7. J. P. Serre, Divisibilité de certaines fonctions arithmétiques, L'Ens. Math., 22 (1976) 227-260.

8. J.-P. Serre, Quelques applications du théorème de densité de Chebotarev, Publ. Math. IHES, 54 (1982) 123-201

9. G. Shimura, Introduction to the arithmetic theory of automorphic forms, Publ. Math. Soc. Japan 11, Princeton University Press, Princeton 1971.

10 I. M. Vinogradov, The method of trigonometrical sums in the Theory of numbers, London, Interscience, 1947.

V. Kumar Murty
School of Mathematics
Tata Institute of Fundamental Research
Homi Bhabha Road
Bombay 400 005

ON THE DISTRIBUTION OF ROUND NUMBERS

Carl Pomerance*
Department of Mathematics
University of Georgia
Athens, Georgia 30602/USA

Dedicated to Paul Erdös on his seventieth birthday.

§1. Introduction.

"Round" numbers are defined in Hardy and Wright [4], p. 358 as integers with an abnormally large number of prime factors. Of course, this is not a precise mathematical definition. One reason is that there are two natural ways of counting the number of prime factors of an integer: with multiplicity and without. We write

$$\Omega(n) = \sum_{p^a | n} 1 , \quad \omega(n) = \sum_{p | n} 1$$

where p denotes a prime and a denotes a positive integer.

The normal behavior of $\Omega(n)$ and $\omega(n)$ are the same. A famous result of Hardy and Ramanujan [3] is that for each $\varepsilon > 0$ the set of n for which

$$|\Omega(n) - \log\log n| > (\log\log n)^{1/2+\varepsilon}$$

or

$$|\omega(n) - \log\log n| > (\log\log n)^{1/2+\varepsilon}$$

has asymptotic density 0. However, the maximal orders are different. Trivially we have

$$\Omega(n) \leq \frac{\log n}{\log 2}$$

and from the prime number theorem we have

*Supported in part by a National Science Foundation grant.

$$\omega(n) \leq \pi(\log n) + O\left(\frac{\log n}{\exp(\sqrt{\log\log n})}\right)$$

$$= \frac{\log n}{\log\log n} + (1+o(1))\frac{\log n}{(\log\log n)^2}$$

(of course, the expansion can be taken further). These upper bounds are the true maximal orders for $\Omega(n)$ and $\omega(n)$.

To study, then, the distribution of round numbers, one should look at the functions

$$N(x,y) = \sum_{n\leq x,\,\Omega(n)=y} 1 ,$$

$$\pi(x,y) = \sum_{n\leq x,\,\omega(n)=y} 1 ,$$

where y denotes a natural number that is considerably bigger than $\log\log x$.

There is much known about these functions when y is of the order of magnitude of $\log\log x$ or smaller. The situation for $N(x,y)$ for all y is now very good due to a recent result of Nicolas who obtains an asymptotic formula for $N(x,y)$ for each integer $y \geq (2+\varepsilon)\log\log x$. We will describe these results below.

The situation for $\pi(x,y)$ for large y is much poorer. This is partly due to the fact that numbers round in the $\Omega(n)$ sense are usually divisible by a large power of 2, while numbers round in the $\omega(n)$ sense are divisible by many different primes. Thus the study of $N(x,y)$ for large y almost reduces to a study of the multiples of large powers of 2, while the situation with $\pi(x,y)$ is not so simple. It is my goal in this paper to establish new upper and lower bound inequalities for $\pi(x,y)$ for large y that begin to pin down this function.

Before stating the principal results, let us first review what is known about $\pi(x,y)$ and $N(x,y)$. For y a fixed positive integer, Landau [6] proved in 1900 that

$$\pi(x,y) \sim N(x,y) \sim \frac{x}{\log x}\cdot\frac{(\log\log x)^{y-1}}{(y-1)!} \quad \text{as} \quad x\to\infty .$$

The same result was obtained by Erdös [1] for those y with

$$y - \log\log x = O(\sqrt{\log\log x}) .$$

Asymptotic formulae for a wider range were obtained by Sathe [10] and Selberg [11]. Let

$$(1.1) \qquad F(z) = \frac{1}{\Gamma(1+z)}\prod_p\left(1+\frac{z}{p-1}\right)\left(1-\frac{1}{p}\right)^z ,$$

$$G(z) = \frac{1}{\Gamma(1+z)}\prod_p\left(1-\frac{z}{p}\right)^{-1}\left(1-\frac{1}{p}\right)^z ,$$

where the products are over all primes. Note that $G(z)$ has poles at the primes, but $F(z)$ is defined everywhere. The Sathe-Selberg results are that for any $B>0$ we have uniformly for $x\geq 3$ and $1\leq y\leq B\log\log x$

$$(1.2) \quad \pi(x,y) = \frac{x}{\log x}\cdot\frac{(\log\log x)^{y-1}}{(y-1)!}F\left(\frac{y-1}{\log\log x}\right)\left(1+O\left(\frac{y}{(\log\log x)^2}\right)\right)$$

and that for any $\varepsilon>0$, we have uniformly for $x\geq 3$ and $1\leq y\leq(2-\varepsilon)\log\log x$,

$$N(x,y) = \frac{x}{\log x}\cdot\frac{(\log\log x)^{y-1}}{(y-1)!}G\left(\frac{y-1}{\log\log x}\right)\left(1+O\left(\frac{y}{(\log\log x)^2}\right)\right) .$$

Selberg went on to give an asymptotic formula for $N(x,y)$ for $(2+\varepsilon)\log\log x < y < B\log\log x$, stating that the two results could be blended together for $y\approx 2\log\log x$. Nicolas's new result [7] is valid for $(2+\varepsilon)\log\log x < y \leq \log x/\log 2$:

$$N(x,y) = \frac{Cx}{2^y}\log\frac{x}{2^y}\left(1+O\left(\log^b\frac{3x}{2^y}\right)\right)$$

where $b<0$ and

$$C = \frac{1}{4}\prod_{p>2}\left(1+\frac{1}{p(p-2)}\right) .$$

Very little is known about $\pi(x,y)$ if $y/\log\log x\to\infty$. Part of the basic landscape is the Hardy-Ramanujan [3] inequality: there are absolute constants c_1, c_2 with

$$(1.3) \qquad \pi(x,y) \leq c_1\frac{x}{\log x}\cdot\frac{(\log\log x+c_2)^{y-1}}{(y-1)!} , \quad x\geq 3 .$$

This inequality can be proved by induction and without the prime number theorem. A very nice feature of (1.3) is its universality - there are essentially no restrictions placed on x and y.

To complement (1.3) with a lower bound we have an argument of Erdös and Nicolas [2]. Let $N(t)$ denote the product of the first t primes. Then evidently $\omega(pN(y-1)) = y$ for any prime $p \nmid N(y-1)$. Thus we have

$$(1.4) \qquad \pi(x,y) \geq \pi\left(\frac{x}{N(y-1)}\right) - y + 1 .$$

It is perhaps surprising that (1.3) and (1.4), by some measures, are close together for some ranges of the variable y. For example, in [2] it is shown that if $0 < c < 1$ is fixed, then

$$\sum_{y > c \log x/\log\log x} \pi(x,y) = x^{1-c+o(1)} \quad \text{as } x \to \infty .$$

The proof uses (1.4) and an inequality slightly weaker than (1.3). Similar results, sometimes for generalizations of $\pi(x,y)$ that count restricted kinds of prime factors are obtained in [9].

