6. Sums involving primes

Lemma 6.1. If $t > -1$ is a non-zero real number then

$$\frac{t}{1+t} < \log(1 + t) < t.$$  

In particular if $t \in (0, 1/2]$ then

$$\log(1 - t) > -2t.$$  

Proof. Let

$$f: (-1, \infty) \rightarrow \mathbb{R}$$

be the function

$$f(u) = \log(1 + u).$$

Then $f$ is continuous and differentiable, so that by the mean value theorem we may find

$$s \in \begin{cases} [0, t] & \text{if } t > 0 \\ [t, 0] & \text{if } t < 0 \end{cases}$$

such that

$$\frac{1}{1 + s} = f'(s)$$

$$= \frac{f(t) - f(0)}{t - 0}$$

$$= \frac{\log(1 + t) - \log 1}{t - 0}$$

$$= \frac{\log(1 + t)}{t}.$$  

It follows that

$$\log(1 + t) = \frac{t}{1 + s}.$$  

Now use the fact that the RHS is a decreasing function of $s$, if $t > 0$ and an increasing function of $s$, if $t < 0$ to get the first inequality.

Now suppose that $t \in (0, 1/2]$, so that $-t \in [-1/2, 0)$. As $-t \geq -1/2$, $1 - t \geq 1/2$ and so

$$\log(1 - t) > \frac{-t}{1 - t}$$

$$\geq -2t.$$  

□

Theorem 6.2.

$$\sum_{p \leq x} \frac{1}{p} > \frac{1}{2} \log \log x.$$  

In particular
\[ \sum_{p} \frac{1}{p} \]
diverges.

Proof. We start with the inequality of (5.1),
\[ \prod_{p \leq x} \left( 1 - \frac{1}{p} \right) < \frac{1}{\log x}. \]
If we take logs of both sides, we get
\[ \sum_{p \leq x} \log(1 - \frac{1}{p}) = \log \prod_{p \leq x} \left( 1 - \frac{1}{p} \right) < \log \frac{1}{\log x} \]
\[ = - \log \log x. \]

Now (6.1) implies that
\[ \left( 1 - \frac{1}{p} \right) > \frac{-2}{p} \]
for every prime \( p \) and so
\[ \sum_{p \leq x} \frac{2}{p} > \log \log x. \quad \square \]

**Theorem 6.3.**
\[ \sum_{p \leq x} \frac{\log p}{p} = \log x + O \left( \frac{\log x}{\log \log x} \right). \]

Proof. We saw in (3.3) that if we write \( n = p^r m \), where \( m \) is coprime to \( p \), then
\[ r = \frac{n}{p} + \frac{n^2}{p^2} + \frac{n^3}{p^3} + \ldots \]

It follows that
\[ n! = \prod_{p \leq n} p^{\frac{n}{p} + \frac{n^2}{p^2} + \frac{n^3}{p^3} + \ldots}, \]
so that
\[ \log n! = \sum_{p \leq n} \left( \frac{n}{p} + \frac{n^2}{p^2} + \frac{n^3}{p^3} + \ldots \right) \log p. \]

For the first term in every sum on the RHS we have
\[ \sum_{p \leq n} \frac{n}{p} \log p \leq \sum_{p \leq n} \frac{n}{p} \log p. \]
On the other hand,

\[
\sum_{p \leq n} \frac{n}{p} \log p \geq \sum_{p \leq n} \left( \frac{n}{p} - 1 \right) \log p
\]

\[
= \sum_{p \leq n} \frac{n}{p} \log p - \sum_{p \leq n} \log p
\]

\[
\geq \sum_{p \leq n} \frac{n}{p} \log p - \log n \sum_{p \leq n} 1.
\]

\[
= \sum_{p \leq n} \frac{n}{p} \log p - \pi(n) \log n.
\]

Note that

\[
\sum_{p \leq n} \left( \frac{n}{p^2} + \frac{n}{p^3} + \ldots \right) \log p \leq \sum_{p \leq n} \left( \frac{n}{p^2} + \frac{n}{p^3} + \ldots \right) \log p.
\]

