Analytic Number Theory
Problem Set 3
Solutions

Problem 1

Obtain an asymptotic estimate with error term $O(x^{1/3})$ for the number of squarefull integers $\le x$, i.e., for the quantity

$$S(x) = \#\{n \le x : p|n \Rightarrow p^2|n\}.$$

Solution

The squarefull numbers are exactly the integers of the form $n = a^2b^3$, where a and b are positive integers. Moreover, by examining the exponents in the prime factorization of n it is easy to see that such a representation exists with *squarefree* b, and that under this additional restriction a and b are uniquely determined. Therefore, we have

$$\begin{aligned} S(x) &= \#\{(a,b) \in \mathbb{N}^2 : \mu^2(b) = 1, a^2b^3 \le x\} \\ &= \sum_{b \le x^{1/3}} \mu^2(b) \sum_{a \le (x/b^3)^{1/2}} 1 = \sum_{b \le x^{1/3}} \mu^2(b) \left(x^{1/2} b^{-3/2} + O(1) \right) \\ &= x^{1/2} A + O(x^{1/2} R(x^{1/3})) + O(x^{1/3}), \end{aligned}$$

where

$$A = \sum_{b=1}^{\infty} \frac{\mu^2(b)}{b^{3/2}}, \quad R(y) = \sum_{b>y} \frac{\mu^2(b)}{b^{3/2}}.$$

Since $R(y) \le \sum_{n>y} n^{-3/2} \ll y^{-1/2}$ for $y \ge 2$, the error terms above are of order $O(x^{1/3})$, and we obtain $S(x) = A\sqrt{x} + O(x^{1/3})$.

Remark. Alternatively, one could use a convolution argument as follows: Let f denote the characteristic function of squarefull integers. Then f is multiplicative with $f(p) = 0$ and $f(p^m) = 1$ for $m \ge 2$. Thus, a natural approximation to f is the characteristic function of the squares, s, which is multiplicative with $s(p^m) = 1$ if m is even, and 0 otherwise. Expressing f as a convolution $f = s * g$, we find (using the multiplicativity of f and s) that g is multiplicative with $g(p^m) = 1$ if $m = 3$ and 0 else, so that $g(n) = 1$ if $n = m^3$ with m squarefree, and $g(n) = 0$ else. Using this and the fact that $\sum_{n \le y} s(n) = \sum_{m^2 \le y} s(n) = \sqrt{y} + O(1)$, we obtain

$$\sum_{n \le x} f(n) = \sum_{d \le x} g(d) \sum_{n \le x/d} s(n) = \sum_{d \le x} g(d) \sqrt{x/d} + O\left(\sum_{n \le x} |g(d)| \right),$$

which as above reduces to $A\sqrt{x} + O(x^{1/3})$.

Problem 2

Given an arithmetic function $a(n)$, $n = 1, 2, \dots$, and a real number $\alpha > -1$ define a mean value $M_\alpha(a)$ by

$$M_\alpha(a) = \lim_{x\to\infty} \frac{1+\alpha}{x^{1+\alpha}} \sum_{n\le x} n^\alpha a(n),$$

provided the limit exist. (In particular, $M_0(a) = M(a)$ is the usual asymptotic mean value of a.) Prove, using a rigorous $\epsilon - x_0$ argument, that the mean value $M_\alpha(a)$ exists if and only if the ordinary mean value $M(a) = M_0(a)$ exists. (As a consequence, if one of the mean values $M_\alpha(a)$, $\alpha > -1$, exists, then all of these mean values exist.)

Solution

Fix $\alpha > -1$ and an arithmetic function $a(n)$ and write M_α and M for $M_\alpha(a)$ and $M(a)$, respectively. Define corresponding partial sums

$$T_\alpha(x) = \sum_{n\le x} n^\alpha a(n), \quad S(x) = T_0(x) = \sum_{n\le x} a(n),$$

so that

$$M_\alpha = \lim_{x\to\infty} \frac{(1+\alpha)T_\alpha(x)}{x^{1+\alpha}}, \quad M = M_0 = \lim_{x\to\infty} \frac{S(x)}{x},$$

provided the limits exist.

