18.104, Solutions to fifth problem set

Problem 1 (page 294, #4). This asked you to show that

$$\sum_{p \le N \text{ prime}} \frac{1}{p} > \log\log N - \frac{1}{2}. \tag{$*$}$$

The natural starting point is the inequality proved on page 290:

$$\sum_{p \le N \text{ prime}} \log \frac{1}{1 - \frac{1}{p}} > \log\log N. \tag{1}$$

The hint suggested using the power series expansion

$$\log(1+x) = x - x^2/2 + x^3/3 - x^4/4 \cdots, \tag{2}$$

which is valid for $-1 < x \le 1$. Many of you rewrote the argument of log in (1) as $1 + \frac{1}{p-1}$, and then applied the expansion (2) with $x = \frac{1}{p-1}$. This is natural, but it's a very difficult path. The alternating signs in the expansion are terrible to keep track of, to begin with.

What works better is to write

$$\log \frac{1}{1 - \frac{1}{p}} = -\log(1 - \frac{1}{p}),$$

and then to apply (2) with $x = -\frac{1}{p}$. Then all the terms in the expansion have the same sign, and we get

$$\sum_{p \le N \text{ prime}} \log \frac{1}{1 - \frac{1}{p}} = \sum_{p \le N \text{ prime}} (\frac{1}{p} + \frac{1}{2p^2} + \frac{1}{3p^3} + \cdots). \tag{3}$$

Combining (3) and (1), we get the inequality

$$\sum_{p \le N \text{ prime}} \frac{1}{p} > \log\log N - \sum_{p \le N \text{ prime}} \sum_{k=2}^{\infty} \frac{1}{kp^k}. \tag{4}$$

In order to prove ($*$), it is therefore enough to prove the inequality

$$\sum_{p \text{ prime}} \sum_{k=2}^{\infty} \frac{1}{kp^k} \le \frac{1}{2} \tag{$**$}$$

This is actually tough, as far as I can see. I'll explain a straightforward approach that doesn't work, and then see how to sharpen it to succeed.

First, replace the outer sum over primes by a sum over all integers $p \ge 2$. Then interchange the order of summation. This gives

$$\sum_{p \text{ prime}} \sum_{k=2}^{\infty} \frac{1}{kp^k} < \sum_{k=2}^{\infty} \sum_{p=2}^{\infty} \frac{1}{kp^k}. \tag{5}$$

Now the inside sum (by making a picture in the usual way) is smaller than

$$\int_1^{\infty} \frac{dx}{kx^k}.$$

This integral is equal to $\frac{1}{k(k-1)}$. Plugging back into (5) gives

$$\sum_{p \text{ prime}} \sum_{k=2}^{\infty} \frac{1}{kp^k} < \sum_{k=2}^{\infty} \frac{1}{k(k-1)} = 1. \tag{6}$$

This is close to (∗∗) but not quite there; it gives the inequality in the problem with 1 instead of 1/2, which is not good enough.

All I could find to get 1/2 was something pretty ugly. Fix a prime number $m+1$. We'll split the sum in (∗∗) into two parts: primes smaller than m, and primes equal to or greater than $m+1$. For the second part, we'll replace the sum by a sum over all integers starting with $m+1$, and proceed exactly as I did in what didn't work above. So here we go...

$$\sum_{p \text{ prime}} \sum_{k=2}^{\infty} \frac{1}{kp^k} < \sum_{p \le m \text{ prime}} \sum_{k=2}^{\infty} \frac{1}{kp^k} + \sum_{p=m+1}^{\infty} \sum_{k=2}^{\infty} \frac{1}{kp^k}. \tag{7}$$

Call the two terms on the right A and B. For B, we proceed more or less as before:

$$\begin{aligned} B = \sum_{k=2}^{\infty} \sum_{p=m+1}^{\infty} \frac{1}{kp^k} &< \sum_{k=2}^{\infty} \int_m^{\infty} \frac{dx}{kx^k} \\ &= \sum_{k=2}^{\infty} \frac{1}{k(k-1)m^{k-1}} \\ &< \sum_{k=2}^{\infty} \frac{1}{k(k-1)m} = 1/m. \end{aligned} \tag{8}$$

