \sum_i e_{i,j} \]
denote the number of 1’s in column $j$; we require each $m_j \geq 2$ (else the prime corresponding to that column cannot participate in a product that gives a square). We require that $M$ is transitive. That is, for any $j$, $2 \leq j \leq \ell$ there exists a sequence of row indices $i_1, \ldots, i_g$, and column indices $j_1, \ldots, j_{g-1}$, such that
\[ e_{i_1,1} = e_{i_g,j} = 1; \text{ and, } e_{i_h,j_h} = e_{i_h+1,j_h} = 1 \text{ for } 1 \leq h \leq g-1. \]
In other words we do not study $M$ if, after a permutation, it can be split into a block diagonal matrix with more than one block, since this would correspond to independent squares.

It is convenient to keep in mind two reformulations:

**Integer version:** Given primes $p_1 < p_2 < \ldots < p_\ell$, we assign, to each row, an integer
\[ n_i = \prod_{1 \leq j \leq \ell} p_j^{e_{i,j}}, \text{ for } 1 \leq i \leq r. \]

**Graph version:** Take a graph $G(M)$ with $r$ vertices, where $v_i$ is adjacent to $v_I$ with an edge of colour $p_j$ if $p_j$ divides both $n_i$ and $n_I$ (or, equivalently, $e_{i,j} = e_{I,j} = 1$). Notice that transitive is equivalent to connected in this formulation.

### 4.4 A Property of $\Delta(M)$

Let us re-order the rows of $M$ so that, in the graph theory version, each new vertex connects to the graph that we already have, which is always possible as the overall graph is connected. Let
\[ \ell_I = \# \{ j : \exists i \leq I \text{ with } e_{i,j} = 1 \}, \]
the number of columns with a 1 in or before the $I$th row, and
\[ N_I := \sum_{i \leq I, j \leq \ell} e_{i,j}, \]
the number of 1s up to and including in the $I$th row. Define
\[ \Delta_I = N_I - I - \ell_I + 1; \]
and then
\[ \Delta(M) := N(M) - r(M) - \ell(M) + 1 = \Delta_r, \]
where $r = r(M)$ is the number of rows of $M$. 

13
Now $N_1 = \ell_1$ and therefore $\Delta_1 = 0$. Let us consider the transition when we add in the $(I+1)$th row. The condition that each new row connects to what we already have means that
$$\ell_{I+1} - \ell_I \leq N_{I+1} - N_I - 1;$$
and so
$$\Delta_{I+1} = N_{I+1} - I - \ell_{I+1} = N_I - I - \ell_I + (N_{I+1} - N_I) - (\ell_{I+1} - \ell_I) \geq N_I - I - \ell_I + 1 = \Delta_I.$$
Therefore
$$\Delta(M) = \Delta_r \geq \Delta_{r-1} \geq \ldots \geq \Delta_2 \geq \Delta_1 = 0. \quad (21)$$

4.5 We Only Need Consider the Case $\Delta(M) = 0$

We wish to count how often we expect to get the integers $\prod_{1 \leq j \leq \ell} p_{j}^{e_{i,j}}$ times $y_0$-smooth, $i = 1, \ldots, r$, in our sequence of integers $a_1, \ldots, a_J$. This count is
$$\sim \binom{J}{r} |\text{Orbit}_{\text{Rows}}(M)| \prod_{1 \leq i \leq r} \Psi(x/\prod_{1 \leq j \leq \ell} p_{j}^{e_{i,j}}, y_0), \quad (22)$$
where by $\text{Orbit}_{\text{Rows}}(M)$ we mean the set of distinct matrices produced by permuting the rows of $M$.

Now, since $r$ is to be thought of as fixed as we let $x$ and $J$ tend to infinity, we have that
$$\binom{J}{r} \sim \frac{J^r}{r!};$$

furthermore, we know that\footnote{This is a consequence of the “Orbit-Stabilizer Theorem” from elementary group theory. It follows from the fact that the cosets of $\text{Aut}_{\text{Rows}}(M)$ in the permutation group on the $r$ rows of $M$, correspond to the distinct matrices (orbit elements) gotten by performing row interchanges on $M$.}
$$r! = |\text{Orbit}_{\text{Rows}}(M)| \cdot |\text{Aut}_{\text{Rows}}(M)|$$

It follows that the quantity in (22) is
$$= \frac{J^r}{|\text{Aut}_{\text{Rows}}(M)|} \prod_{1 \leq i \leq r} \frac{\Psi(x/\prod_{1 \leq j \leq \ell} p_{j}^{e_{i,j}}, y_0)}{x} \prod_{1 \leq i \leq r} \frac{1}{m_j\cdot p_j}, \quad (23)$$
where $m_j = \sum_i e_{i,j} \geq 2,$ and $\text{Aut}_{\text{Rows}}(M)$ denotes the number of ways to obtain exactly the same matrix by permuting the rows (this corresponds to permuting identical integers...}
that occur). Summing the last quantity in (23) over all $y_0 < p_1 < p_2 < \ldots < p_\ell$, we obtain, by the prime number theorem,

$$\sim \frac{(\eta \pi(y_0))^r}{|\text{Aut}_{\text{Rows}}(M)|} \int_{y_0 < v_1 < v_2 < \ldots < v_\ell} \prod_{1 \leq j \leq \ell} \frac{dv_j}{v_j^{m_j \alpha_j} \log v_j}$$

$$\sim \frac{\eta^r \pi(y_0)^{r+\ell-\sum_j m_j}}{|\text{Aut}_{\text{Rows}}(M)|} \int_{1 < t_1 < t_2 < \ldots < t_\ell} \prod_{1 \leq j \leq \ell} \frac{dt_j}{t_j^{m_j \alpha_j}}$$

$$= \frac{\eta^r \pi(y_0)^{1-\Delta(M)}}{|\text{Aut}_{\text{Rows}}(M)|} \prod_{1 \leq j \leq \ell} \frac{1}{\nu_j},$$

(24)

where

$$\nu_j := \sum_{i=j}^\ell (m_i - 1),$$

In the second row we used the approximation $\log v_j \sim \log y_0$ (because this range of values of $v_j$ gives the main contribution to the integral); the fact that for $v_j$ in this range

$$v_j^{\alpha_j} \sim v_j / \log y_0;$$

and, finally, we made the substitution $t_j = v_j / y_0$.

Evidently, the last row in (24) is only significant if $\Delta(M) \leq 0$ (any single $M$ with $\Delta(M) \leq 0$ gives a larger contribution than all the other $M$ combined); so by (21) we see that $\Delta(M) = 0$. We deduce that each $\Delta_I = 0$; and so

$$\ell_{I+1} - \ell_I = N_{I+1} - N_I - 1$$

for each $I$; that is each new vertex connects with a unique colour to previous vertices. In other words all but one of the primes used in $n_{I+1}$ is new. This can actually be seen easily, in retrospect, in the formula $\ell = N - (r - 1)$ – that is, we take the total number of $i$’s, subtract the unique prime that connects each new row with the previous rows, and this gives us all of the primes used.

So the set $\mathcal{M}$ of “appropriate matrices” is the set of such matrices $M$ with $\Delta(M) = 0$; that is

$$N(M) = r(M) + \ell(M) - 1.$$

(25)

We claim that if $M \in \mathcal{M}$ then all cycles in its graph are monochromatic: If not, then consider a minimal cycle in the graph, where not all the edges are of the same color. We first show that, in fact, each edge in the cycle has a different color. To see this, we start with a cycle where not all edges are of the same color, but where at least two edges have the same color. Say we arrange the vertices $v_1, \ldots, v_k$ of this cycle so that the edge $(v_1, v_2)$ has the same color as $(v_j, v_{j+1})$, for some $2 \leq j \leq k - 1$, or the same color as $(v_k, v_1)$. If we are in the former case, then we reduce to the smaller cycle $v_2, v_3, \ldots, v_j$, where not all edges have the same color; and, if we are in the latter case, we reduce to the smaller cycle $v_2, v_3, \ldots, v_k$, where again not all the edges have the same color. Thus,
if the cycle has not all edges of the same color, but has more than one edge of the same color, then it cannot be a minimal cycle.

Now let \( I \) be the number of vertices in our minimal cycle of different colored edges, and reorder the rows of \( M \) so that this cycle appears as the first \( I \) rows. \(^4\) Then

\[
N_I \geq 2I + (\ell_I - I) = \ell_I + I.
\]

The term \( 2I \) accounts for the fact that each prime corresponding to a different colored edge in the cycle must divide at least two members of the cycle, and the \( \ell_I - I \) accounts for the remaining primes that divide members of the cycle (that don’t correspond to the different colored edges). This then gives \( \Delta_I \geq 1 \); and thus by (21) we have \( \Delta(M) \geq 1 \), a contradiction. We conclude that every cycle in our graph is monochromatic.