Assume first that the ordinary mean value, M, exists. Partial summation gives, for any $x \ge 1$,

$$T_\alpha(x) = x^\alpha S(x) - \int_1^x \alpha t^{\alpha-1} S(t)dt.$$

Dividing both sides by $x^{1+\alpha}$ and setting

$$t_\alpha(x) = \frac{T_\alpha(x)}{x^{1+\alpha}}, \quad s(x) = t_0(x) = \frac{S(x)}{x},$$

we get

$$t_\alpha(x) = s(x) - \alpha x^{-\alpha-1} \int_1^x t^\alpha s(t)dt = s(x) - \alpha x^{-\alpha-1} I(x),$$

say.

Now, by assumption, the limit $M = \lim_{x\to\infty} s(x)$ exists, and we need to show that the limit $M_\alpha = (1+\alpha)\lim_{x\to\infty} t_\alpha(x)$ exists as well. By the above identity, this (and the equality of the two limits M_α and M) will follow provided we can show that

$$\lim_{x\to\infty} x^{-1-\alpha} I(x) = \frac{M}{1+\alpha}. \tag{1}$$

Let $\epsilon > 0$ be given and choose $x_0 = x_0(\epsilon) \geq 1$ such that

$$|s(x) - M| \leq \epsilon \quad (x \geq x_0).$$

Using the identity

$$\frac{x^{1+\alpha}}{1+\alpha} = \int_0^x t^\alpha dt$$

and defining, for convenience, $s(t) = 0$ for $0 \leq t < 1$, we have, for $x \geq x_0$,

$$I(x) = \int_0^x t^\alpha s(t) dt = \int_0^x t^\alpha (s(t) - M) dt + \frac{x^{1+\alpha} M}{1+\alpha}.$$

It follows that

$$\begin{aligned} \left| I(x) - \frac{x^{1+\alpha} M}{1+\alpha} \right| &\leq \int_0^{x_0} t^\alpha (|s(t)| + M) dt + \epsilon \int_{x_0}^x t^\alpha dt \\ &\leq C_0 + \epsilon \frac{x^{1+\alpha}}{1+\alpha}, \end{aligned}$$

where $C_0 = C_0(\epsilon)$ is the value of the first integral above and is a constant independent of x. Hence

$$\limsup_{x\to\infty} \left| x^{-1-\alpha} I(x) - \frac{M}{1+\alpha} \right| \leq \epsilon.$$

Since ϵ can be chosen arbitrarily small, (1) follows.

Thus, the existence of M implies that of M_α.

Conversely, assume that M_α exists. By partial summation we have

$$S(x) = x^{-\alpha} T_\alpha(x) - \int_1^x (-\alpha) t^{-\alpha-1} T_\alpha(t) dt.$$

Introducing again the normalized quantities $s(x)$ and $t_\alpha(x)$, this becomes

$$s(x) = t_\alpha(x) + \alpha x^{-1} \int_1^x t_\alpha(t) dt = t_\alpha(x) + \alpha x^{-1} J(x),$$

say. The existence of M_α implies that $\lim_{x\to\infty} t_\alpha(x) = M_\alpha/(1+\alpha)$. Thus, in order to conclude that the limit $M = \lim_{x\to\infty} s(x)$ exists as well and is equal to M_α, it suffices to show that the integral $J(x)$ satisfies

$$\lim_{x\to\infty} x^{-1} J(x) = \frac{M_\alpha}{1+\alpha}. \tag{2}$$

This can be proved in exactly the same way as (1) using an $\epsilon - x_0(\epsilon)$ argument.