In Kolesnik and Straus [5] an "asymptotic formula" is given for $\pi(x,y)$ valid for $y < (\log x)^{3/5-\varepsilon}$ unconditionally and for $y < (1/2 - \varepsilon)\log x/\log\log x$ assuming the Riemann Hypothesis. The function they prove asymptotic to $\pi(x,y)$ is a sum of many terms, the number of terms depends on the choice of x and y, and the terms are not all of the same sign. It is not clear which, if any of the terms, gives the main contribution. The only thing guaranteed in the theorem is that for the stated range the error term is $o(\pi(x,y))$. Thus with this result it is impossible to tell which of (1.3) and (1.4) is nearer to the truth at $y = [\sqrt{\log x}]$, say.

Recently, Nicolas and Tenenbaum [8] discovered the elegant inequality

$$\sum_{n \leq x} \tau_k(n) \leq \frac{x(\log x + k - 1)^{k-1}}{(k-1)!}$$

valid for all $x \geq 1$ and all integers $k \geq 2$. The function $\tau_k(n)$ is defined as the number of ordered k-tuples $a_1, \ldots, a_k$ of positive integers with $a_1 \ldots a_k = n$. Thus $\tau_2(n)$ is the usual divisor function. The proof of the Nicolas-Tenenbaum inequality can be seen easily by an induction argument. The interest here is that it is easy to show that $k^{\omega(n)} \leq \tau_k(n)$, so that

$$(1.5) \qquad \pi(x,y) \leq k^{-y} \sum_{n \leq x} \tau_k(n) \leq \frac{x(\log x + k - 1)^{k-1}}{k^y (k-1)!}$$

for any integer $k \geq 2$. One can then choose k so as to minimize the right side of (1.5), obtaining an upper bound for $\pi(x,y)$ that is superior to (1.3) for $y \geq (\log\log x)^{1+\varepsilon}$.

The principal results of this paper are also inequalities for $\pi(x,y)$. Let

$$L = L(x,y) = \log\log x - \log y - \log\log y .$$

We show that uniformly for $x \geq 3$ and integers y with

$$(1.6) \qquad (\log\log x)^2 \leq y \leq \frac{\log x}{3 \log\log x},$$

we have

$$(1.7) \qquad \pi(x,y) = \frac{x}{y!} \exp\{y(\log L + \frac{\log L}{L} + O(\frac{1}{L}))\} .$$

We also show that uniformly for $x \geq 3$ and y with

$$(1.8) \qquad y = \left[\frac{c \log x}{\log\log x}\right], \quad \frac{1}{3} \leq c \leq 1 - \frac{1}{\log\log x},$$

we have

$$(1.9) \qquad \pi(x,y) = \frac{x}{y^y}(1-c)^y \exp\{O(y)\} .$$

Also, for

$$y \geq \frac{\log x}{\log\log x} - \frac{\log x}{(\log\log x)^2},$$

we have

$$(1.10) \qquad \pi(x,y) \leq \exp\{O(y)\} .$$

The methods for proving (1.7) could be stretched for smaller values of y than given by (1.6), but the proof gets messier as y gets smaller. Using an idea of Nicolas to introduce a factor $1/\log x$, one can prove by these methods that

$$\pi(x,y) = \frac{x}{y! \log x} \exp\{y(\log L + \frac{\log L}{L} + O(\frac{1}{L}))\}$$

for integers y with

$$(\text{loglog } x)^{1+\varepsilon} \le y \le \frac{\log x}{3 \text{ loglog } x}$$

for any fixed $\varepsilon > 0$. The implied constant depends on the choice of ε. Note that for $y \ge (\text{loglog } x)^2$, the factor $1/\log x$ may be absorbed into the error factor $\exp\{O(y/L)\}$.

To see the quality of the new inequalities (1.7), (1.9), and (1.10), some comparisons are now made with (1.3), (1.4), and (1.5). For $y = [(\text{loglog } x)^c]$, $c \ge 2$ fixed, (1.3) and (1.4) give (using $\log_2 x = \text{loglog } x$, etc.)

$$\frac{x}{y!} \exp\{y(-\log_4 x + O(1))\} \le \pi(x,y)$$

$$\le \frac{x}{y!} \exp\{y(\log_3 x + O(\frac{1}{\log_2 x}))\} ,$$

(1.5) gives

$$\pi(x,y) \le \frac{x}{y!} \exp\{y(\log_3 x - \frac{(c-1)\log_3 x}{\log_2 x} + O(\frac{1}{\log_2 x}))\} ,$$

and (1.7) gives

$$\pi(x,y) = \frac{x}{y!} \exp\{y(\log_3 x - \frac{(c-1)\log_3 x + \log_4 x}{\log_2 x} + O(\frac{1}{\log_2 x}))\} .$$

For $y = [(\log x)^c]$, $0 < c < 1$ fixed, (1.3) and (1.4) give

$$\frac{x}{y!} \exp\{y(-\log_3 x + O(1))\} \le \pi(x,y)$$

$$\le \frac{x}{y!} \exp\{y(\log_3 x + O(\frac{1}{\log_2 x}))\} ,$$

(1.5) gives

$$\pi(x,y) \le \frac{x}{y!} \exp\{y(\log_3 x + \log(1-c) + O(\frac{\log_3 x}{\log_2 x}))\} ,$$

and (1.7) gives

$$\pi(x,y) = \frac{x}{y!} \exp\{y(\log_3 x + \log(1-c) + O(\frac{1}{\log_2 x}))\} .$$

For $y = [c \log x/\text{loglog } x]$, $0 < c < 1$ fixed, (1.3) and (1.4) give

$$x^{1-c} \exp\{O(y)\} \le \pi(x,y) \le x^{1-c} \exp\{y(2 \log_3 x + O(1))\} ,$$

(1.5) gives

$$\pi(x,y) \le x^{1-c} \exp\{y(\log_3 x + O(\log_4 x))\} ,$$

and (1.7) and (1.9) give

$$\pi(x,y) = x^{1-c} \exp\{y(\log_3 x + O(1))\} .$$

Finally at $y = [\log x/\text{loglog } x]$, (1.3) and (1.4) give

$$\exp\{O(y)\} \le \pi(x,y) \le \exp\{O(y \log_3 x)\} ,$$

(1.5) gives

$$\pi(x,y) \le \exp\{O(y \log_4 x)\} ,$$

and (1.10) gives

$$\pi(x,y) \le \exp\{O(y)\} .$$

One question which can be partially answered here is the issue of the range of y for which (1.2) is valid. This equation has been proved only when $y \le B \text{ loglog } x$ where B is an arbitrary but fixed constant. However, the right side continues to make sense for larger values of y and one might wonder how close an approximation it gives for large y. From (1.7) there is a constant $c_0 > 0$ such that (1.2) is false for all $y \ge \exp(c_0\sqrt{\text{loglog } x})$. It is probable that a refinement of the proof of (1.7) would show (1.2) is false for all $y \ge \exp(\varepsilon\sqrt{\text{loglog } x})$ for any fixed $\varepsilon > 0$, but we do not undertake this project here.

§2. The basic ideas.

Underlying our results in the subsequent sections are some essentially simple combinatorial methods. In this section these basic ideas are described.