### 4.6 Rank \((M) = \ell(M)\)

Next we show that if \( M \in \mathcal{M} \) then

\[
\text{rank}(M) = \ell(M),
\]

by induction on \( \ell \).

For \( \ell = 0, 1 \) this is trivial. Otherwise, as there are no cycles the graph must end in a “leaf”; that is a vertex of degree one. Suppose this corresponds to row \( r \) and color \( \ell \). We now construct a new matrix \( M' \) which is matrix \( M \) less column \( \ell \), and any rows that only contained a 1 in the \( \ell \)th column. The new graph now consists of \( m_\ell - 1 \) disjoint subgraphs, each of which corresponds to an element of \( \mathcal{M} \). Thus the rank of \( M \) is given by 1 (corresponding to the \( r \)th row, which acts as a pivot element in Gaussian elimination on the \( \ell \)th column) plus the sum of the ranks of new connected subgraphs. By the induction hypothesis, they each have rank equal to the number of their primes, thus in total we have \( 1 + (\ell - 1) = \ell \), as claimed.

### 4.7 The key identity

In this subsection we will show the following crucial identity concerning the set \( \mathcal{M} \) of matrices with the property given by (25).

**Lemma 3** For any set \( R \) of rows, let \( M = M_R \) be the matrix created by these rows (allowing multiplicity). If \( M_R \in \mathcal{M} \) then

\[
\sum_{S \subseteq R, \ M_S \in \mathcal{M}} (-1)^{N(S)} = r(M) - \text{rank}(M).
\] \( \text{(26)} \)

Furthermore, if we let \( N \geq 2 \) be an even integer, then we have the overestimate

\[
\sum_{S \subseteq R, N(S) \leq N, \ M_S \in \mathcal{M}} (-1)^{N(S)} \geq r(M) - \text{rank}(M),
\] \( \text{(27)} \)

\(^4\)This we are allowed to do, because the connectivity of successive rows can be maintained, and because we will still have \( \Delta(M) = 0 \) after this permutation of rows.
and for \( N \geq 3 \) odd we have the underestimate
\[
\sum_{S \subset R, N(S) \leq N, M_S \in M} (-1)^{N(S)} \leq r(M) - \text{rank}(M). \tag{28}
\]

**Proof.** It is easy to show that the lemma holds in the case where \(|R| = 2\). Assume that, for proof by induction, it holds for all \( R \) satisfying \(|R| \leq r - 1\). We will now prove the lemma for \(|R| = r\).

If \( M_R \in \mathcal{M} \) then the last row cannot add in any new primes; but we have seen that any row after the first uses only one old prime. Therefore the last row corresponds to an integer that is a prime.

We just proved that there exist integers in our set (i.e. the integers which correspond to the rows of \( M \)) that are prime; and such primes must divide at least one other integer of our set, as \( M_R \in \mathcal{M} \). Call this prime \( p_\ell \).

**Case 1: \( p_\ell \) divides at least three elements from our set.**

In this case we partition the set of rows into three subsets: \( R_0 \), the rows without a 1 in the \( \ell \)th column; \( R_1 \), the rows with a 1 in the \( \ell \)th column, but no other 1s (that is, rows which correspond to the prime \( p_\ell \)); and \( R_2 \), the rows with a 1 in the \( \ell \)th column, as well as other 1s. Note that \(|R_1| \geq 1\) and \(|R_1| + |R_2| \geq 3\) by hypothesis.

Write each \( S \subset R \) with \( M_S \in \mathcal{M} \) as \( S = S_0 \cup S_1 \cup S_2 \) where \( S_i \subset R_i \). If we fix \( S_0 \) and \( S_2 \) with \(|S_2| \geq 2\) then \( S_0 \cup S_2 \in \mathcal{M} \) if and only if \( S_0 \cup S_1 \cup S_2 \in \mathcal{M} \) for any \( S_1 \subset R_1 \). Therefore the contribution of all of these \( S \) to the sum in (26) is
\[
(-1)^{N(S_0) + N(S_2)} \sum_{S_1 \subset R_1} (-1)^{|S_1|} = (-1)^{N(S_0) + N(S_2)}(1 - 1)^{|R_1|} = 0 \tag{29}
\]

Now consider those sets \( S \) with \(|S_2| = 1\). In this case we must have \(|S_1| \geq 1\) and equally we have \( S_0 \cup \{p_\ell\} \cup S_2 \in \mathcal{M} \) if and only if \( S_0 \cup S_1 \cup S_2 \in \mathcal{M} \) for any \( S_1 \subset R_1 \) with \(|S_1| \geq 1\). Therefore the contribution of all of these \( S \) to the sum in (26) is
\[
(-1)^{N(S_0) + N(S_2)} \sum_{S_1 \subset R_1, |S_1| \geq 1} (-1)^{|S_1|} = (-1)^{N(S_0) + N(S_2)}((1 - 1)^{|R_1|} - 1)
\]
\[
= (-1)^{N(S_0 \cup \{p_\ell\} \cup S_2)}. \tag{30}
\]

Regardless of whether \(|S_2| = 1\) or \(|S_2| \geq 2\), we get that if we truncate the sums (29) or (30) to all those \( S_1 \subset R_1 \) with
\[
N(S_1) = |S_1| \leq N - N(S_0) - N(S_2),
\]
then the total sum is \( \leq 0 \) if \( N \) is odd, and is \( \geq 0 \) if \( N \) is even; furthermore, note that we get that these truncations are 0 in two cases: If \( N - N(S_0) - N(S_2) \leq 0 \) (which means that the
above sums are empty, and therefore 0 by convention), or if \( N - N(S_0) - N(S_2) \geq N(R_1) \) (which means that we have the complete sum over all \( S_1 \subset R_1 \)).

It remains to handle those \( S \) where \( |S_2| = 0 \). We begin by defining certain sets \( H_j \) and \( T_j \): If the elements of \( R_2 \) correspond to the integers \( h_1, \ldots, h_k \) then let \( H_j \) be the connected component of the subgraph containing \( h_j \), of the graph obtained by removing all rows divisible by \( p\ell \) except \( h_j \). Let \( T_j = H_j \cup \{ p\ell \} \). Note that if \( S_2 = \{ h_j \} \) then \( S_0 \cup \{ p\ell \} \cup S_2 \subset T_j \) (in the paragraph immediately above).

Note that if \( S \) has \( |S_2| = 0 \), then \( S = S_0 \cup \{ p\ell, h_j \} \) and therefore, combining all of this information,

\[
\sum_{S_1 \subset R_1} (-1)^{|S_1|} = (1 - 1)^{|R_1|} = 1 = |R_1| - 1.
\]

Furthermore, if we truncate this sum to all those \( S_1 \) satisfying

\[
N(S_1) = |S_1| \leq N,
\]

then the sum is \( |R_1| - 1 \) if \( N \geq 2 \) is even, and the sum is \( |R_1| - 1 \) if \( N \geq 3 \) is odd.

Finally note that if \( S \subset T_j \) with \( M_S \in \mathcal{M} \) then either \( |S_2| = 0 \) or \( S = S_0 \cup \{ p\ell, h_j \} \) and therefore, combining all of this information,

\[
\sum_{S \subset R, M_S \in \mathcal{M}, N(S) \leq N} (-1)^{N(S)} = |R_1| - 1 + \sum_{j=1}^{k} \sum_{S \subset T_j, M_S \in \mathcal{M}} (-1)^{N(S)} = |R_1| - 1 + \sum_{j=1}^{k} (r(T_j) - \ell(T_j))
\]

by the induction hypothesis (as each \( |T_j| < |M| \)). Also by the induction hypothesis, along with what we worked out above for \( N \) even and odd, in all possibilities for \( |S_2| \) (i.e. \( |S_2| = 0, 1 \) or exceeds 1), we have that for \( N \geq 3 \) odd,

\[
\sum_{S \subset R, M_S \in \mathcal{M}, N(S) \leq N} (-1)^{N(S)} \leq |R_1| - 1 + \sum_{j=1}^{k} (r(T_j) - \ell(T_j));
\]

and for \( N \geq 2 \) even,

\[
\sum_{S \subset R, M_S \in \mathcal{M}, N(S) \leq N} (-1)^{N(S)} \geq |R_1| - 1 + \sum_{j=1}^{k} (r(T_j) - \ell(T_j)).
\]

The \( T_j \) less the rows \( \{ p\ell \} \) is a partition of the rows of \( M \) less the rows \( \{ p\ell \} \), and so

\[
\sum_j (r(T_j) - 1) = r(M) - |R_1|.
\]
The primes in $T_j$ other than $p_\ell$ is a partition of the primes in $M$ other than $p_\ell$, and so
\[ \sum_j (\ell(T_j) - 1) = \ell(M) - 1. \]
Combining this information gives (26), (27), and (28).