Remark. The hypothesis $\alpha > -1$ is crucial in this argument since we need that $x^{1+\alpha}$ tends to infinity as $x \to \infty$. When $\alpha \le -1$ this is no longer the case. In fact, when $\alpha = -1$, the above definition of M_α makes no longer sense since it involves a factor $1+\alpha$, but the logarithmic mean value $\lim_{x\to\infty}(1/\log x)\sum_{n\le x} a(n)n^{-1}$ can be viewed as the limiting case of M_α as $\alpha \to -1$, and we know that the existence of the logarithmic mean value is not equivalent to that of the ordinary mean value.

Problem 3

Let f be an arithmetic function having a non-zero mean value $M(f) = A$, and let α be a fixed real number. Obtain an asymptotic formula for the sums $\sum_{n \le x} f(n) n^{i\alpha}$.

Solution

Let $S(x) = \sum_{n \le x} f(n)$, so that $\lim_{x\to\infty} S(x)/x = A$. By partial summation we have

$$\sum_{n \le x} f(n) n^{i\alpha} = x^{i\alpha} S(x) - i\alpha \int_1^x t^{i\alpha - 1} S(t) dt.$$

Since $\lim_{t\to\infty} S(t)/t = A$, given $\epsilon > 0$ one can find a $t_0 = t_0(\epsilon) \ge 1$ such that $|S(t) - At| \le \epsilon t$ for $t \ge t_0$. Thus, if $t \ge t_0$ then

$$|t^{i\alpha} S(t) - At^{1+i\alpha}| \le \epsilon t$$

and therefore

$$\left| \int_1^x t^{i\alpha-1} S(t) dt - A \int_1^x t^{i\alpha} dt \right| \le \int_1^{t_0} (|S(t)| + |A|) dt + \epsilon \int_{t_0}^x dt \le K_0(\epsilon) + \epsilon x$$

with a suitable constant $K_0(\epsilon)$. Since

$$x^{1+i\alpha} - i\alpha \int_1^x t^{i\alpha} dt = x^{1+i\alpha} \left(1 - \frac{i\alpha}{1+i\alpha} \right) + \frac{i\alpha}{1+i\alpha} = \frac{1}{1+i\alpha} x^{1+i\alpha} + O(1),$$

it follows that

$$\limsup_{x\to\infty} \frac{1}{x} \left| \sum_{n \le x} f(n) n^{i\alpha} - \frac{1}{1+i\alpha} A x^{1+i\alpha} \right| \le \epsilon.$$

Since ϵ was arbitray, this proves that $\sum_{n \le x} f(n) n^{i\alpha} \sim \frac{A}{1+i\alpha} x^{1+i\alpha}$ as $x \to \infty$. (In particular, this shows that the mean value of $f(n) n^{i\alpha}$ does not exist for $\alpha \ne 0$.)

Problem 4

Obtain an estimate, similar to Dirichlet's estimate for $\sum_{n\le x} d(n)$, for the sum $\sum_{n\le x} 2^{\omega(n)}$. (You can leave constants appearing in this estimate unspecified.)

Solution

Let $S(x) = \sum_{n\le x} 2^{\omega(n)}$. We will show that

$$S(x) = Ax\log x + Bx + O\left(\sqrt{x}\log x\right) \quad (x \ge 2), \tag{1}$$

where A and B are constants. Except for the extra log factor in the error term, this estimate is of the same quality as Dirichlet's estiamte for the sums of the divisor function, $D(x) = \sum_{n\le x} d(n)$, which states

$$D(x) = x\log x + (2\gamma - 1)x + O(\sqrt{x}) \quad (x \ge 1). \tag{2}$$

We will deduce (1) from (2) by applying the convolution method with $f(n) = 2^{\omega(n)}$ and $f_0(n) = d(n)$ as the approximating function. Accordingly, we define g by $f = f_0 * g$. Since f and f_0 are multiplicative, g is multiplicative as well, and a simple calculation shows that $g(p^m) = -1$ if $m = 2$ and $g(p^m) = 0$ otherwise. Thus,

$$g(n) = \begin{cases} \mu(m) & \text{if } n = m^2, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