In sections 3 and 4, instead of directly estimating $\pi(x,y)$, we instead estimate the function

$$s(x,y) = \sum_{\substack{n \le x \\ \omega(n)=y}} \frac{1}{n}.$$

It is trivial to pass from an upper bound for $s(x,y)$ to one for $\pi(x,y)$ via the inequality

$$\pi(x,y) \le x s(x,y). \tag{2.1}$$

We can also use lower bounds for $s(x,y)$:

Proposition 2.1. If x is sufficiently large and $y > 0$ is an integer, then

$$\pi(x,y) \ge \frac{1}{3y \log x} s\left(\frac{x}{3 \log x}, y-1\right).$$

Proof. We have for x sufficiently large

$$\pi(x,y) = \frac{1}{y} \sum_{\substack{m \le x \\ \omega(m)=y-1}} \sum_{\substack{p^a \le x/m \\ (p,m)=1}} 1$$

$$\ge \frac{1}{y} \sum_{\substack{m \le x \\ \omega(m)=y-1}} (\pi(x/m) - (y-1))$$

$$\ge \frac{1}{2y} \sum_{\substack{m \le x/3 \log x \\ \omega(m)=y-1}} \pi(x/m)$$

$$\ge \frac{x}{3y \log x} s\left(\frac{x}{3 \log x}, y-1\right),$$

which establishes the proposition.

An upper bound for $s(x,y)$ is evidently obtained from the inequality

$$s(x,y) \le \Big(\sum_{p^a \le x} p^{-a}\Big)^y / y!, \tag{2.2}$$

and from this an estimate similar to the Hardy-Ramanujan inequality (1.3) can be derived. There is waste in the inequality (2.2) from several sources. One such source is that the "square-free" terms

formed by using the multinomial theorem to expand the right side of (2.2) are reciprocals of integers n with $\omega(n) = y$. Every n with $n \le x$ and $\omega(n) = y$ occurs this way, but many larger values of n also occur.

This source of error can be ameliorated by partitioning the interval $(1,x]$. If $n \le x$ and $\omega(n) = y$, then the number of $p^a \| n$ with $p^a \ge x^{w/y}$ is at most y/w. Thus, for any choice of $w > 1$ we have

$$s(x,y) \le \sum_{0 \le i \le y/w} \frac{1}{i!}\Big(\sum_{x^{w/y} \le p^a \le x} p^{-a}\Big)^i \frac{1}{(y-i)!}\Big(\sum_{p^a < x^{w/y}} p^{-a}\Big)^{y-i}$$

and choosing w optimally gives an estimate that is much better than (2.2).

Continuing with this idea, we partition $(1,x]$ into many intervals and obtain finer estimates. It turns out that the interval which gives the principal contribution is $(y, x^{1/y}]$. Note that

$$\sum_{y < p^a \le x^{1/y}} p^{-a} = \log\log x - \log y - \log\log y + o(1),$$

which explains why it is convenient to state the result (1.7) in terms of the function L.

A lower bound for $s(x,y)$ can be obtained by examining the expression

$$\frac{1}{y!}\Big(\sum_{y^2 < p \le x^{1/y}} p^{-1}\Big)^y.$$

Using the multinomial theorem to expand the power we find "good" terms $1/n$ where n is square-free, $\omega(n) = y$, $n \le x$ and "bad" terms c/n where $\omega(n) < y$ and the coefficient c is at most $1/2$. Since these bad numbers n are divisible by the square of a large prime (at least y^2), the sum of their reciprocals can be shown to be negligible provided

$$\sum_{y^2 < p \le x^{1/y}} p^{-1}$$

is bounded away from 0. This sum is about $L - \log 2$ and is bounded away from 0 if $y < (1/2 - \varepsilon)\log x/\log\log x$. This partially explains why the upper bound in (1.6) is needed in the proof of (1.7).

Using only this idea, a much better lower bound estimate than (1.4) can be proved. We get a slightly better result by also considering contributions from higher intervals.

It is interesting to note that the lower bound estimates we obtain in (1.7) and (1.9) are also valid for the smaller function

$$\pi_0(x,y) = \sum_{n\le x,\, \omega(n)=y} \mu^2(n) .$$

Concerning the lower bound in (1.6), the methods of this paper are valid for smaller values of y, say down to $2 \log\log x$, but the situation gets more complicated. It may, in fact, be possible to extend (1.2) somewhat, so it seems reasonable to leave the estimates for smaller values of y to that line of attack.

Our results are not very good when y is larger than $\log x/\log\log x$. It is possible that the methods of [2], Theorem 1 are applicable here.

An outline of the remaining sections is as follows. In section 3 we establish the lower bound in (1.7) using the attack outlined above. In section 4 we obtain the corresponding upper bound. In section 5 we prove the lower bound in (1.9) by a different method: we directly count integers $n \le x$ with $\omega(n) = y$, rather than work with the sum of their reciprocals. In section 6 we establish the upper bounds in both (1.9) and (1.10), again by directly counting the relevant integers.

§3. A lower bound.

In this section we establish a result that is a little stronger than the lower bound implicit in (1.7).

<u>Theorem 3.1</u>. There is an absolute constant x_0 such that for all $x \ge x_0$ and integers y with

$$\log\log x (\log\log\log x)^2 \le y \le \frac{\log x}{3 \log\log x} ,$$

we have uniformly

$$\pi(x,y) \ge \frac{x}{y!\log x} \exp\{y(\log L + \frac{\log L}{L} + O(\tfrac{1}{L}))\}$$

where $L = \log\log x - \log y - \log\log y$.

<u>Proof</u>. Let $L' = L + 20$ and let $k = [\log L'] - 2$. Let

$$I_{-1} = (y^2, x^{2e^{-1}/y}] , \quad I_i = (x^{2e^{i-1}/y}, x^{2e^i/y}] \text{ for } i = 0,\dots,k-1.$$

Let

$$s_i(u,v) = \sum_{\substack{n\le u,\, \omega(n)=v \\ p|n \Rightarrow p\in I_i}} \frac{1}{n} .$$

Then if $u_{-1}u_0\cdots u_{k-1} \le x$ and $v_{-1} + v_0 +\dots+ v_{k-1} = y$, we have

$$s(x,y) \ge \prod_{i=-1}^{k-1} s_i(u_i,v_i) . \tag{3.1}$$

We apply inequality (3.1) with

$$u_{-1} = x^{2e^{-1}} , \qquad u_i = x^{2e^i/L'} , \text{ for } i = 0,\dots,k-1$$

(3.2)

$$v_{-1} = y-k[y/L'] , \qquad v_i = [y/L'] , \text{ for } i = 0,\dots,k-1 .$$

Note that if n_i is the product of v_i primes in I_i, then necessarily $n \le u_i$ for $i = -1,0,\dots,k-1$. To see this, first note that for $i = -1$, we have

$$n_{-1} \le (x^{2e^{-1}/y})^{y-k[y/L']} \le x^{2e^{-1}(1-k/L' + k/y)} \le u_{-1}$$

since for $x \ge x_0$, $y \ge L'$. For $i = 0,\dots,k-1$, we have

$$n_i \le (x^{2e^i/y})^{[y/L']} \le x^{2e^i/L'} = u_i .$$

Next note that

$$u_{-1}u_0\cdots u_{k-1} \le x \text{ and } v_{-1} + v_0 +\dots+ v_{k-1} = y .$$

The second statement is obvious. For the first, we have

$$u_{-1}u_0\cdots u_{k-1} \le x^{2e^{-1} + 2(e^k-1)/(e-1)L'}$$

and

$$2e^{-1} + 2e^{k}/(e-1)L' \le 2e^{-1} + 2e^{-2}(e-1)^{-1} < 1$$

by our choice of k. It is thus valid to apply the inequality (3.1) with the choices of u_i, v_i given in (3.2).