**Case 2 :** $p_\ell$ divides exactly two elements from our set.
Suppose these two elements are $n_r = p_\ell$ and $n_{r-1} = p_\ell q$ for some integer $q$. If $q = 1$ this is our whole graph and (26), (27) and (28) all hold, so we may assume $q > 1$. If $n_j \neq q$ for all $j$, then we create $M_1 \in \mathcal{M}$ with $r - 1$ rows, the first $r - 2$ the same, and with $n_{r-1} = q$. We have
\[ N(M_1) = N(M) - 2, \quad r(M_1) = r(M) - 1, \quad \ell(M_1) = \ell(M) - 1. \]

We claim that there is a 1-1 correspondence between the subsets $S \subset R(M)$ with $M_S \in \mathcal{M}$ and the subsets $T \subset R(M_1)$ with $(M_1)_T \in \mathcal{M}$. The key observation to make is that $p_\ell \in S$ (ie row $r$) if and only if $p_\ell q \in S$ (ie row $r - 1$), since $M_S \in \mathcal{M}$. Thus if rows $r - 1$ and $r$ are in $S$ then $S$ corresponds to $T$ (ie $T = S_1$), which we obtain by replacing rows $r - 1$ and $r$ of $S$ by row $r - 1$ of $T$ which corresponds to $q$. Otherwise we let $S = T$. Either way $(-1)^{N(S)} = (-1)^{N(T)}$ and so
\[ \sum_{S \subset R, \ell(M_S) = \ell(M)} (-1)^{N(S)} = \sum_{T \subset R(M_1), (M_1)_T \in \mathcal{M}} (-1)^{N(T)} = r(M_1) - \ell(M_1) = r(M) - \ell(M), \]
by the induction hypothesis. Further, we have that for $N$ even,
\[ \sum_{S \subset R, N(S) \leq N, \ell(M_S) = \ell(M)} (-1)^{N(S)} \geq r(M) - \ell(M). \]
The analogous inequality holds in the case where $N$ is odd. Thus, we have that (26), (27) and (28) all hold.

Finally, suppose that $n_j = q$ for some $j$, say $n_{r-2} = q$. Then $q$ must be prime else there would be a non-monochromatic cycle in $M \in \mathcal{M}$. But since prime $q$ is in our set it can only divide two of the integers of the set (by our previous deductions) and these are $n_{r-2}$ and $n_{r-1}$. However this is then the whole graph and we observe that (26), (27), and (28) all hold.

### 4.8 Counting Configurations I

For later arguments we will need to have some precise estimates for how often a particular matrix $M \in \mathcal{M}$ occurs as a connected component in $A_{y_0}$, which is just what the following lemma gives:
Lemma 4. Suppose that $M \in \mathcal{M}$. Then, the expected number of times that $M$ occurs as a connected component of $a_1, \ldots, a_J$, $J = \eta J_0$, among those $a_i$ that are $c y_0$-smooth (where the primes corresponding to column of $M$ lie in $[y_0, c y_0]$) is

$$\sim \frac{\eta^{-\ell} e^{-\ell \pi(y_0)}}{|Aut(M)|} \prod_{j=1}^{\ell} \left( (m_j - 2)! \sum_{i=0}^{m_j-2} e^{-i \eta} e^{-m_j/2} - e^{-c \eta} e^{-i} \right).$$

Proof. Let $r$ be the number of rows of $M$, and let $\ell$ be the number of columns of $M$. Given that $b$ is a randomly chosen $c y_0$-smooth number $\leq x$, the probability that it is coprime to the list of primes $q_1, \ldots, q_\ell \in (y_0, c y_0)$ is, by (14),

$$\psi(x, c y_0) - \sum_{j \geq 1} (-1)^{j+1} \sum_{\{p_1, \ldots, p_j\} \subset \{q_1, \ldots, q_\ell\}} \psi(x/p_1 \cdots p_j, c y_0) \psi(x, c y_0)$$

$$= 1 - \frac{\sum_{i=1}^{\ell} \psi(x/q_i, c y_0)}{\psi(x, c y_0)} + O(2^\ell / \pi(y_0)^2).$$

Again by (14) we have that for fixed $c$ this is

$$1 - \sum_{i=1}^{\ell} \frac{1}{q_i} + o \left( \frac{1}{\pi(y_0)} \right).$$

Now suppose that $N$ satisfies

$$r \leq N \ll \pi(y_0),$$

and that $b_1, b_2, \ldots, b_N$ are independently, uniformly chosen $c y_0$-smooth numbers. The probability that the numbers $b_{r+1}, b_{r+2}, \ldots, b_N$ are all coprime to $q_1, \ldots, q_\ell$ is

$$\left( 1 - \sum_{i=1}^{\ell} \frac{1}{q_i} + o \left( \frac{1}{\pi(y_0)} \right) \right)^{N-r} \sim \exp \left( -(N-r) \log(y_0) \sum_{i=1}^{\ell} \frac{1}{q_i} \right).$$

If $r = o(N)$, then this last quantity is

$$\sim y_0^{-N \sum_{i=1}^{\ell} 1/q_i}.$$
Next, we work out the expected number of times that \( b_1, \ldots, b_r \) correspond to a connected component of the whole graph formed using \( b_1, \ldots, b_N \), where this connected component corresponds to the matrix \( M \). From what we worked out in (22), (23) and (24) we know that the expected number of times that a sequence of randomly chosen numbers \( a_1, \ldots, a_r \leq x \) corresponds to the matrix \( M \), where the \( j \)th column of \( M \) is to correspond to the particular prime \( q_j \), is

\[
\sim |\text{ Orbit}_{\text{Rows}}(M)| \prod_{1 \leq i \leq r} \frac{\psi(x/\prod_{1 \leq j \leq \ell} q_j^{e_{i,j}}, y_0)}{x},
\]

where the \( i,j \) entry of \( M \) is \( e_{i,j} \). Using (14) this is

\[
\sim |\text{ Orbit}_{\text{Rows}}(M)| (\Psi(x, y_0)/x)^r \prod_{1 \leq j \leq \ell} \frac{1}{q_j^{m_j}},
\]

where \( m_j \) is the sum of entries in the \( j \)th column of \( M \). Now, if we condition this on the assumption that all the \( a_1, \ldots, a_r \) are \( cy_0 \)-smooth, then this expectation becomes

\[
\sim |\text{ Orbit}_{\text{Rows}}(M)| C^r \prod_{1 \leq j \leq \ell} \frac{1}{q_j^{m_j}} \sim |\text{ Orbit}_{\text{Rows}}(M)| e^{-r} \prod_{1 \leq j \leq \ell} \frac{1}{q_j^{m_j}},
\]

where \( C \sim 1/c \) is the proportion of \( y_0 \)-smooths to \( cy_0 \)-smooths in \([1, x]\).

It now follows that the probability that \( b_1, \ldots, b_N \) (chosen independently and uniformly from among the \( cy_0 \)-smooths \( \leq x \)) have the property that \( b_1, \ldots, b_r \) correspond to the matrix \( M \) (with components arranged so that the \( i \)th component corresponds to prime \( q_i \)) and that the remaining numbers \( b_{r+1}, \ldots, b_N \) are coprime to \( q_1, \ldots, q_{\ell} \), is

\[
\sim y_0^{-N \sum_{i=1}^r 1/q_i} |\text{ Orbit}_{\text{Rows}}(M)| e^{-r} \prod_{1 \leq j \leq \ell} \frac{1}{q_j^{m_j}}.
\]

We next need to sum this quantity over all sequences of primes \( q_1, \ldots, q_{\ell} \in [y_0, cy_0] \). When we do this, we get

\[
\sim |\text{ Orbit}_{\text{Rows}}(M)| e^{-r} \prod_{j=1}^{\ell} \sum_{q_j \text{ prime}} y_0^{-N/q_j} \frac{1}{q_j^{m_j}}
\]

\[
\sim \frac{|\text{ Orbit}_{\text{Rows}}(M)| e^{-r}}{(\log y_0)^{-N(M)+\ell}} \prod_{j=1}^{\ell} \int_{y_0}^{c y_0} y_0^{-N/t_j} \frac{1}{t_j^{m_j}} dt_j
\]

\[
\sim \frac{|\text{ Orbit}_{\text{Rows}}(M)| e^{-r}}{(\log y_0)^{-N(M)+\ell}} \prod_{j=1}^{\ell} \int_{1/y_0}^{1/c y_0} y_0^{-N/u_j m_j-2} du_j
\]