For $x \ge 1$ we have

$$\begin{aligned} S(x) &= \sum_{n\le x} f(n) = \sum_{n\le x}\sum_{d|n} g(d) f_0(n/d) \\ &= \sum_{d\le x} g(d) \sum_{n\le x/d} f_0(n) = \sum_{d\le x} g(d) D(x/d). \end{aligned}$$

Substituting the estimate (2) for $D(x/d)$, we get

$$\begin{aligned} S(x) &= (x\log x)\sum_{d\le x}\frac{g(d)}{d} - x\sum_{d\le x}\frac{g(d)\log d}{d} \\ &\quad + (2\gamma-1)x\sum_{d\le x}\frac{g(d)}{d} + O\left(\sqrt{x}\sum_{d\le x}\frac{|g(d)|}{\sqrt{d}}\right). \end{aligned} \tag{4}$$

The sums involving $g(d)$ can be evaluated as follows, using the formula (3):

$$\sum_{d\le x}\frac{g(d)}{d}=\sum_{m\le\sqrt{x}}\frac{\mu(m)}{m^2}=\sum_{m=1}^{\infty}\frac{\mu(m)}{m^2}-\sum_{m>\sqrt{x}}\frac{\mu(m)}{m^2} = \frac{6}{\pi^2}+O\left(\frac{1}{\sqrt{x}}\right)\quad (x\ge 1), \tag{5}$$

since $\sum_{n\ge y} n^{-2}=O(1/y)$ for $y\ge 1$; and

$$\sum_{d\le x}\frac{g(d)\log d}{d}=\sum_{m\le\sqrt{x}}\frac{\mu(m)\log m^2}{m^2} = 2\sum_{m=1}^{\infty}\frac{\mu(m)\log m}{m^2}-\sum_{m>\sqrt{x}}\frac{\mu(m)\log m}{m^2} = C+O\left(\frac{\log x}{\sqrt{x}}\right)\quad (x\ge 2), \tag{6}$$

where C is a constant.

$$\sum_{n>y}\frac{\log n}{n^2}\ll\sum_{y<n\le y^2}\frac{\log y^2}{n^2}+\sum_{n>y^2}\frac{n^{1/2}}{n^2} = O\left(\frac{\log y}{y}\right)+O\left(\frac{1}{(y^2)^{1/2}}\right)=O\left(\frac{\log y}{y}\right)\quad (y\ge 2).$$

Substituting (5) and (6) into (3) we get the asserted estimate (1) with $A=6/\pi^2$ and $B=-C+(2\gamma-1)\pi^2/6$.

Remark. Using other convolution formulas such as $2^\omega=1*\mu^2$ would only give an error term $O(x)$ (unless one would use the Dirichlet hyperbola method to estimate the double sums that arise in the process). This is because the function 1, while being well behaved, is a very poor approximation to 2^ω, in contrast to the divisor function. (Recall that what counts in the case of multiplicative functions is their values at primes.)

Problem 5

Using the Dirichlet hyperbola method (or some other method), obtain an estimate for the sum $\sum_{n\le x} d(n)/n$ with an error term $O((\log x)/\sqrt{x})$. (Note the error term. Exercise 2 in Section 3 of Apostol asks for an estimate of the same sum, but only with error $O(1)$.)

Solution

We will show that

$$\text{(1)}\qquad \sum_{n\le x}\frac{d(n)}{n}=\frac{1}{2}\log^2 x+2\gamma\log x+C+O\left(\frac{\log x}{\sqrt{x}}\right)\quad (x\ge 2),$$

where C is a constant. We have

$$\sum_{n\le x}\frac{d(n)}{n}=\sum_{n\le x}\frac{1}{n}\sum_{\substack{k,l\le x\\ kl=n}}1=\sum_{k\le x}\frac{1}{k}\sum_{\substack{l\le x\\ kl\le x}}\frac{1}{l}.$$