Consider the expression

$$\frac{1}{v_i!}\Big(\sum_{p\in I_i} \frac{1}{p}\Big)^{v_i} .$$

Multiplying this out, we have a sum of reciprocals of integers $n_i \le u_i$ where each prime in n_i is in I_i and $\Omega(n_i) = v_i$. If n_i is square-free (so that $\omega(n_i) = v_i$), the coefficient of $1/n_i$ is 1. For the non-square-free n_i's, the coefficient is at most 1/2. Therefore that part of the sum consisting of the non-square-free n_i's is majorized by

$$\frac{1}{2}\Big(\sum_{p\in I_i} \frac{1}{p^2}\Big)\frac{1}{(v_i-2)!}\Big(\sum_{p\in I_i} \frac{1}{p}\Big)^{v_i-2}$$

$$= \frac{1}{(v_i-2)!}\Big(\sum_{p\in I_i} \frac{1}{p}\Big)^{v_i-2} O\Big(\frac{1}{y^2\log y}\Big)$$

$$= O\Big(\frac{1}{v_i!\log y}\Big(\sum_{p\in I_i} \frac{1}{p}\Big)^{v_i}\Big)$$

where the last equality follows from the fact that $y \le \log x/3 \operatorname{loglog} x$ implies $\Sigma_{p\in I_i} 1/p \ge c > 0$ for some absolute constant c for $x \ge x_0$. This estimate is uniform for $i = -1,0,\dots,k-1$.

Therefore we have

$$s_i(u_i,v_i) \ge \Big(1 + O\Big(\frac{1}{\log y}\Big)\Big)\frac{1}{v_i!}\Big(\sum_{p\in I_i} \frac{1}{p}\Big)^{v_i}$$

so that from (3.1) we have

$$(3.3) \qquad s(x,y) \ge c \prod_{i=-1}^{k-1} \frac{1}{v_i!}\Big(\sum_{p\in I_i} \frac{1}{p}\Big)^{v_i}$$

for some absolute constant $c > 0$.

We first estimate the factorials in (3.3). By Stirling's formula we have

$$\log \prod_{i=-1}^{k-1} v_i! = \sum_{i=-1}^{k-1} (v_i \log v_i - v_i + O(\log v_i))$$

$$= y\Big(1 - \frac{k}{L'}\Big)\Big(\log y + \log\Big(1 - \frac{k}{L'}\Big) - 1\Big) + \frac{ky}{L'}(\log y - \log L' - 1)$$

$$+ O(k \log(y/L) + \log y)$$

$$= y \log y - y + y\Big(\log\Big(1 - \frac{k}{L'}\Big) - \frac{k}{L'} \log\Big(1 - \frac{k}{L'}\Big) - \frac{k}{L'} \log L'\Big)$$

$$+ O(k \log(y/L) + \log y)$$

$$(3.4) \qquad = \log y! + y\Big(-\frac{k}{L} - \frac{k}{L} \log L + O\Big(\frac{k^2}{L^2}\Big)\Big) .$$

Next note that by the prime number theorem with error term, we have

$$\log\Big(\sum_{p\in I_{-1}} \frac{1}{p}\Big)^{v_{-1}} = v_{-1} \log(L - 1 + O(e^{-c\sqrt{\log y}}))$$

$$= y\Big(1 - \frac{k}{L'}\Big)\log L + O(k \log L) + O(y/L)$$

$$(3.5) \qquad = y\Big(\log L - \frac{k}{L} \log L + O\Big(\frac{1}{L}\Big)\Big)$$

and

$$\sum_{i=0}^{k-1} \log\Big(\sum_{p\in I_i} \frac{1}{p}\Big)^{v_i} = \sum_{i=0}^{k-1} v_i \log(1 + O(e^{-c\sqrt{(\log x)/y}}))$$

$$= O(y\, e^{-c\sqrt{\operatorname{loglog} x}}) = O(y/L) .$$

Putting these estimates and (3.4) into (3.3), we have

$$s(x,y) \ge \frac{1}{y!} \exp\{y(\log L + \frac{k}{L} + O(\frac{1}{L}))\} .$$

Our theorem now follows from Proposition 2.1.

§4. An upper bound for the same range.

In this section we complete our proof of (1.7).

Theorem 4.1. Uniformly for all $x \ge 3$ and integers $y \ge 3$ satisfying

$$(\log\log x)^2 \le y \le \frac{\log x}{3\log\log x}$$

we have

$$\pi(x,y) \le \frac{x}{y!}\exp\{y(\log L + \frac{\log L}{L} + o(\tfrac{1}{L}))\}$$

where $L = \log\log x - \log y - \log\log y$.

Proof. We may assume x is large. Let

$$K = [\log y] + 1 .$$

Let

$$I_{-2} = (0, y^{1/2}] , \quad I_{-1} = (y^{1/2}, y] , \quad I_0 = (y, x^{1/y}] ,$$

$$I_i = (x^{e^{i-1}/y}, x^{e^i/y}] \quad \text{for} \quad i = 1,\dots,K .$$

Then

$$(4.1) \qquad s(x,y) \le \sum_{a_{-2},\dots,a_K}{}' \prod_{i=-2}^{K} \frac{1}{a_i!}\Big(\sum_{p^a\in I_i} \frac{1}{p^a}\Big)^{a_i}$$

where Σ' indicates the sum is over all ordered partitions of $y = a_{-2} + \dots + a_K$ into non-negative summands where

$$(4.2) \qquad a_{-2} \le (2 + o(1))\sqrt{y}/\log y , \quad a_{-1} \le (1 + o(1))y/\log y ,$$

$$\text{and} \quad y^{a_0} \prod_{i=1}^{K} (x^{e^{i-1}/y})^{a_i} \le x .$$

Our strategy is to majorize the summand of Σ' in (4.1) over all allowable ordered partitions and then multiply this upper bound by the number of such partitions.

We let

$$\alpha_i = a_i/y .$$

Using the inequality $1/u! \le e^u u^{-u}$, we have

$$(4.3) \qquad \prod_{i=-2}^{K} \frac{1}{a_i!} \le \exp\{\sum_{i=-2}^{K} (a_i - a_i \log a_i)\}$$

$$= \exp\{y(1 - \log y - \sum_{i=-2}^{K} \alpha_i \log \alpha_i)\}$$

where we interpret $0 \log 0$ as 0 should some $\alpha_i = 0$.