Each of these integrals can be evaluated using integration-by-parts, to produce:

\[
\int_{1/y_0}^{1/c y_0} y_0^{-N/u_j m_j-2} du_j = \frac{(m_j - 2)!}{N \log y_0} \sum_{i=0}^{m_j-2} \frac{y_0^{-N/c y_0} e^{-m_j+2+i} - y_0^{-N/y_0}}{(m_j - 2 - i)! y_0^{m_j-2-i} (N \log y_0)^i}.
\]
Summing over all sequences of primes $q_1, \ldots, q_\ell$ as we just did will not quite give us the expected number of times that $M$ appears among $b_1, \ldots, b_N$ – we still need to divide by the possible permutations of a given sequence of primes $q_1, \ldots, q_\ell$ that leave the graph fixed, which is $|\text{AutColumns}(M)|$; and, we need to multiply by the number of places among the $N$ numbers $b_1, \ldots, b_N$ that $M$ can occur, which is $\binom{N}{r}$. When we do this, we get that the expected number of times $M$ as a connected component, conditioned on all the $b_i$ being $c y_0$-smooth, is

$$\sim \left( \frac{N^r}{r!} \right) \frac{|\text{OrbitRows}(M)| e^{-\ell} N^{-\ell}}{|\text{AutColumns}(M)| (\log y_0)^{-N(M)+2\ell}} \times \prod_{j=1}^\ell \left( (m_j - 2)! \sum_{i=0}^{m_j-2} \frac{y_0^{-N/cy_0} c^{-m_j+2+i} - y_0^{-N/y_0}}{(m_j - 2 - i)! y_0^{m_j-2-i} (N \log y_0)^i} \right).$$

Now, since $|\text{Aut}(M)| = |\text{OrbitRows}(M)||\text{AutColumns}(M)|$, since $|\text{OrbitRows}(M)||\text{OrbitRows}(M)| = r!$, we deduce that the above is

$$\sim \frac{N^{r-\ell} e^{-\ell}}{|\text{Aut}(M)| (\log y_0)^{-N(M)+2\ell}} \times \prod_{j=1}^\ell \left( (m_j - 2)! \sum_{i=0}^{m_j-2} \frac{y_0^{-N/cy_0} c^{-m_j+2+i} - y_0^{-N/y_0}}{(m_j - 2 - i)! y_0^{m_j-2-i} (N \log y_0)^i} \right).$$

This is only the conditional expectation for the number of times $M$ appears, given that $b_1, \ldots, b_N \leq x$ are $c y_0$-smooth. If we let $N$ be the number of $c y_0$-smooths that we select among the numbers $a_1, \ldots, a_J$, then to finish the proof of our lemma, we still need to sum over all possibilities for $N \leq J$, and multiply by the appropriate probabilities. Fortunately, we have that with probability $1 - o(1)$,

$$N \sim (\psi(x,c y_0)/x) J.$$  

For a more precise estimate we can use a binomial expansion in combination with Stirling’s formula and ugly estimates; or, we can use a cleaner, more conceptual tool, such as Chernoff’s inequality:

**Chernoff’s Inequality.** Suppose that $X_1, \ldots, X_n$ are independent random variables such that $\mathbb{E}(X_i) = 0$ and $|X_i| \leq 1$ for all $i$. Let $X = \sum_{i=1}^n X_i$ and let $\sigma^2$ be the variance of $X$. Then,

$$\text{Prob}(|X| \geq k \sigma) \leq 2 e^{-k^2/4}.$$
In our case we let $\alpha = \psi(x, cy_0)/x$ be the probability of picking a $cy_0$-smooth integer, and then let define $X_1, \ldots, X_J$ according to the rule

$$X_i = \begin{cases} 1 - \alpha, & \text{if } a_i \text{ is } cy_0 \text{-smooth;} \\ -\alpha, & \text{otherwise.} \end{cases}$$

Then, $E(X_i) = 0$ and $|X_i| < 1$. Since the $X_i$ are independent,

$$\sigma^2 = J \text{Var}(X_i) = J(\psi(x, cy_0)/x)(1 - \psi(x, cy_0)/x) \sim J\psi(x, cy_0)/x.$$

So, by Chernoff’s inequality, we deduce

$$\text{Prob}(|N - E(N)| > k \sqrt{J\psi(x, cy_0)/x}) = \text{Prob}(|X| > k \sqrt{J\psi(x, cy_0)/x}) = e^{-O(k^2)}.$$

This decays so rapidly with $k$ that we can just plug in

$$N = \frac{J\psi(x, cy_0)}{x} = \frac{J\psi(x, y_0)}{x} \cdot \frac{\psi(x, cy_0)}{\psi(x, y_0)} \sim c\eta \pi(y_0).$$

in (31) to get our expectation. When we do this, we get

$$\sim \frac{\eta^{r-\ell}}{\ell \pi(y_0)^{\ell+r-N(M)}} \prod_{j=1}^{\ell} \frac{(m_j - 2)! \sum_{i=0}^{m_j-2} e^{-n c_m j^2} - e^{-n c_i}}{(m_j - i - 2)! \eta^i}.$$

Using the fact that for $M \in \mathcal{M}$ we get $r + \ell = N(M) = 1$, our lemma follows.

### 4.9 Counting configurations II

Recalling that $A_{y_0}$ is the matrix $M$ formed from our sequence $a_1, \ldots, a_J$ upon deleting columns corresponding to primes $p \leq \pi(y_0)$, we wish to determine the expected value of $r(A_{y_0}) - \text{rank}(A_{y_0})$. We partition $A_{y_0}$ into connected components $M_1, \ldots, M_k$; and then

$$r(A_{y_0}) - \text{rank}(A_{y_0}) = \sum_j r(M_j) - \text{rank}(M_j),$$

which equals

$$\sum_j \sum_{S \subseteq R(M_j)} (-1)^{N(S)} = \sum_{S \subseteq R(A_{y_0})} (-1)^{N(S)}.$$

By combining (24), (33), and the identity

$$\sum_{\sigma \in S_\ell} \prod_{j=1}^{\ell} \frac{1}{c_{\sigma(i)}} = \prod_{j=1}^{\ell} \frac{1}{c_{\ell}}.$$
Here $S_\ell$ is the symmetric group on $1, \ldots, \ell$, and we will use $c_i = m_i - 1$ we have that if we sum over all orderings of the primes, we obtain
\[
E(r(A_{y_0}) - \text{rank}(A_{y_0})) \sim (\eta + f(\eta)) \pi(y_0)
\]
where
\[
f(\eta) := \sum_{M \in \mathcal{M}^*} \frac{(-1)^{N(M)}}{|\text{Aut}_{\text{Cols}}(M)| \cdot |\text{Aut}_{\text{Rows}}(M)|} \cdot \frac{\eta^{r(M)}}{\prod_{j=1}^{\ell} (m_j - 1)}.
\]
Here $\text{Aut}_{\text{Cols}}(M)$ denotes the number of ways to obtain exactly the same matrix $M$ when permuting the columns; and $\mathcal{M}^* = \mathcal{M}/\sim$ where two matrices are “equivalent” if one can be obtained from the other by permuting rows and columns.

### 4.10 Husimi graphs

We need to look more closely at the graphs that arise here. Notice that they are simple graphs. Moreover they have only monochromatic cycles, and so these cycles are subsets of a complete graph (between the vertices divisible by the prime corresponding to the color of the cycle). Thus we see that any two-connected subgraph is actually a complete graph. This is precisely the definition of a Husimi graph (see [5]). These graphs have a rich history, inspiring the combinatorial theory of species, and are central to the thermodynamical study of imperfect gases (see [5] for references and discussion). Moreover any Husimi graph corresponds to a matrix $M \in \mathcal{M}$; except the graph which is a single point, without edges. However if we add this into our set $\mathcal{M}^*$ (to create, say, $\mathcal{M}^+$) this would give a new term $\frac{(-1)^0}{1!} \cdot \eta^1 = \eta$, so that $f^+(\eta) = \eta + f(\eta)$; how convenient!

A key observation is that for a Husimi graph $G$ we have
\[
\text{Aut}(G) \cong \text{Aut}_{\text{Rows}}(M) \times \text{Aut}_{\text{Cols}}(M).
\]
To see this note that if $\sigma \in \text{Aut}(G)$ then it must define a permutation of the colors of $G$; that is an element $\tau \in \text{Aut}_{\text{Cols}}(M)$. Then $\tau^{-1}\sigma \in \text{Aut}(G)$ is an automorphism of $G$ that leaves the colors alone; and therefore must permute the elements of each given color. However if two vertices of the same color in $G$ are each adjacent to an edge of another color then permuting them would permute those colors which is impossible. Therefore $\tau^{-1}\sigma$ only permutes the vertices of a given color which are not adjacent to edges of any other color; and these correspond to automorphisms of the rows of $M$ containing just one 1. However this is all of $\text{Aut}_{\text{Rows}}(M)$ since if two rows of $M$ are identical then they must contain a single element, else $G$ would contain a non-monochromatic cycle.