For A, we undo the series expansion (3) that we began with:

$$\begin{aligned} A &= \sum_{p \le m \text{ prime}} \sum_{k=2}^{\infty} \frac{1}{kp^k} \\ &= \sum_{p \le m \text{ prime}} \left(\log\left(\frac{1}{1-\frac{1}{p}} \right) - \frac{1}{p} \right) \end{aligned}. \tag{9}$$

I'm going to use $m = 6$, so that the primes 2, 3, and 5 appear in (A), and B starts at $p = 7$. Using a calculator reveals that

$$\begin{aligned} \log\left(\frac{1}{1-\frac{1}{2}}\right) - \frac{1}{2} &< .1932, \\ \log\left(\frac{1}{1-\frac{1}{3}}\right) - \frac{1}{3} &< .0722, \\ \log\left(\frac{1}{1-\frac{1}{5}}\right) - \frac{1}{5} &< .0232. \end{aligned}$$

Adding and plugging into (8) gives $A < .2886$. By (7), $B < .1667$. Therefore $A + B < .4553 < 1/2$, and this is (**).

Problem 2 (page 295, #6). Most of you had at least the main idea here. If the density exists, then it can be computed as

$$\lim_{k\to\infty} A(n_k)/n_k$$

for any sequence n_k of positive integers tending to infinity. (This is because any subsequence of a convergent sequence converges to the same limit.) Now look at two special sequences of integers:

$$n_k = 10^k - 1, \qquad m_k = 2 \cdot 10^{k-1}.$$

These two numbers are preceded by exactly the same integers with first digit one: the 10^{k-1} k-digit ones, the 10^{k-2} with $k-1$ digits, and so on down to the 10^0 one-digit number 1. Therefore

$$A(n_k) = A(m_k) = \sum_{i=0}^{k-1} 10^i = (10^k - 1)/(10-1) = (10^k - 1)/9.$$

We now calculate

$$\lim_{k\to\infty} A(n_k)/n_k = \lim_{k\to\infty}(10^k - 1)/(9\cdot(10^k - 1)) = 1/9.$$

Similarly

$$\lim_{k\to\infty} A(m_k)/m_k = \lim_{k\to\infty}\frac{10^k - 1}{9\cdot(2\cdot 10^{k-1})} = \lim_{k\to\infty}(10/18 - 1/(18\cdot 10^{k-1}) = 10/18 = 5/9.$$

Because these two subsequences have different limits, the density of A is not defined.

Problem 3 (page 295, #8). Suppose that $A = \{a_k | k = 1, 2, \ldots\}$ is a collection of positive integers such that $\sum_{k=1}^{\infty}\frac{1}{a_k}$ converges. We want to show that A has density zero. This means that for every $\epsilon > 0$, we can find an $N_d(\epsilon)$ so that if $x \geq N_d(\epsilon)$, then $A(x)/x < \epsilon$. (The subscript d stands for "density.") Here $A(x)$ is the number of k such that $a_k \leq x$.

So suppose that A does *not* have density zero. This means that for some $\epsilon > 0$, we can find arbitrarily large numbers x so that $A(x) > \epsilon x$. Replacing the variable x by $y = x/\epsilon$, we can find arbitrarily large numbers y so that

$$A(y/\epsilon) > y.$$

In particular, this means that we can choose a sequence of positive even integers $2n_1, 2n_2, \ldots$ with the properties

$$A(2n_m/\epsilon) \geq 2n_m, \tag{1}$$

and

$$n_m \geq 2n_{m-1} \qquad (m = 2, 3, 4\ldots). \tag{2}$$

It follows easily by induction that

$$n_m \geq n_1 + n_2 + \cdots + n_{m-1} \qquad (m = 2, 3, 4\ldots). \tag{3}$$

In light of the definition of $A(x)$, condition (1) means that there are at least $2n_m$ integers k with the property that $a_k \leq 2n_m/\epsilon$, or equivalently

$$\frac{1}{a_k} \geq \epsilon/2n_m. \tag{4}$$

Now we choose a subsequence of a_k in the following way. First choose n_1 distinct integers $k_1^1, \ldots, k_{n_1}^1$ so that

$$\frac{1}{a_{k_i^1}} \geq \epsilon/2n_1. \tag{5}$$

This is possible by (4). Next, choose n_2 integers $k_1^2, \ldots, k_{n_2}^2$, distinct from each other and from the first n_1, with the property that