Note that

$$\sum_{p^a\in I_{-2}} p^{-a} = \log\log y + O(1) ,$$

$$\sum_{p^a\in I_{-1}} p^{-a} = \log 2 + O(e^{-c\sqrt{\log y}}) < 1 , \quad \text{for} \quad x \ge x_0 ,$$

$$\sum_{p^a\in I_0} p^{-a} = L + O(e^{-c\sqrt{\log y}}) ,$$

$$\sum_{p^a\in I_i} p^{-a} = 1 + O(e^{-c\sqrt{(\log x)/y}}) , \quad \text{for} \quad i = 1,\dots,K .$$

Therefore

$$(4.4) \qquad \prod_{i=-2}^{K} \Big(\sum_{p^a\in I_i} p^{-a}\Big)^{a_i} \le \exp\{y(\alpha_0 \log L + o(\tfrac{1}{L}))\} .$$

Indeed from (4.2) we have

$$\alpha_{-2} \log(\log\log y + O(1)) = o(\tfrac{1}{L}) ,$$

$$\alpha_{-1} \log(\log 2 + O(e^{-c\sqrt{\log y}})) \le 0 ,$$

$$\sum_{i=1}^{K} \alpha_i \log(1 + O(e^{-c\sqrt{(\log x)/y}})) = O(e^{-c\sqrt{\log\log x}}) = o(\tfrac{1}{L}) .$$

Therefore, from (4.3) and (4.4) we have

$$(4.5) \qquad \prod_{i=-2}^{K} \frac{1}{a_i!}\Big(\sum_{p^a\in I_i} \frac{1}{p^a}\Big)^{a_i} \le \frac{1}{y!}\exp\{y(\alpha_0 \log L - \sum_{i=-2}^{K} \alpha_i \log\alpha_i + o(\tfrac{1}{L}))\} .$$

Thus we should consider the maximum value taken by the expression in the α's. We have the

Lemma. Let

$$F(\alpha_{-2},\dots,\alpha_K) = \alpha_0 \log L - \sum_{i=-2}^{K} \alpha_i \log \alpha_i$$

where the $\alpha_i \ge 0$ satisfy the constraints

$$\sum_{i=-2}^{K} \alpha_i = 1 \quad \text{and} \quad \sum_{i=1}^{K} e^{i-1} \alpha_i \le 1 .$$

Then

$$F(\alpha_{-2},\dots,\alpha_K) \le \log L + \frac{\log L}{L} + O(\tfrac{1}{L}) .$$

Before we prove the lemma, we note how the theorem follows almost immediately from it and (4.5). Indeed all we need do is multiply by the number of ordered partitions $a_{-2}, a_{-1}, \dots, a_K$ and this number is majorized by

$$(y+1)^{K+3} = \exp\{O((\log y)^2)\} = \exp\{o(y/L)\} .$$

To prove the lemma we first show we may take L large. Indeed when L is bounded the only assertion of the lemma is that F is bounded. And this assertion is easy to see, since the contribution to F from the α_i with $i \ge 2$ is at most

$$\sum_{i\ge 2} (i-1)e^{-i+1} = O(1) .$$

In proving the lemma we shall find it convenient to replace F with a larger function G which has infinitely many variables. This is really only to streamline the arguments since it is easy to see that the contribution to F from the α_i with $i \ge 2 \log L$ is negligible. Let

$$G(\alpha_{-2},\alpha_{-1},\dots) = \alpha_0 \log L - \sum_{i\ge -2} \alpha_i \log \alpha_i$$

where each $\alpha_i \ge 0$,

$$(4.6) \qquad \sum_{i\ge -2} \alpha_i = 1 \quad \text{and} \quad \sum_{i\ge 1} \alpha_i e^{i-1} = 1 .$$

Then the maximum value of F is at most the maximum value of G.

We next note that if G attains its maximum value at $(\alpha_{-2},\alpha_{-1},\dots)$ then each $\alpha_i > 0$. We show this for $i \ge 2$, the proof of the remaining cases is simpler. So suppose $i \ge 2$, $\alpha_i = 0$ and $\alpha_{i-1} > 0$. Let

$$\alpha_0' = \alpha_0 + (1 - \tfrac{1}{e})\alpha_{i-1}^e ,$$

$$\alpha_{i-1}' = \alpha_{i-1} - \alpha_{i-1}^e ,$$

$$\alpha_i' = \tfrac{1}{e}\alpha_{i-1}^e ,$$

$$\alpha_j' = \alpha_j \quad \text{for } j \ne 0, i-1, i .$$

Then $\alpha_{-2}', \alpha_{-1}', \dots$ satisfy the constraints (4.6) and $G(\alpha_{-2}',\alpha_{-1}',\dots) > G(\alpha_{-2},\alpha_{-1},\dots)$.

Thus the maximum point of G occurs at an interior point and may be located by the method of Lagrange multipliers. This method gives our maximum point the shape

$$\alpha_{-2} = \alpha_{-1} = e^{-\lambda}, \quad \alpha_0 = e^{-\lambda} L, \quad \alpha_i = e^{-\lambda}e^{-\mu e^{i-1}} \quad \text{for } i \ge 1 .$$

The constraints (4.6) give us the equations

$$(4.7) \qquad 2 + L + \sum_{j\ge 0} e^{-\mu e^j} = e^{\lambda} = \sum_{j\ge 0} e^j e^{-\mu e^j}$$

and at this point we have

$$(4.8) \qquad G(\alpha_{-2},\alpha_{-1},\dots) = \lambda + \mu .$$

Thus it remains to estimate $\lambda + \mu$. First note we may assume $0 < \mu \le 1$, for if $\mu > 1$, (4.7) implies L is bounded above (in fact, (4.7) implies $L < 0$). Thus we have

(4.9) $$\sum_{j\ge 0} e^{-\mu e^j} = \int_0^\infty e^{-\mu e^t}\,dt + O(1)$$

$$= \int_0^{\log 1/\mu} e^{-\mu e^t}\,dt + O(1)$$

$$= \int_0^{\log 1/\mu} (1 + O(\mu e^t))\,dt + O(1)$$

$$= \log 1/\mu + O(1) .$$

The function $f(t) = e^t e^{-\mu e^t}$ on $[0,\infty)$ increases to its maximum value of $1/e\mu$ at $t = \log 1/\mu$ and then decreases. The value at $1 + \log 1/\mu$ is e^{1-e}/μ. Therefore

$$\frac{e^{1-e}}{\mu} < \sum_{j\ge 0} e^j e^{-\mu e^j} \le \int_0^\infty e^t e^{-\mu e^t}\,dt + O(\tfrac{1}{\mu}) = O(\tfrac{1}{\mu}) ,$$

so that from (4.7) we have

(4.10) $$e^\lambda \asymp \frac{1}{\mu} .$$

Thus from (4.7), (4.9), (4.10), and our assumption that L is large, we have

(4.11) $$L \asymp \frac{1}{\mu}$$

so that

(4.12) $$\log L = \log 1/\mu + O(1) .$$

Thus from (4.7), (4.9), and (4.12) we have

$$\lambda = \log(L + \log L + O(1)) = \log L + \frac{\log L}{L} + O(\tfrac{1}{L})$$

and the lemma and the theorem now follow from (4.8) and (4.11).

§5. A lower bound for the distribution of rounder numbers.

In this section we establish the lower bound in (1.9) .

Theorem 5.1. There is an absolute constant A such that for $y = [c \log x/\log\log x]$ where

$$1/3 \le c \le 1 - 1/\log\log x$$

we have uniformly

$$\pi(x,y) \ge x\, y^{-y} (1-c)^y e^{Ay}$$

for $x \ge 3$.