Let $\text{Hu}(j_2, j_3, \ldots)$ denote the number of Husimi graphs with $j_i$ blocks of size $i$ for each $i$, on
\[
r = 1 + \sum_{i \geq 2} (i - 1)j_i
\]
vertices (this corresponds to there being $j_i$ primes which divide exactly $i$ of the integers in our set with $\ell = \sum j_i$ and $N(M) = \sum i j_i$). The Mayer-Husimi formula (which is (42) in [5]) gives
\[
\text{Hu}(j_2, j_3, \ldots) = \frac{(r - 1)!}{\prod_{i \geq 2} ((i - 1)j_i)!} \cdot r^{\ell - 1}.
\]
These are counted in the right way, in that we count all distinct labellings. Therefore, using this formula in (35), taking into account (36), we obtain, for $|\eta| < \eta_1$,

$$f(\eta) = \sum_{j_1, j_2, \ldots} \frac{(-1)^{r+\ell-1}}{\prod_{i \geq 2} (i-1)!^{j_i}(i-1)!^{j_i}} \cdot \eta^r.$$  
(39)

where $r$ is as in (37) and $\ell = \sum j_i$. We wish to determine $\eta_0$ such that as $\eta \to \eta_0^-$, $f(\eta)$ tends to 1.

To uncover the properties of $f(\eta)$ we will need, we begin with the simple identity

$$\exp\left(\sum_{i=1}^{\infty} c_i \right) = \sum_{k_1, k_2, \ldots \geq 0} \prod_{i \geq 1} \frac{c_i^{k_i}}{k_i!}. \tag{40}$$

Note that this series converges absolutely for any sequence of numbers $c_1, c_2, \ldots$ where $c_1 + c_2 + \cdots$ converges; and so, the terms in the series on the right-hand-side can be added in any order we please, and we get the same answer.

If we let

$$A(T) := -\sum_{j \geq 1} \frac{(-1)^j T^j}{j \cdot j!},$$

then our identity (40) gives us that

$$\exp(rA(\eta t)) = \sum_{j_1, j_2, \ldots \geq 0} \prod_{i \geq 2} \frac{1}{j_i!} \left(\frac{(-1)^{r+\ell-1} \eta^r}{(i-1)!} \right)^{j_i}.$$  

Thus, the coefficient of $t^{r-1}$ in $\exp(rA(\eta t))$ is

$$\sum_{j_1, j_2, \ldots \geq 0} \prod_{i \geq 2} \frac{1}{j_i!} \left(\frac{(-1)^{r+\ell-1} \eta^r}{(i-1)!} \right)^{j_i}.$$  

(Here, $\ell = j_2 + j_3 + \cdots$) It follows that

$$f'(\eta) = \sum_{r \geq 1} \frac{\text{coeff of } t^{r-1} \text{ in } \exp(rA(\eta t))}{r}. \tag{41}$$

### 4.11 Lagrange Inversion and a Formula for $f'(\eta)$

We begin by stating the Lagrange Inversion formula:
Theorem 3 Suppose that \( z = g(w) \) is analytic at \( w = 0 \), and that \( g'(0) \neq 0 \). Let \( a = g(0) \). Then, there exists a local inverse \( w = h(z) \) to \( g(w) \), defined in the neighborhood of \( a \), whose Taylor expansion about \( z = a \) is given by

\[
h(z) = \sum_{r=1}^{\infty} \left( \frac{d}{dw} \right)^{r-1} \left( \frac{w}{g(w) - a} \right)^r \bigg|_{w=0} \frac{(z-a)^r}{r!}.
\]

A corollary of this theorem, and Taylor’s theorem, is

Corollary 1 In Lagrange’s theorem if \( g(w) = w/\varphi(w) \), where \( \varphi(w) \neq 0 \), which in particular implies \( g(0) = 0 \), then \( h(z) \) takes the form

\[
h(z) = \sum_{r=1}^{\infty} c_{r-1} z^r,
\]

where \( c_j \) is the coefficient of \( z^j \) in the Taylor series expansion of \( \varphi(z) \).

Applying this corollary with \( \varphi(w) = e^{A(\eta w)} \), we find that \( g(w) = we^{-A(\eta w)} \) and has inverse \( h(z) \), where

\[
h(1) = \sum_{r \geq 1} \text{coeff. of } z^{r-1} \text{ in } \exp(rA(\eta z)) = f'(\eta).
\]

Thus, since

\[
1 = g(h(1)) = h(1)e^{-A(\eta h(1))} = f'(\eta)e^{-A(\eta f'(\eta))},
\]

we deduce

\[
f'(\eta) = e^{A(\eta f'(\eta))}.
\]  

(42)

Of course note that this only holds for \( \eta \) in some neighborhood about 0, and in the next subsection we will determine this neighborhood.

4.12 The Radius of Convergence of \( f'(\eta) \)

By (42) we have that if \( \eta \) is a non-negative real number such that \( f'(\eta) \) converges, then \( f'(\eta) > 0 \). It follows that if there exists \( \eta \) for which \( f(\eta) = 1 \), then it is unique.

In fact, we know more than just that \( f'(\eta) > 0 \): First, observe that since

\[
A'(T) = \frac{1 - e^{-T}}{T},
\]

we have that \( A'(T) > 0 \) for all \( T > 0 \); and, as \( A(0) = 0 \), we further deduce that \( A(T) > 0 \) for all \( T > 0 \). Thus, since \( A(\eta f'(\eta)) > 0 \), we deduce

\[
f'(\eta) > 1.
\]
Define
\[ B_\eta(y) := A(\eta y) - \log y. \]
Find \( \eta \) such that \( f'(\eta) = y \) is the same as finding \( \eta \) satisfying \( B_\eta(y) = 0 \). Now \( B_\eta(1) = A(\eta) > 0 \), and \( B'_\eta(y) = -e^{-\eta y}/y; \) and so, for \( y \geq 1 \),
\begin{equation}
B_\eta(y) = B_\eta(1) + \int_1^y B'_\eta(t) dt = A(\eta) - \int_1^y \frac{e^{-\eta t}}{t} dt. \tag{43}
\end{equation}
In particular,
\[ B_\eta(\infty) = A(\eta) - \int_1^\infty \frac{e^{-\eta y}}{y} dy. \]
If we want \( \eta \) to satisfy \( B_\eta(\infty) = 0 \), then we want that
\[ A(\eta) = \int_1^\infty \frac{e^{-\eta y}}{y} dy = \int_\eta^\infty \frac{e^{-v}}{v} dv. \tag{44} \]
The function \( A(T) \) also satisfies the relation
\[ A(T) = A(0) + \int_0^T A'(v) dv = \int_0^T \frac{(1 - e^{-v})}{v} dv; \tag{45} \]
and so, combining this with (44) we seek \( \eta \) satisfying
\[ \int_0^\eta \frac{(1 - e^{-v})}{v} dv = \int_\eta^\infty \frac{e^{-v}}{v} dv. \tag{46} \]
We can rewrite this as
\[ \int_1^\eta \frac{dv}{v} = \int_1^\infty \frac{e^{-v}}{v} dv - \int_0^1 \frac{(1 - e^{-v})}{v} dv. \tag{47} \]
As the left-hand-side integrates to \( \log(\eta) \), on exponentiating we find that
\[ \eta = \exp \left( \int_1^\infty \frac{e^{-v} dv}{v} - \int_0^1 \frac{(1 - e^{-v})}{v} dv \right) = e^{-\gamma} = .561459\ldots \tag{48} \]
This establishes that \( \eta_1 \leq e^{-\gamma} \). To deduce that \( \eta_1 = e^{-\gamma} \), we must show that the series (41) converges absolutely for \( |\eta| < e^{-\gamma} \). To this end, let \( 1 \leq y < \infty \) be given.

To deduce that the series (41) converges absolutely for \( |\eta| < e^{-\gamma} \), we first note that the equation \( y = e^{A(\eta y)} \) implicitly defines an analytic function \( y \) in a neighborhood of \( \eta = 0 \); that function equals the series (41) in the largest open disk where the series converges absolutely. This disk \( |\eta| < e^{-\gamma} \) is also the largest disk where \( y \) is analytic, because from (44) through (48) we know that \( y \) has a singularity at \( \eta = e^{-\gamma} \). By analytic continuation, we will have that \( y \) equals the series (41) in the largest circle about \( \eta = 0 \) where the function \( y \) has no singularities. From what we worked out above, the series (41) converges for all positive real numbers less than \( e^{-\gamma} \); and so, it converges absolutely for all complex numbers of size less than \( e^{-\gamma} \).
4.13 Showing that $f(\eta)$ approaches 1 as $\eta \to \eta_0^-$

For $x > 0$ define

$$g(x) = e^{A(x)}.$$  

Let $\tau = \eta f'(\eta)$ so that $\eta = \tau/g(\tau)$ by (42), and $f'(\eta) = g(\tau)$. Now, since

$$\frac{g'(x)}{g(x)} = A'(x) = \frac{1 - e^{-x}}{x},$$

we have

$$f(\eta) = \int_0^\eta f'(u)du = \int_0^{\tau=\tau} g(t)d(t/g) = \int_0^\tau \left(1 - t\frac{g'}{g}\right)dt = \int_0^\tau e^{-t}dt = 1 - e^{-\tau} = 1 - e^{-\eta f'(\eta)}.$$  

(49)

Thus $f(\eta) < 1$ when $f'(\eta)$ converges, i.e., when $0 \leq \eta < \eta_0$, and $f(\eta) \to 1$ and $\eta \to \eta_0^-$.