We now split up the last sum according to the Dirichlet hyperbola method into

$$\begin{aligned}\sum_{k\le\sqrt{x}}\frac{1}{k}\sum_{l\le x/k}\frac{1}{l}+\sum_{l\le\sqrt{x}}\frac{1}{l}\sum_{k\le x/l}\frac{1}{k}-\sum_{k\le\sqrt{x}}\frac{1}{k}\sum_{l\le\sqrt{x}}\frac{1}{l}\\ =\Sigma_1+\Sigma_2-(\Sigma_3)^2,\end{aligned}$$

say. Applying the estimates for $\sum_{n\le y}1/n$ and $\sum_{n\le y}(\log n)/n$ from class and from the preceding problem, we get, for $x\ge 2$,

$$\begin{aligned}(\Sigma_3)^2&=\left(\log\sqrt{x}+\gamma+O\left(\frac{1}{\sqrt{x}}\right)\right)^2\\ &=\frac{1}{4}\log^2 x+\gamma\log x+\gamma^2+O\left(\frac{\log x}{\sqrt{x}}\right),\end{aligned}$$

and

$$\begin{aligned}
\Sigma_1 = \Sigma_2 &= \sum_{k\le\sqrt{x}} \frac{1}{k}\left(\log\frac{x}{k} + \gamma + O\left(\frac{k}{x}\right)\right) \\
&= (\log x + \gamma)\sum_{k\le\sqrt{x}}\frac{1}{k} - \sum_{k\le\sqrt{x}}\frac{\log k}{k} + O\left(\sum_{k\le\sqrt{x}}\frac{1}{x}\right) \\
&= (\log x + \gamma)\left(\log\sqrt{x} + \gamma + O\left(\frac{1}{\sqrt{x}}\right)\right) \\
&\quad - \frac{1}{2}\left(\log\sqrt{x}\right)^2 - A + O\left(\frac{\log\sqrt{x}}{\sqrt{x}}\right) + O\left(\frac{1}{\sqrt{x}}\right) \\
&= \frac{3}{8}\log^2 x + \frac{3}{2}\gamma\log x + B + O\left(\frac{\log x}{\sqrt{x}}\right),
\end{aligned}$$

where A is the constant from the preceding problem and $B = \gamma^2 - A$. Combining these estimates gives (1) with constant $C = 2B - \gamma^2$.

Remark. An alternative approach is to apply partial summation to "remove" the factors $1/n$ in the sum $\sum_{n\le x} d(n)/n$ and use Dirichlet's estimate for $\sum_{n\le x} d(n)$. In general, with partial summation arguments, the outcome is usually hard to predict, and one just has to try it and see how far one gets; in some cases, it works very well, while in others it fails completely. It turns out that partial summation does work quite well in this case, involves less calculation than the above argument (though, of course, the bulk of the argument is in the proof of Dirichlet's estimate for $\sum_{n\le x} d(n)$, which one simply quotes), and in fact gives a slightly stronger result, involving an error term of size $O(1/\sqrt{x})$. Try it!

Problem 6

Let $q_1 = 1, q_2 = 2, q_3 = 3, q_4 = 5\ldots$ denote the sequence of squarefree numbers.

(i) Obtain an asymptotic estimate with error term $O(\sqrt{n})$ for q_n.

(ii) Show that there are arbitrarily large gaps in the sequence $\{q_n\}$, i.e., $\limsup_{n\to\infty}(q_{n+1} - q_n) = \infty$. (Hint: Chinese Remainder Theorem.)

(iii) (Bonus problem) Prove the stronger bound

$$\limsup_{n\to\infty} \frac{q_{n+1} - q_n}{\log n/\log\log n} \ge \frac{1}{2}.$$

(iv) (Superbonus problem—deadline 10/31/05) Prove that (iii) holds with 1/2 replaced by the constant $\pi^2/12$, i.e., that the limsup above is at least $\pi^2/12$.

(According to the famous number theorist Paul Erdös, this result can be proved by "a simple argument using the Chinese Remainder Theorem and the Prime Number Theorem", so all you need to do to earn the superbonus points is to supply this "simple argument" ....)