Proof. Let $T = \frac{1}{y}\log x - \log\log x$, so that

$$1 < T \le 2\log\log x + o(1) .$$

Let

$$k = [(\log\log x)^2] - 1 , \quad \alpha = \frac{1}{T+1} , \quad \beta = \frac{T}{T+1} .$$

Let

$$\alpha_0 = 1 - \sum_{i=1}^{k} \alpha\beta^i ,$$

so that

$$\alpha_0 = 1 - \alpha\beta\,\frac{1-\beta^k}{1-\beta} = 1 - \beta(1-\beta^k) > 1 - \beta = \alpha > 0 .$$

We have

(5.1) $$0 \le T - \sum_{i=1}^{k} i\alpha\beta^i = o(1) \qquad (\text{as } x \to \infty).$$

Indeed

$$\sum_{i=1}^{k} i\,\alpha\beta^i = \alpha\beta\,\frac{1-\beta^{k+1} - (k+1)\beta^k(1-\beta)}{(1-\beta)^2}$$

$$= T - T\beta^{k+1} - (k+1)\beta^{k+1} .$$

Thus (5.1) will be proved if we show $\beta^{k+1} = o(1/k)$ as $x \to \infty$. But

$$\log \beta^{k+1} \le (\log\log x)^2 \log(1 - \frac{1}{T+1})$$

$$< -(\log\log x)^2/(T+1)$$

$$\le -\frac{1}{2}\log\log x + O(1) ,$$

so that

$$(5.2) \qquad \beta^{k+1} = O(1/\sqrt{\log x}) = o(1/k) .$$

Now let

$$I_0 = (0, \log x] , \quad I_i = (e^{i-1}\log x, e^i \log x] \quad \text{for } i = 1,\dots,k .$$

Let

$$a_i = [\alpha\beta^i y] \quad \text{for } i = 1,\dots,k , \quad a_0 = y - \sum_{i=1}^{k} a_i .$$

Let $\pi(I)$ denote the number of primes in the interval I. We have

$$\pi(I_i) \ge c_1 \frac{e^i \log x}{i + \log\log x} \quad \text{for } i = 0,1,\dots,k$$

where $c_1 > 0$ is some absolute constant.

Say we construct an integer n by choosing a_i different primes in I_i and multiply them all together for $i = 0,1,\dots,k$. Then

$$n \le \prod_{i=0}^{k} (e^i \log x)^{a_i} = (\log x)^y \exp\{\sum_{i=1}^{k} i\, a_i\}$$

$$\le (\log x)^y e^{Ty} = x ,$$

by (5.1). Also, each such n has $\omega(n) = y$. Thus $\pi(x,y)$ is at least the number of choices of such n. That is,

$$(5.3) \quad \pi(x,y) \ge \prod_{i=0}^{k} \binom{\pi(I_i)}{a_i} \ge \prod_{i=0}^{k} \left(\frac{\pi(I_i)}{a_i}\right)^{a_i}$$

$$\ge \prod_{i=0}^{k} \left(\frac{c_1 e^i \log x}{a_i (i + \log\log x)}\right)^{a_i}$$

$$= \exp\{\sum_{i=0}^{k} (-a_i \log a_i + i\, a_i - a_i \log(i + \log\log x))\}(c_1 \log x)^y .$$

From (5.1) we have

$$(5.4) \quad \sum_{i=1}^{k} i\, a_i = \sum_{i=1}^{k} i[\alpha\beta^i y] = Ty + o(y) + O(\sum_{i=1}^{k} i) = Ty + o(y) .$$

We now show that

$$(5.5) \quad \sum_{i=0}^{k} a_i \log(i + \log\log x) = y \log\log\log x + O(y) .$$

Indeed, this sum is

$$\sum_{j=0}^{[k/\log\log x]} \sum_{\substack{[i/\log\log x]=j \\ i\le k}} a_i \log((j+1)\log\log x) + O(\sum_{i=0}^{k} a_i)$$

$$= \sum_{i=0}^{k} a_i \log\log\log x + \sum_j \sum_i a_i \log(j+1) + O(y)$$

$$= y \log\log\log x + y \sum_{j\ne 0} \sum_i \alpha\beta^i \log j + O(y)$$

$$= y \log\log\log x + O(y \sum_{j\ne 0} \beta^{[j \log\log x]} \log j) + O(y)$$

$$= y \log\log\log x + O(y \sum_{j\ne 0} (1-c_2)^j \log j) + O(y)$$

$$= y \log\log\log x + O(y)$$

where $1 > c_2 > 0$ is an absolute constant.

Completing the estimate of the right side of (5.3) we are going to show that

$$(5.6) \quad -\sum_{i=0}^{k} a_i \log a_i = -y \log y + y \log T + O(y) .$$

But first we note that the theorem will follow immediately from (5.3)-(5.6) , for we have

$$\pi(x,y) \geq \exp\{y(-\log y + \log T + T - \log\log\log x + \log\log x + O(1))\}$$

$$= x\,y^{-y}\exp\{y(\log T - \log\log\log x + O(1))\}$$

$$= x\,y^{-y}\exp\{y\log(1/c - 1) + O(y)\}$$

$$= x\,y^{-y}(1-c)^{y}\exp\{O(y)\} .$$

To see (5.6), we first estimate the error introduced by replacing $a_0 \log a_0$ with $\alpha_0 y \log(\alpha_0 y)$ and $a_i \log a_i$ for $i = 1,\dots,k$ with $\alpha\beta^i y \log(\alpha\beta^i y)$. If $\alpha\beta^i y < 1$, then

$$|\alpha\beta^i y\log(\alpha\beta^i y) - a_i \log a_i| = |\alpha\beta^i y \log(\alpha\beta^i y)| \leq 1/e .$$

If $\alpha\beta^i y \geq 1$, then

$$|\alpha\beta^i y\log(\alpha\beta^i y) - a_i \log a_i| \leq 1 + \log(\alpha\beta^i y) < 1 + \log y .$$

Thus

$$(5.7)\qquad \Big|\sum_{i=1}^{k} \alpha\beta^i y\log(\alpha\beta^i y) - \sum_{i=1}^{k} a_i \log a_i\Big| = O(k \log y) = o(y) .$$

Next note that $\alpha y < \alpha_0 y \leq a_0 < \alpha_0 y + k$, so that

$$(5.8)\qquad |\alpha_0 y\log(\alpha_0 y) - a_0 \log a_0| \ll k \log y = o(y).$$

From (5.7) and (5.8) we have

$$(5.9)\qquad -\sum_{i=0}^{k} a_i \log a_i = -\alpha_0 y\log(\alpha_0 y) - \sum_{i=1}^{k} \alpha\beta^i y\log(\alpha\beta^i y) + o(y)$$

$$= -y(\alpha_0(\log \alpha_0 + \log y) + \sum_{i=1}^{k} \alpha\beta^i(\log\alpha + i\log\beta + \log y)) + o(y)$$

$$= -y(\log y + \alpha_0\log\alpha_0 + \beta\log\alpha - \beta^{k+1}\log\alpha + T\log\beta) + o(y)$$

using (5.1). We now use (5.2) and the definitions of our letters to see that

$$\alpha_0 \log \alpha_0 = O(1) , \quad \beta \log \alpha = -\log T + O(1)$$

$$-\beta^{k+1}\log\alpha = O\Big(\frac{\log T}{\sqrt{\log x}}\Big) = o(1) , \quad T\log\beta = O(1) .$$

Putting these estimates into (5.9) we have (5.6) and thus the theorem.

§6. An upper bound for the distribution of the roundest numbers.

In this section we show the upper bound implicit in (1.9) and also the inequality (1.10). Although these are stated as separate results, there is a common proof.