4.14 From expectation to probability

To complete the proof of our upper bound, we must show that

$$\text{Prob}(T < (e^{-\gamma} + o(1))J_0(x)) = 1 + o(1),$$

and to do this it suffices to prove that for a fixed matrix $M$ satisfying $\Delta(M) = 0$ as above, the number of times that $M$ appears as a submatrix of $A_{y_0}$ is, with probability $1 - o(1)$, very close to the expected value. For then, by (32), (33) and (34), we will have that

$$J - \text{rank}(A_{y_0}) > \pi(y_0)$$

holds with probability $1 - o(1)$ if $J = (e^{-\gamma} + o(1))J_0(x)$.

To show that each $M$ appears the expected number of times with high probability, we will use Chebychev’s inequality. We will need to define some random variables: Suppose we have chosen $a_1, \ldots, a_J \leq x$ randomly, then given a subset $S$ of $\{1, \ldots, J\}$, let $X_M(S)$ equal 1 if when we represent the sequence $\{a_s\}_{s \in S}$ as vectors over $F_2$, and remove the columns corresponding to primes less than or equal to $y_0$, some permutation of the rows that result correspond to the matrix $M$; and, let $X_M(S)$ equal 0 otherwise. Then, the total number of times that $M$ occurs in $A_{y_0}$ is

$$Y_M = \sum_{S \subseteq \{1, \ldots, J\}} X_M(S).$$

From what we worked out in (38) and (24) we have that

$$\mathbb{E}(Y_M) \sim c_M \cdot \pi(y_0)
$$
where $c_M$ is a constant that depends only on $M$. Since each of the $X_M(S)$ have the same probability distributions, we have that for every $S \subset \{1, \ldots, J\}$,

$$\mathbb{E}(X_M(S)) \sim \left( \frac{J}{|M|} \right)^{-1} c_M \pi(y_0).$$

We also have that

$$\mathbb{E}(Y_M^2) = \sum_{S_1, S_2 \subset \{1, \ldots, J\} \atop |S_1|=|S_2|=|M|} \mathbb{E}(X_M(S_1)X_M(S_2)).$$

Now, if $S_1$ and $S_2$ are disjoint, then $X_M(S_1)$ and $X_M(S_2)$ are independent, and we deduce

$$\mathbb{E}(X_M(S_1)X_M(S_2)) = \mathbb{E}(X_M(S_1))\mathbb{E}(X_M(S_2)).$$

Let us now consider the case where $S_1$ and $S_2$ are not disjoint. If both $X_M(S_1)$ and $X_M(S_2)$ equal 1, then we must have that if we let $S = S_1 \cup S_2$, then the sequence $\{a_s\}_{s \in S}$ corresponds to some matrix $M'$ having $|S|$ vertices. This matrix $M'$ may not satisfy $\Delta(M) = 0$, but at least we always will have that for $T \subset \{1, \ldots, J\}$, $|T| = |S|$,

$$\mathbb{E}(X_{M'}(T)) \lesssim \left( \frac{J}{|T|} \right)^{-1} c_{M'} \pi(y_0),$$

because we know that if $\Delta(M') \neq 0$, then $\mathbb{E}(Y_{M'}) = o(\pi(y_0))$. Also note that there are at most a finite number of possible matrices $M'$ that can occur, given that $S_1$ and $S_2$ are to correspond to the matrix $M$.

Thus, if $|S_1 \cap S_2| = j$, then

$$\mathbb{E}(X_M(S_1)X_M(S_2)) \ll_M \left( \frac{J}{2|M| - j} \right)^{-1} \pi(y_0).$$

It follows that

$$\mathbb{E}(Y_M^2) - \mathbb{E}(Y_M)^2 \ll_M \pi(y_0) \sum_{j=1}^{|M|} \left( \frac{J}{2|M| - j} \right)^{-1} \sum_{S_1, S_2 \subset \{1, \ldots, J\} \atop |S_1|=|S_2|=|M|} 1$$

$$= \pi(y_0) \sum_{j=1}^{|M|} \left( \frac{J}{2|M| - j} \right)^{-1} \left( \frac{J}{|M| - j, |M| - j, j, J - 2|M| + j} \right)$$

$$\ll_M \pi(y_0).$$

Chebychev’s inequality now implies that

$$\text{Prob}(|Y_M - \mathbb{E}(Y_M)| > \epsilon \mathbb{E}(Y_M)) \ll_M \frac{\mathbb{E}(Y_M^2) - \mathbb{E}(Y_M)^2}{\epsilon^2 \mathbb{E}(Y_M)^2} \ll_M \frac{1}{\epsilon^2 \pi(y_0)}.$$

So, with probability $1 - o(1)$,

$$Y_M \sim \mathbb{E}(Y_M),$$

and we are done.
5 Proof of the lower bound on $T$ in Theorem 1

5.1 Proof strategy

To show that

$$\text{Prob}(T > (\pi/4)(e^{-\gamma} - o(1))J_0(x)) = 1 - o(1),$$

we show, for a certain value of $y$, that the expected number of non-trivial subsets $S$ of $\{1, \ldots, J\}$ for which $\prod_{i \in S} a_i$ is a square is $o(1)$, for $J(x) = (\pi/4)(e^{-\gamma} - o(1))J_0(x)$.

5.2 Structure of a square product

We begin with the following proposition.

**Proposition 2** Select integers $a_1, \ldots, a_J$ at random from $[1, x]$. The probability that there exists a subsequence $I$ of the $a_i$ with

$$2 \leq |I| \leq \frac{\log x}{2\log \log x},$$

such that $\prod_{a \in I} a \in \mathbb{Z}^2$

is $O(J^2 \log x/x)$ provided $J < x^{o(1)}$.

**Proof.** Suppose that $b_1, \ldots, b_k$ were chosen at random from $[1, x]$. The probability that $b_1 b_2 \ldots b_k$ is a square equals

$$x^{-k}\{|b_1, \ldots, b_k \leq x : b_1 b_2 \ldots b_k \in \mathbb{Z}^2\}|.$$

Now write each $b_i$ uniquely as

$$b_i = c_i u_i^2,$$

where $c_i$ is squarefree.

Assuming that $b_1 \ldots b_k$ is a square, which implies $c_1 \ldots c_k$ is a square, define the doubly indexed sequence $c_{i,j}$, where $i, j = 1, \ldots, k$ and $i \neq j$, to be any satisfying the relations

$$c_{i,j} = c_{j,i}; \text{ and, } c_i = \prod_{j \neq i} c_{i,j}. \quad (50)$$

The fact that such $c_{i,j}$ exist can be seen as follows: For each prime $p$ dividing $c_1 \cdots c_k$, we will need to decide which $c_{i,j}$ that $p$ divides; and, to do this, suppose that $p$ divides $c_{i_1, \ldots, c_{i_{2t}}}$ (the reason it is $2t$ is that all the $c_i$ are square-free and have product a square). Then, the following $c_{i,j}$ are to be divisible by $p$, and no others:

$$c_{i_1, i_2}, c_{i_2, i_1}, c_{i_3, i_4}, c_{i_4, i_3}, \ldots, c_{i_{2t-1}, i_{2t}}, c_{i_{2t}, i_{2t-1}}.$$ 

If this process leaves some $c_{i,j}$ not divisible by any prime $p|c_1 \cdots c_k$, then we set $c_{i,j} = 1$.  

30
Given $c_1,\ldots,c_k$, the number of sequences $b_1,\ldots,b_k$ satisfying $b_i = c_iu_i^2$ is the number of possibilities for the numbers $u_i$, which is $\leq (x/c_i)^{1/2}$; and so, the probability that $b_1\cdots b_k$ is a square is

$$
\leq \frac{1}{x^k} \sum_{c_{i,j} \leq x} \prod_{i=1}^{k} \left( \frac{x}{\prod_{j \neq i} c_{i,j}} \right)^{1/2}
$$

$$
\leq \frac{1}{x^{k/2}} \sum_{1 \leq i < j \leq k} \left( \sum_{c_{i,j} \leq x} \frac{1}{c_{i,j}} \right) \leq \frac{1}{x^{k/2}} (1 + \log x)^{k(k-1)/2}
$$

(51)

since each $c_{i,j}$ appears twice in the above product. Therefore the probability that there exists $I \subset \{1,2,\ldots,J\}$ for which $\prod_{i \in I} a_i \in \mathbb{Z}^2$ with $|I| = k$ is

$$
\leq \left( \frac{J}{k} \right) \frac{1}{x^{k/2}} (1 + \log x)^{k(k-1)/2} \leq \left( \frac{J^2(1 + \log x)^{k-1}}{x} \right)^{k/2}
$$

which gives $O(J^2 \log x/x)$ for $k = 2$, and is $\leq 1/x$ for $3 \leq k \leq \log x/2 \log \log x$. □

5.3 The main argument

In this subsection, we prove that

$$
\text{Prob}(T > (\pi/4)(e^{-\gamma} - o(1))J_0(x)) = 1 - o(1).
$$

We first observe that, as a consequence of the upper bound developed in Section 4, we may assume that $T < J_0(x)$ holds with probability $1 - o(1)$ (in fact, we know much more is true). Furthermore, from Proposition 2 we know that we need only focus on subsequences $I$ of $a_1,\ldots,a_J$ (where $J = T < J_0(x)$) of length exceeding log $x/2 \log \log x$, that have product equal to a square.