$$\frac{1}{a_{k_i^2}} \geq \epsilon/2n_2. \tag{6}$$

According to (4) there are $2n_2$ such k to choose from; since we are excluding less than $2n_1$, we can do this. Continue in this way. At the mth step, choose n_m integers $k_1^m, \ldots, k_{n_m}^m$, distinct from each other and from the first $n_1 + \cdots + n_{m-1}$, with the property that

$$\frac{1}{a_{k_i^m}} \geq \epsilon/2n_m. \tag{7}$$

According to (4) there are at least $2n_m$ possible k to choose from; since we are excluding $n_1 + \cdots + n_{m-1}$, (3) guarantees that there are at least n_m left.

Now we examine the sum

$$\sum_{m=1}^{\infty}\sum_{i=1}^{n_m}\frac{1}{a_{k_i^m}}.$$

The sum runs over a subset of the (positive) terms in the sum we are interested in; so if this sum diverges, the larger one must as well. According to (7), the mth sum here is at least $\epsilon/2$; so the outer sum is greater than or equal to

$$\sum_{m=1}^{\infty} \epsilon/2,$$

which diverges.

Most of you remembered that the set of primes has density zero, but $\sum \frac{1}{p}$ diverges; so the converse is not true.

Problem 3 (page 296, #12). A good way to approach this is to look at the sum

$$\sum_{i=S(N)+1}^{S(N+1)} \frac{1}{i} = H(S(N+1) - H(S(N)). \tag{1}$$

In light of the definition of $S(N)$, we have

$$H(S(N)) = N + \epsilon(N), \qquad 0 \leq \epsilon(N)) < 1/S(N). \tag{2}$$

Comparing these two formulas gives

$$\sum_{i=S(N)+1}^{S(N+1)} \frac{1}{i} = 1 + \epsilon(N+1) - \epsilon(N) = 1 + \delta(N), \qquad |\delta(N) < 1/S(N)|. \tag{3}$$

On the other hand, we can compare the sum with integrals. This gives

$$\int_{S(N)}^{S(N+1)} \frac{dx}{x} > \sum_{i=S(N)+1}^{S(N+1)} \frac{1}{i} > \int_{S(N)}^{S(N+1)} \frac{dx}{x+1}. \tag{4}$$

Computing the integrals gives

$$\log\left(\frac{S(N+1)}{S(N)}\right) > \sum_{i=S(N)+1}^{S(N+1)} \frac{1}{i} > \log\left(\frac{S(N+1)+1}{S(N)+1}\right). \tag{5}$$

The last argument for log can be written as

$$\frac{S(N+1)}{S(N)} \frac{1+\frac{1}{S(N+1)}}{1+\frac{1}{S(N)}};$$

so the properties of log give

$$\log\left(\frac{S(N+1)}{S(N)}\right) > \sum_{i=S(N)+1}^{S(N+1)} \frac{1}{i} > \log\left(\frac{S(N+1)}{S(N)}\right) + \log\left(1+\frac{1}{S(N+1)}\right) - \log\left(1+\frac{1}{S(N)}\right). \tag{6}$$

For positive t, it's easy to check that $\log(1+t) < t$; so we can write the last inequalities as

$$\sum_{i=S(N)+1}^{S(N+1)} \frac{1}{i} = \log\left(\frac{S(N+1)}{S(N)}\right) - e(N), \qquad 0 < e(N) < 1/S(N). \tag{8}$$

Comparing (8) and (2) gives

$$\log\left(\frac{S(N+1)}{S(N)}\right) = 1 + \delta(N) + e(N) = 1 + f(N), \qquad |f(N)| < 2/S(N).$$

Since $S(N) \to \infty$, it follows at once that

$$\lim_{N\to\infty} \log\left(\frac{S(N+1)}{S(N)}\right) = 1.$$

Exponentiating gives the result of the problem. (Notice that we also get a pretty good bound on the error: at the Nth term it's at most $6/S(N)$. (The extra factor of three comes from the derivative of the exponential function near 1; why is that relevant?) In the table on page 296, the actual error is sometimes more than $1/S(N)$, but seems to be smaller than $2/S(N)$.)