Solution

(i) Using the estimate $\sum_{n\le x}\mu^2(n) = (6/\pi^2)x + O(\sqrt{x})$, we get

$$n = \sum_{m\le q_n} \mu^2(m) = \frac{6}{\pi^2} q_n + O(\sqrt{q_n}) \quad (n\to\infty),$$

From this we deduce first the crude bound $n \ll q_n$, which enables us to replace the error term $O(\sqrt{q_n})$ by $O(\sqrt{n})$. Multiplying the through with the constant $\pi^2/6$, we then get the desired estimate for q_n:

$$q_n = (\pi^2/6)n + O(\sqrt{n}).$$

(ii) Let p_i denote the ith prime. To obtain an interval of the type $(N, N+k]$ that is free of squarefree numbers, we consider the system of congruences

$$N \equiv -i \bmod p_i^2 \quad (i = 1, \ldots, k). \tag{1}$$

If N is a solution to (1), then for $i = 1, \ldots, k$, $N + i$ is divisible by p_i^2, and the interval $(N, N+k]$ is therefore free of squarefree numbers. Thus, to show that there are arbitrarily large gaps between squarefree numbers, it suffices to show that (1) has a positive solution N for any given positive integer k. Since the moduli in (1) are pairwise relatively prime, this is an immediate consequence of the Chinese Remainder Theorem.

(iii) To obtain an explicit lower bound, note that the system (1) has a solution $N = N_k$ with $1 \le N_k \le \prod_{i=1}^k p_i^2 = P_k^2$, where P_k is the product of the first k primes. Now, by the PNT, $\log P_k = \theta(p_k) \sim p_k \sim k \log k$ as $k \to \infty$, Hence, as $k \to \infty$, $\log N_k \le 2(1 + o(1))k \log k$, and therefore $k \ge (1+o(1))(1/2) \log N_k / \log\log N_k$. Now let n_k denote the largest index n such that $q_n \le N_k$. Since by construction the interval $(N_k, N_k + k]$ contains no squarefree numbers, we have $q_{n_k+1} > N_k + k$, and hence

$$q_{n_k+1} - q_{n_k} \ge k \ge (1 + o(1)) \frac{\log N_k}{2 \log\log N_k} \tag{2}$$

as $k \to \infty$. Now, since by part (i), $q_{n_k} = (1 + o(1))cn_k$, with $c = \pi^2/6$, we have $cn_k(1 + o(1)) \le N_k < c(1 + o(1))(n_k + 1)$ and hence

(3)
$$\log N_k = \log n_k + \log c + o(1) = (1 + o(1)) \log n_k$$
(4)
$$\log\log N_k = \log\left((1 + o(1)) \log n_k\right) = \log\log n_k + o(1) = (1 + o(1)) \log\log n_k.$$

Combining (1)–(3) yields

$$\limsup_{k\to\infty} \frac{q_{n_k+1} - q_{n_k}}{2(\log n_k)/(\log\log n_k)} \ge 1,$$

which proves the desired lower bound.

(iv) Deferred.

Problem 7

Show that, $\phi(n) \geq n/4$ holds for at least 1/3 of all positive integers n (in the sense that if A is the set of such n, then $\liminf_{x\to\infty}(1/x)\#\{n \leq x : n \in A\} \geq 1/3$). (Hint: use the fact that (1) $\sum_{n\leq x}\phi(n) \sim (3/\pi^2)x^2$ (which was proved in class) or (2) $\sum_{n\leq x}\phi(n)/n \sim (6/\pi^2)x$ (an easy consequence of Wintner's theorem, or of (1), by partial summation).)