Throughout we shall write $a_i = b_id_i$ where $P(b_i) \leq y$ and where either $d_i = 1$ or $p(d_i) > y$ (here, $p(n)$ denotes the smallest prime divisor of $n$), for $1 \leq i \leq k$. If $a_1,\ldots,a_k$ are chosen at random from $[1,x]$ then

$$
\text{Prob}(a_1\ldots a_k \in \mathbb{Z}^2) \leq \text{Prob}(d_1\ldots d_k \in \mathbb{Z}^2)
$$

$$
= \sum_{d_1,\ldots,d_k \geq 1} \prod_{i=1}^{k} \frac{\Psi(x/d_i,y)}{x} \quad \text{for} \quad d_i = 1 \text{ or } p(d_i) > y
$$

$$
\leq \left( 1 + o(1) \right) \frac{\Psi(x,y)}{x} \sum_{n=1}^{T} \frac{\tau_k(n^2)}{n^{2\alpha}}
$$

(52)

by Proposition 1.
Out of \( J = \eta J_0 \) integers, the number of \( k \)-tuples is \( \binom{J}{k} \ll (eJ/k)^k \); and so the expected number of \( k \)-tuples whose product is a square is
\[
\ll \left( e + o(1) \right) \frac{\eta y}{k \log y_0} \frac{\Psi(x, y)/y}{\Psi(x, y_0)/y_0} \prod_{p \leq y} \left( 1 + \frac{\tau_k(p^2)}{p^{2\alpha}} + \frac{\tau_k(p^4)}{p^{4\alpha}} + \cdots \right). \tag{53}
\]

We now consider \( k \) in two different ranges, and in both ranges we will select different values for \( y \), so as to give good upper bounds for (53):

- First, if
  \[
  \frac{\log x}{2 \log \log x} < k \leq y_0^{1/4},
  \]
  then let \( y = y_0^{1/3} \) so that \( k = o(y_0^\alpha) \). Therefore the Euler product in (53) is
  \[
  \leq \exp \left( O \left( \sum_{p \leq y} \frac{k^2}{p^{2\alpha}} \right) \right) \leq \exp \left( O \left( \frac{k^2 y^{2(1-\alpha)}}{y \log y} \right) \right) = e^{o(k)}.
  \]

  Now \( \Psi(x, y_0^\gamma) = x/y_0^{1/\gamma + o(1)} \) and therefore the quantity in (53) is
  \[
  \leq \left( \frac{1/y_0^{3+o(1)}}{k/y_0^{2+o(1)}} \right)^k \leq y_0^{-k + o(k)},
  \]
  which is \( < 1/x^2 \) in this first range for \( k \).

- Next, we consider the range
  \[
  y_0^{1/4} \leq k = y_0^\beta \leq J \leq J_0.
  \]
  In this case we will choose \( y \) so that \( \lfloor k/C \rfloor = \pi(y) \), and then will optimize the \( C \) later. For this choice of \( y \) a simple calculation reveals that
  \[
  1 + \frac{\tau_k(p^2)}{p^{2\alpha}} + \frac{\tau_k(p^4)}{p^{4\alpha}} + \cdots \sim 1 + \frac{(k/p^\alpha)^2}{2!} + \frac{(k/p^\alpha)^4}{4!} + \cdots = \frac{e^{k/p^\alpha} + e^{-k/p^\alpha}}{2}.
  \]
  In order to evaluate (53) we need to product this over primes \( p > y \). The logarithm of this product equals
  \[
  \sum_{p \leq y, \text{prime}} \log(e^{k/p^\alpha}/2 + e^{-k/p^\alpha}/2) \sim \int_{y}^{\infty} \frac{1}{\log t} \log(e^{k/t^\alpha}/2 + e^{-k/t^\alpha}/2) dt.
  \]
  Letting \( z = k/t^\alpha \), from (13) this last integral is
  \[
  \sim \int_{0}^{C/\beta^2} \frac{\log(e^z/2 + e^{-z}/2)}{\log(k/z)} \frac{(k/z)^{1/\alpha}}{z} dz.
  \]
Now, $z^{1/\alpha} \sim z$ for $z \in [0, C]$, and from (13) we know
\[
k^{1/\alpha} = \frac{y^\beta}{\alpha} = \frac{y_0^{\beta/(1-(1-\alpha))}}{\alpha} \sim \frac{y^\beta y_0^{(1-\alpha)}}{\alpha} \sim k \beta^{-2} \log y.
\]
So, the integral is
\[
\sim g(\beta, C) := k \beta^{-2} \int_0^C \frac{\log(e^{2z} + e^{-2z})}{z^2} dz.
\]
It follows that the quantity in (53) is bounded from above by
\[
(1 + o(1)) \frac{e^{1 + g(\beta,C)/k \eta}}{C} \frac{\Psi(x, y)/y}{\Psi(x, y_0)/y_0}^k.
\]
(54)

Now, for any fixed $C$ we have as a consequence of Lemma 1 that (54) is $o_C(1/x^2)$ unless $\beta = 1 + o_C(1)$; and so, we really only need to consider $k = y_0^{1+o(1)}$, as the total expected number of $k$-tuples for other values of $k$ add only to $o(1)$. Thus, since $y_0$ maximizes $\Psi(x, y)/y$ for $y = y_0$, we deduce that (53) is at most
\[
(1 + o(1)) \frac{e^{1 + g(1,C)/k \eta}}{C} \frac{\Psi(x, y)/y}{\Psi(x, y_0)/y_0}^k \leq (1 + o(1)) \frac{e^{1 + g(1,C)/k \eta}}{C}^k.
\]
(55)

Now, to minimize this over $C > 0$, we must minimize $g(1, C)/k - \log(C)$. However
\[
\frac{d}{dC} \left( \frac{g(1, C)}{k} - \log(C) \right) = \frac{\log(e^{C/2} + e^{-C/2})}{C^2} - \frac{1}{C} < 0;
\]
and so we minimize by letting $C \to \infty$. One can easily verify that
\[
\lim_{C \to \infty} \left( \frac{g(1, C)}{k} - \log(C) \right) = \int_1^\infty \log(1 + e^{-2z}) \frac{dz}{z^2} - \log 2 + \int_0^1 \log(\frac{e^z + e^{-z}}{2}) \frac{dz}{z^2},
\]
which yields, using Maple, that
\[
\lim_{L \to \infty} \frac{\exp(g(1, L)/k)}{L} = 0.83425...
\]
By taking $L$ arbitrarily large, we have from (55) that (53) is at most $(2.26773 + o(1)) \eta^k$. This implies the second part of Theorem 1, because if $\eta = (\pi/4)(e^{-\gamma} - o(1))$, then this quantity is $o(1)$.

In fact, with a little more work, one can show that if $J < (\pi/4)(e^{-\gamma} - o(1))J_0(x)$, then the expected number of $k$-tuples among our $J$ that has product equal to a square, over all $k > \log x/2 \log \log x$, is only $o(1/\sqrt{x})$. □
6 Proof of Theorem 2

6.1 Proof of Theorem 2, parts (a) and (b)

From the results in the previous section we know that if \( J < (\pi/4)(e^{-\gamma} - o(1))J_0(x) \), and
\[
k > \frac{\log x}{2\log \log x},
\]
then the probability that some subset of \( k \) elements chosen from among our \( J \) has product equal to a square is \( o(1/\sqrt{x}) \). This, together with Proposition 2, extends the range to \( k \geq 2 \) (also with probability \( o(1/\sqrt{x}) \)). Thus, if \( T < (\pi/4)(e^{-\gamma} - o(1))J_0(x) \), with probability \( 1 - o(1) \) we have that all subsequences of our \( a_i \) that has product equal to a square, consist of a single element (i.e. \( k = 1 \)), which is therefore itself a square. This proves Part (a) of Theorem 2.