<u>Theorem 6.1</u>. Uniformly for $1/3 \leq c \leq 1 - 1/\log\log x$, we have, with $y = [c \log x/\log\log x]$,

$$\pi(x,y) \leq x\,y^{-y}(1-c)^{y}e^{O(y)} .$$

Uniformly for $y \geq (1 - 1/\log\log x)\log x/\log\log x$, we have

$$\pi(x,y) \leq e^{O(y)} .$$

<u>Proof</u>. Consider the intervals

$$I_0 = (1, \log x] , \quad I_i = (e^{i-1}\log x, e^{i}\log x] \quad \text{for } i = 1,2,\dots,k$$

where $k = [\log x]$. Fix an ordered partition of y into non-negative summands

$$(6.1)\qquad y = a_0 + a_1 + \dots + a_k .$$

We first count $N(a_0,\dots,a_k)$, the number of $n \leq x$ with $\omega(n) = y$ and $\sum_{p|n,\,p\in I_i} 1 = a_i$ for each $i = 0,\dots,k$. In fact we shall get and upper bound for

$$N = \max_{a_0,\dots,a_k} N(a_0,\dots,a_k)$$

and then multiply N by the number of ordered partitions (6.1) for which $N(a_0,\ldots,a_k) > 0$ to obtain an upper bound for $\pi(x,y)$.

Define α_i for $i = 0,1,\ldots,k$ by the equation

$$a_i = \alpha_i y .$$

In order for $N(a_0,\ldots,a_k)$ to be positive, the α_i's must satisfy

$$(6.2)\qquad \sum_{i=0}^{k} \alpha_i = 1 , \quad \sum_{i=1}^{k} (i-1)\alpha_i \le \frac{1}{y}(\log x - (y-a_0)\log\log x - \log M(a_0))$$

where $M(t)$ is the product of the first t primes. To see the latter statement in (6.2), note that if $n \le x$ is counted by $N(a_0,\ldots,a_k)$, then

$$x \ge M(a_0)\prod_{i=1}^{k} (e^{i-1}\log x)^{a_i} = M(a_0)\exp\{\sum_{i=1}^{k}(i-1)\alpha_i y\}(\log x)^{y-a_0} ,$$

so that (6.2) immediately follows.

We next note that for $x \ge x_0$,

$$(6.3)\quad a_0 \log\log x - \log M(a_0) \le y(\log\log x - \log y - \log\log y + 1) .$$

Indeed, there is a t_0 so that for $t \ge t_0$ we have

$$\log M(t) \ge t(\log t + \log\log t) .$$

If x_0 is large enough, then (6.3) is true for $a_0 < t_0$. For $a_0 \ge t_0$, we have

$$a_0 \log\log x - \log M(a_0) \le a_0(\log\log x - \log a_0 - \log\log a_0)$$

$$= -\alpha_0 y \log \alpha_0 + a_0(\log\log x - \log y - \log\log a_0) .$$

Since the latter term is increasing in the variable a_0 for $2 \le a_0 \le y$ and since $-\alpha_0 \log \alpha_0 \le 1/e$, we have (6.3).

Since our theorem is true for x small, we assume that x is large, and in particular, that (6.3) holds. Putting (6.3) in (6.2), we have

$$(6.4)\quad \sum_{i=0}^{k} \alpha_i = 1 , \quad \sum_{i=1}^{k} (i-1)\alpha_i \le \frac{1}{y}\log x - \log y - \log\log y + 1 .$$

We now prove that if $\alpha_0,\alpha_1,\ldots,\alpha_k$ are non-negative and satisfy

$$(6.5)\qquad \sum_{i=0}^{k} \alpha_i = 1 \quad\text{and}\quad \sum_{i=1}^{k}(i-1)\alpha_i = T - 1 ,$$

where $T \ge 1$, then

$$(6.6)\qquad -\sum_{i=0}^{k} \alpha_i \log \alpha_i \le \log T + 2 ,$$

where we interpret $0 \log 0$ as 0. The proof is similar to, but easier than the proof of the lemma in section 4. We use the method of Lagrange multipliers. The choice of $\alpha_0,\ldots,\alpha_k$ which maximizes $-\sum \alpha_i \log \alpha_i$ either has some $\alpha_i = 0$ or 1 or there is a λ and μ with

$$-\log \alpha_i - 1 = \lambda + \mu(i-1) , \quad i = 0,1,\ldots,k .$$

These equations imply the α_i are in geometric progression, so there is an α, β with

$$(6.7)\qquad \alpha_i = \alpha\beta^i , \quad i = 0,1,\ldots,k .$$

As in the proof of the lemma in section 4, it is easy to show that the maximum occurs at a point where all α_i satisfy $0 < \alpha_i < 1$, so we may assume (6.7) holds. Moreover, it is clear that at the maximum point we have $\alpha_0 \le \alpha_1 \le \cdots \le \alpha_k$, so that we have $\beta \le 1$. If $\beta = 1$, then the α_i are all equal to α and (6.5) implies $\alpha = 1/(k+1)$ and $k(k-1)/2(k+1) = T - 1$, so that

$$-\sum \alpha_i \log \alpha_i = \log(k+1) < \log T + 1 .$$

Thus we may assume $0 < \beta < 1$.

It is also clear that the maximum of $-\sum \alpha_i \log \alpha_i$ increases with k , so we now assume (6.7) is an infinite geometric progression. Thus (6.5) translates to

$$1 = \alpha/(1-\beta) , \quad T = \sum_{i=0}^{\infty} i\,\alpha_i = \alpha\beta/(1-\beta)^2 .$$

Solving for α and β , we have

$$\alpha = \frac{1}{T+1} , \quad \beta = \frac{T}{T+1} .$$

Therefore

$$-\sum_{i=0}^{\infty} \alpha_i \log \alpha_i = -\sum \alpha\beta^i(\log \alpha + i \log \beta)$$

$$= -\log \alpha - T \log \beta$$

$$= \log(T+1) - T \log \frac{T}{T+1} < \log T + 2 .$$

Thus (6.6) is proved.

Let $\pi'(I)$ denote the number of prime powers in the interval I. Then for α_i satisfying (6.4) we have

$$\text{(6.8)} \quad N(a_0,\dots,a_k) \le \prod_{i=0}^{k} \binom{\pi'(I_i)}{a_i} \le \binom{2 \log x/\operatorname{loglog} x}{a_0} \prod_{i=1}^{k} \binom{e^i \log x/\operatorname{loglog} x}{a_i}$$

$$\le \Big(\prod_{i=0}^{k} 1/a_i!\Big)\Big(\frac{2 \log x}{\operatorname{loglog} x}\Big)^{a_0} \prod_{i=1}^{k} \Big(\frac{e^i \log x}{\operatorname{loglog} x}\Big)^{a_i}$$

$$= \Big(\prod_{i=0}^{k} 1/a_i!\Big)\Big(\frac{\log x}{\operatorname{loglog} x}\Big)^{y} 2^{a_0} \exp\Big(\sum_{i=1}^{k} i\, \alpha_i\, y\Big)$$

$$\le \Big(\prod_{i=0}^{k} 1/a_i!\Big)\Big(\frac{\log x}{\operatorname{loglog} x}\Big)^{y} \exp(\log x - y \log y - y \operatorname{loglog} y + 2y) .$$