Solution

Let $A = \{n \in \mathbb{N} : \phi(n) \geq n/4\}$, and let $B = \mathbb{N}\backslash A$. We want to show that $\underline{d}(A) \geq 1/3$, where

$$\underline{d}(A) = \liminf_{x\to\infty} \frac{1}{x}\#\{n \leq x : n \in A\}$$

is the lower asymptotic density of A. Since $\phi(n)/n \leq 1$ for any n and $\phi(n)/n \leq 1/4$ for $n \in B$, we have, for any $x \geq 1$,

$$\frac{1}{x}\sum_{n\leq x}\frac{\phi(n)}{n} \leq \frac{1}{x}\sum_{n\leq x, n\in B}\frac{1}{4} + \frac{1}{x}\sum_{n\leq x, n\in A} 1 \leq \frac{1}{4} + \frac{1}{x}\sum_{n\leq x, n\in A} 1$$

Letting $x \to \infty$, the first term on the left converges to $6/\pi^2 = 0.6079\ldots$. Thus the liminf of the second term, i.e., $\underline{d},(A)$ is $\geq 6/\pi^2 - 1/4 > 1/3$, as claimed.

Problem 8

Euler's proof of the infinitude of primes shows that $(*)\ \sum_{p\le x} 1/p \ge \log\log x - C$, for some constant C and all sufficiently large x. This is a remarkably good lower bound for the sum of reciprocals of primes (it is off by only a term $O(1)$), so it is of some interest to see what this bound implies for $\pi(x)$. The answer is, surprisingly little, as the following problems show.

(i) Deduce from $(*)$, *without using any other information about the primes*, that there exists $\delta > 0$ such that $\pi(x) > \delta \log x$ for all sufficiently large x. In other words, show that if A is any sequence of positive integers satisfying

$$\sum_{a\le x, a\in A} \frac{1}{a} \ge \log\log x - C \tag{1}$$

for some constant C and all sufficiently large x, then there exists a constant $\delta > 0$ such that the counting function $A(x) = \#\{a \in A, a \le x\}$ satisfies

$$A(x) \ge \delta \log x \tag{2}$$

for all sufficiently large x.

(ii) (Superbonus problem—deadline 10/31/05) Show that this result is nearly best possible, in the sense that it becomes false if the function $\log x$ on the right-hand side of (2) is replaced by a power $(\log x)^\alpha$ with an exponent α greater than 1. In other words, given $\epsilon > 0$, construct a sequence A of positive integers, satisfying (1) above, but for which the counting function $A(x)$ $A(x) = \#\{a \in A, a \le x\}$ satisfies

$$\liminf_{x\to\infty} A(x) \log x)^{-1-\epsilon} \le 1. \tag{3}$$

Solution

(i) Let $S(x) = \sum_{p\le x} 1/p$. By assumption, there exist constants $c > 0$ and $x_0 \ge 1$ such that

$$S(x) \ge \log\log x - C \quad (x \ge x_0). \tag{1}$$

On the other hand, we have the upper bound

$$S(x) \le \sum_{n \le x} \frac{1}{n} \le 1 + \int_1^x \frac{1}{t} dt = 1 + \log x \quad (x \ge 1). \tag{2}$$

(Of course, much stronger bounds can be obtained if we use some information about the distribution of primes (e.g., Chebyshev's bound $\pi(x) \ll x/\log x$), but we are not allowed to use such information.)

Now let $x \ge x_0$ be given, and let y be a number with $1 \le y \le x$. Our lower bound for $\pi(x)$ is based on the inequality

$$\pi(x) \ge y \sum_{y<p\le x} \frac{1}{p} = y(S(x) - S(y)),$$

with a suitable choice of y. Estimating $S(x)$ from below by (1), and $S(y)$ from above by (2), we obtain, for $x \ge x_0$ and $1 \le y \le x$,

$$\pi(x) \ge y(\log\log x - C - \log y - 1).$$

Choosing y to be of the form $y = \delta \log x$ we have $\log y = \log\log x - \log(1/\delta)$, so our lower bound becomes

$$\delta \log x(\log(1/\delta) - C - 1),$$

which equals $\delta \log x$ if we take $\delta = e^{-C-2}$. Hence, with this choice of δ, we have $\pi(x) \ge \delta \log x$ for all sufficiently large x, as claimed.

(ii) Deferred.