To prove Part (b) of Theorem 2, first suppose we are given that
\[
T < (e^{-\gamma} + \epsilon)J_0(x).
\]
Then, from what we have proved in the previous section, with probability \( 1 - o(\epsilon) \) we will have that
\[
(\pi/4)(e^{-\gamma} - o(1))J_0(x) \leq T < (e^{-\gamma} + \epsilon)J_0(x).
\]
Let
\[
J = T, \quad \text{and} \quad \eta = J/J_0.
\]
Suppose that \( k = y^{\beta + o(1)} \) satisfies
\[
k \geq y^{1/4}, \quad \text{and} \quad k \notin [y_0 \exp(-(c_3 + \epsilon)\sqrt{\log y_0}), y_0 \exp((c_3 + \epsilon)\sqrt{\log y_0})],
\]
where
\[
c_3 = \sqrt{2 - \log 2}.
\]
Given such a \( k \), let \( y \) be any integer such that \( \pi(y) = \lfloor k/C \rfloor \), where \( C > 0 \) is some large number to be chosen later. From the argument just before (54) from the previous section, we know that the expected number of \( k \) tuples that have product equal to a square, among our \( J \) numbers, is at most
\[
\left( (1 + o(1)) \frac{e^{1+g(\beta,C)/k} \Psi(x,y)/y}{\Psi(x,y_0)/y_0} \right)^k.
\]
We also know that the total contribution of all \( k \) tuples where \( \beta \) is not \( 1 + o(1) \), is \( o(1) \); so, we only need to focus on \( \beta = 1 + o(1) \) (i.e. \( k = y^{1+o(1)}_0 \)). As we proved, by taking \( C \) near infinity (but still fixed), this is at most
\[
\left( (2.26773 + o(1)) \frac{\eta \Psi(x,y)/y}{\Psi(x,y_0)/y_0} \right)^k.
\]
Now, applying Lemma 2, we find that for sufficiently small \( \epsilon > 0 \) this is at most
\[
((2.26773 + o(1))(2/e^2 + o(1))\eta)^k < (0.345)^k,
\]
using the fact that \( \eta < e^{-\gamma} + \epsilon \).

So, the expected number of \( k \) tuples with product a square is \( o(1) \) for all \( k \) satisfying (56); and, in fact, the same is true for \( k < y_0^{1/4} \), as can be see by applying the ideas from the previous section, along with Proposition 2. We deduce, then, that with probability \( 1 - o(1) \), any \( k \) tuple with product equal to a square must occur with
\[
k \in [y_0 \exp(-c_3 + \epsilon) \sqrt{\log y_0}, \ y_0 \exp((c_3 + \epsilon) \sqrt{\log y_0})].
\]
This proves part (b) of Theorem 2. \( \square \)

6.2 Proof of Theorem 2, part (c)

In the previous subsection we proved that
\[
|I| \leq y_1 := y_0 \exp((c_3 + \epsilon) \sqrt{\log y_0}),
\]
with probability \( 1 - o(1) \). In this section, among other results, we prove Part (c) of Theorem 2.

**Proposition 3** Write each \( a_i = b_id_i \) with \( y = y_1 \). With probability \( 1 - o(1) \), we have
\[
(i) \ d_1 \ldots d_k \text{ is cubefree, and}
(ii) \ l \ll \log y_0
(iii) \ d_1 \ldots d_l \text{ has either } l - 1 \text{ or } l \text{ distinct prime factors}
(iv) \text{ if } l \to \infty \text{ as } x \to \infty \text{ then } d_1 \ldots d_l \text{ has exactly } l - 1 \text{ distinct prime factors.}
\]

**Proof.** First note that with the choice of \( y = y_1 \), we have \( y/k \log y \to \infty \) and \( y = y_0^{1+o(1)} \), so we know that \( y^\alpha \sim y/\log y \) by (12).

Since no subproduct of \( d_1 \ldots d_l \) is a square, we must have that \( d_1 \ldots d_l \geq y^{2(l-1)} \).

For ease of notation we will relabel, replacing \( d_1 \ldots d_l \) by \( d_1 \ldots d_l \).

In the argument at the start of the previous section (with \( k \) replaced by \( l \)) we restrict our attention to cases in which \( d_1 \ldots d_l \geq y^{2l} \phi(x)^2 \), where \( \phi(x) = y^{O(1)} \). To obtain an upper bound we may, in (52), multiply through the summand by \( (n/y' \phi(x))^{2\theta} \) where we have chosen \( \theta > 0 \) so that \( y^{2\theta} = 2y/\log y \). Then we must multiply the right side of (53) through by \( 1/(y^{2\theta}) \phi(x)^{2\theta} \) and change the terms in the Euler product to \( (1 + \tau_1 p^2)/p^{2(\alpha-\theta)} + \tau_1 p^4)/p^{4(\alpha-\theta)} + \ldots \). By the prime number theorem the Euler product is \( e^{(1+o(1))l} \); and so the expected number of such \( l \)-tuples is
\[
\leq \frac{1}{\phi(x)^{2\theta}} \left( (1 + o(1)) \frac{e^2 \eta}{2} \frac{\Psi(x, y)/y}{\Psi(x, y_0)/y_0} \right)^l \leq \frac{1}{e^l \phi(x)^{2\theta}},
\]
(57)
as \( \eta < 1 + o(1) \), by Lemma 2 for \( y = y_1 \).

Taking \( \phi(x) = 1/y \) above, and as \( 2\theta < 1 \) by definition, we find that the expected number of such products with \( l \gg \log y_0 \) is \( o(1) \). This implies (ii).

Therefore we may assume henceforth that \( l \ll \log y_0 \), which implies that \( 2\theta \sim 1 \). Select any \( \phi(x) \to \infty \); then the expected number of such products with \( l \ll \log y_0 \) is \( o(1) \). This implies (ii).

For \( p \) divides more than two of the \( d_i \) then there must be one or two minimal products of \( d_i \)'s that are squares such that their union is divisible by \( p^3 \). Thus the union of these two products is a set of distinct \( d_i \)'s whose product \( N \) is a powerful number (that is if prime \( q | N \) then \( q^2 | N \) for which \( p^3 | N \). To get an upper bound on the expected number of such sets we proceed much as above but now the Euler product in (53) is replaced by

\[
\sum_{p>y} \left( \frac{\tau(p^3)}{p^{3\alpha}} + \frac{\tau(p^4)}{p^{4\alpha}} + \cdots \right) \prod_{q>y, q \neq p} \left( 1 + \frac{\tau(q^2)}{q^{2\alpha}} + \frac{\tau(q^3)}{q^{3\alpha}} + \frac{\tau(q^4)}{q^{4\alpha}} + \cdots \right);
\]

and we thus get the bound in (55) times a factor which is

\[
\ll \sum_{p>y} \frac{l^3}{p^{3\alpha}} \times \frac{l^3}{y^{3\alpha-1}\log y} \times \frac{l^3}{(y/\log y)^2} \ll \frac{(\log y)^5}{y^2} = o(1);
\]

which implies (i). \( \square \)

**Remark 2.** If we had selected \( y = y_0 \exp(C\sqrt{\log y_0}) \) for large \( C = C_\epsilon \), and let \( \phi(x) = 1 \), we then get the bound \( \ll \epsilon^l \) in (57). Therefore \( d_1 \ldots d_l < y^{2l} \) with probability \( 1 + O(\epsilon) \), and this implies that \( d_1 \ldots d_l \) has exactly \( l - 1 \) prime factors.

## 7 Acknowledgements

We thank Francois Bergeron for pointing out the connection with Husimi graphs, for providing mathematical insight and pointing us to the right references such as [5].

## References

[1] J. Buhler, H. W. Lenstra Jr., and C. Pomerance, *Factoring integers with the number field sieve*, Lecture Notes in Math, 1554, Springer, Berlin, 1993.

[2] J. D. Dixon, *Asymptotically fast factorization of integers*, Math. Comp. 36 (1981), 255-260.

36
[3] E. Friedgut, *Sharp Thresholds of Graph Properties, and the k-SAT Problem*, J. Amer. Math. Soc. **12** (1999), 1017-1054.

[4] A. Hildebrand and G. Tenenbaum, *On integers free of large prime factors*, Trans. Amer. Math. Soc **296** (1986), 265–290.

[5] Pierre Leroux, Enumerative Problems Inspired by Mayer’s Theory of Cluster Integrals, Electronic Journal of Combinatorics, paper R32, May 14, 2004.

[6] C. Pomerance, *The quadratic sieve factoring algorithm*. Advances in cryptology, Paris (1984), 169-182.

[7] C. Pomerance, *The role of smooth numbers in number theoretic algorithms*, Proc. International Cong. of Mathematicians (Zurich, 1994), Birkhäuser **1** (1995), 411 - 422.

[8] C. Pomerance, *Multiplicative independence for random integers*, Analytic number theory: Proceedings of a conference in honor of Heini Halberstam (eds., B.C. Berndt et. al.), Birkhäuser **2** (1996), 703 - 711.

[9] C. Pomerance, *Smooth numbers and the quadratic sieve*, Proc. of an MSRI workshop, J. Buhler and P. Stevenhagen, eds. (to appear).

[10] R. Silverman, *The multiple polynomial quadratic sieve*, Math. Comp. **bf 48** (1987), 329-339.

[11] G. Tenenbaum, Introduction to the analytic and probabilistic theory of numbers, Cambridge Univ. Press 1995.