Now

$$\prod_{i=0}^{k} 1/a_i! < \prod_{i=0}^{k} \exp\{-a_i(\log a_i - 1)\}$$

$$= \exp\{-y \log y + y - \sum_{i=0}^{k} \alpha_i\, y \log \alpha_i\}$$

$$\le \exp\{-y \log y + y \log(\tfrac{1}{y} \log x - \log y - \operatorname{loglog} y + 2) + 3y\}$$

from (6.4), (6.5), and (6.6). Putting this calculation into (6.8) we have

$$\text{(6.9)} \quad N(a_0,\dots,a_k) \le x\cdot\exp\{y(\operatorname{loglog} x - \operatorname{logloglog} x - 2 \log y - \operatorname{loglog} y$$

$$+ \log(\tfrac{1}{y} \log x - \log y - \operatorname{loglog} y + 2) + 5)\}$$

$$\le x\, y^{-y} \exp\{y(-\operatorname{loglog} y + \log(\tfrac{1}{y} \log x - \log y - \operatorname{loglog} y + 2) + O(1))\}.$$

For $y = [c \log x/\operatorname{loglog} x]$ with $1/3 \le c \le 1 - 1/\operatorname{loglog} x$, we have

$$-\operatorname{loglog} y + \log(\tfrac{1}{y} \log x - \log y - \operatorname{loglog} y + 2) = \log(1-c) + O(1)$$

and for $y \ge (1 - 1/\operatorname{loglog} x)\log x/\operatorname{loglog} x$ (but $y = (1 + o(1))\log x/\operatorname{loglog} x$) we have

$$-\operatorname{loglog} y + \log(\tfrac{1}{y} \log x - \log y - \operatorname{loglog} y + 2) = -\operatorname{logloglog} x + O(1).$$

Thus from (6.9), we have

$$N(a_0,\dots,a_k) \le \begin{cases} x\, y^{-y}(1-c)^y \exp\{O(y)\} \\ \exp\{O(y)\} \end{cases}$$

depending on whether y is smaller or greater than $(1 - 1/\operatorname{loglog} x)\log x/\operatorname{loglog} x$.

Thus it only remains to show that the number of ordered partitions $a_0,\dots,a_k$ of y satisfying (6.4) is at most $e^{O(y)}$. Since

$$\frac{1}{y} \log x - \log y - \operatorname{loglog} y \ll \operatorname{loglog} x \ll \log y ,$$

it follows from (6.4) that

$$\sum_{i\ge\sqrt{y}} a_i \ll \frac{1}{\sqrt{y}} \sum_{i\ge\sqrt{y}} (i-1)a_i \ll \sqrt{y} \log y .$$

Thus the number of non-zero a_i for $i \ge \sqrt{y}$ is at most $O(\sqrt{y} \log y)$. If $p(t)$ denotes the number of numerical partitions of t, we thus have that the number of ordered choices of the a_i for $i \ge \sqrt{y}$ is at most (where $A\sqrt{y} \log y$ is an integer and A is a large constant)

$$p(A\sqrt{y} \log y)(A\sqrt{y} \log y)!\binom{k}{A\sqrt{y} \log y} = e^{O(\sqrt{y}(\log y)^2)}.$$

But the number of ordered choices of the a_i for $i < \sqrt{y}$ is at most

$$(y+1)^{\sqrt{y}+1} = e^{O(\sqrt{y} \log y)} .$$

We conclude that the number of choices for ordered partitions $a_0,\dots,a_k$ of y is at most

$$e^{O(\sqrt{y}(\log y)^2)} \leq e^{O(y)},$$

thus concluding the proof of the theorem.

Acknowledgements

I would like to thank the hospitality of the Département de Mathématiques-Informatique at the U.E.R. des Sciences de Limoges where much of this paper was written. I also wish to thank Gerald Tenenbaum who suggested the current proof of the lemma in section 4. I also wish to thank Jean-Louis Nicolas for his interest in the paper.

References

1. P. Erdös, On the integers having exactly K prime factors, Ann. of Math. (2) 49 (1948), 53-66.

2. P. Erdös and J.L. Nicolas, Sur la fonction: nombre de facteurs premiers de n, L'Enseignement Math. 27(1981), 3-27.

3. G.H. Hardy and S. Ramanujan, The normal number of prime factors of a number n, Quarterly J. Math. 48(1917), 76-92.

4. G.H. Hardy and E.M. Wright, An Introduction to the Theory of Numbers, Fourth Ed., Oxford, 1968.

5. G. Kolesnik and E.G. Straus, On the distribution of integers with a given number of prime factors, Acta Arith. 37(1980), 181-199.

6. E. Landau, Handbuch der Lehre von der Verteilung der Primzahlen, Chelsea Publishing Co., New York, 1953.

7. J.L. Nicolas, Sur la distribution des nombres entiers ayant une quantite fixee de facteurs premiers, Acta Arith., to appear.

8. J.L. Nicolas and G. Tenenbaum, private communication.

9. K.K. Norton, On the number of restricted prime factors of an integer. I, Illinois J. Math. 20(1976), 681-705; II, Acta Math. 143 (1979), 9-38; III, L'Enseignement Math. 28(1982), 31-52.

10. L.G. Sathe, On a problem of Hardy on the distribution of integers having a given number of prime factors, J. Indian Math. Soc. (N.S.) 17(1953), 63-141; 18(1954), 27-81.

11. A. Selberg, Note on a paper by L.G. Sathe, J. Indian Math. Soc. (N.S.) 18(1954), 83-87.

Some Remarks in Number Theory

S. Srinivasan

Dedicated to Professor Paul Erdös on the occasion of his Seventieth Birthday

§ 1. (cf. [7(i)]). For given integer $r \geq 1$, prime p, it is observed that

$$(2^{p^{r-1}} - 1,\ (2^{p^r} - 1)/(2^{p^{r-1}} - 1)) = 1.$$

This immediately implies the infinitude of primes $\equiv 1 (\bmod p^r)$. On the same lines, existence of primes $\equiv 1 (\bmod n)$, for all n, can be proved.

§ 2. In the direction of difference between consecutive primes the following remarks may be of some interest.

(I) It is easily seen that the roots ρ' of

$$(1) \ldots \qquad \int_0^1 \alpha\zeta(s,\alpha)d\alpha = 0 \ (\mathrm{Re}\ s > 1)$$

are precisely the complex zeros ρ, of the Riemann's zeta function, shifted by 1. (Here, as usual, $\zeta(s,\alpha) := \sum_{n=0}^{\infty} (n+\alpha)^{-s}$, $0 < \alpha$.)

(II) (cf. [5]; see also [4].) In the context of Grimm's conjecture, the following formulation may be useful. Assume that $n+1,\ldots,n+k$ are composite, and also that $k \geq \exp((\log n)^{\theta})$, with a fixed $\theta < 1$. Then is it true that

$$(2) \ldots \qquad \sum_{n < m \leq n+k} \Omega(m) \geq k(\log\log k + C)$$

holds with 'large' C? (Here, as usual, $\Omega(m)$ denotes the number of prime divisors, counted with multiplicity, of m.) This means, since

$$(3) \ldots \qquad \sum_{n < m \leq n+k} \Omega_k(m) = k(\log\log k + O(1))$$

(where $\Omega_k(m)$ is the number of primes $p \leq k$, $p|m$ counted with multiplicity), whether we have (2) with C exceeding the constant implied by $O(1)$ in (3). Clearly, (2) with arbitrarily large C would yield Piltz' conjecture - even with a $C_{\theta} \longrightarrow \infty$ as $\theta \longrightarrow 1$.

(III) We might conjecture that