Thus

\[
\log n! = \sum_{p \leq x} \frac{n}{p} \log p + O \left( \pi(n) \log n \right) + O \left( n \sum_{p \leq n} \left( \frac{1}{p^2} + \frac{1}{p^3} + \ldots \right) \log p \right)
\]

\[
= n \sum_{p \leq x} \frac{\log p}{p} + O \left( \pi(n) \log n \right) + O \left( n \sum_{p \leq n} \frac{\log p}{p(\log p - 1)} \right).
\]

Since the series

\[
\sum_{k=1}^{\infty} \frac{\log k}{k(k - 1)}
\]

converges, we get

\[
\log n! = n \sum_{p \leq x} \frac{\log p}{p} + O \left( \pi(n) \log n \right) + O(n).
\]

However it is shown in a homework problem that

\[
\log n! = n \log n + O(n).
\]

Putting all of this together we get

\[
n \sum_{p \leq x} \frac{\log p}{p} = n \log n + O(n) + O(\pi(n) \log n).
\]

But by (5.2)

\[
\pi(n) = O \left( \frac{n}{\log \log n} \right).
\]
Dividing through by \( n \) we get
\[
\sum_{p \leq x} \frac{\log p}{p} = \log n + O\left(\frac{\log n}{\log \log n}\right).
\]
This establishes the result when \( x = n \) is an integer. As
\[
\log x = \log_{\lfloor x \rfloor} + O(1),
\]
the result holds for any \( x \). \( \Box \)

We will need the following result, which is obtained using integration by parts.

**Theorem 6.4.** Suppose that \( \lambda_1, \lambda_2, \ldots \) is a sequence of reals such that
\[
\lambda_1 \leq \lambda_2 \leq \ldots
\]
and the limit is infinity. Let \( c_1, c_2, \ldots \) be any sequence of complex numbers and let \( f(x) \) be a function whose derivative is continuous for \( x \geq \lambda_1 \).

If
\[
C(x) = \sum_{\lambda_n \leq x} c_n,
\]
where the sum is over all \( n \) such that \( \lambda_n \leq x \), then
\[
\sum_{\lambda_n \leq x} c_n f(\lambda_n) = C(x) f(x) - \int_{\lambda_1}^x C(t) f'(t) \, dt.
\]

**Proof.** If \( \nu \) is the largest index such that \( \lambda_\nu \leq x \), we have
\[
\sum_{\lambda_n \leq x} c_n f(\lambda_n) = C(\lambda_1) f(\lambda_1) + (C(\lambda_2) - C(\lambda_1)) f(\lambda_2) + \cdots + (C(\lambda_\nu) - C(\lambda_{\nu-1})) f(\lambda_\nu)
\]
\[
= C(\lambda_1) (f(\lambda_1) - f(\lambda_2)) + \cdots + C(\lambda_{\nu-1}) (f(\lambda_{\nu-1}) - f(x)) + C(\lambda_\nu) f(x)
\]
\[
= - \int_{\lambda_1}^x C(t) f'(t) \, dt + C(x) f(x),
\]
since \( C(t) \) is constant over the intervals \( (\lambda_{i-1}, \lambda_i) \) and \( (\lambda_\nu, x) \). \( \Box \)

**Theorem 6.5.**
\[
\sum_{p \leq x} \frac{1}{p} \sim \log \log x.
\]

**Proof.** We are going to apply (6.4) to
\[
\lambda_n = p_n, \quad c_n = \frac{\log p_n}{p_n} \quad \text{and} \quad f(t) = \frac{1}{\log t}.
\]
We have
\[ \sum_{p \leq x} \frac{1}{p} = \sum_{2 < p < x} \left( \frac{\log p}{p} \cdot \frac{1}{\log p} \right) + \frac{1}{2} \]
\[ = \frac{1}{\log x} \sum_{2 < p < x} \frac{\log p}{p} - \int_3^x \left( \sum_{p \leq t} \frac{\log p}{p} \right) \frac{-dt}{t \log^2 t} + O(1) \]
\[ = \frac{1}{\log x} \left( \log x + O \left( \frac{\log x}{\log \log x} \right) \right) + \int_3^x \left( \log t + O \left( \frac{\log t}{\log \log t} \right) \right) \frac{dt}{t \log^2 t} + O(1) \]
\[ = \int_3^x \frac{dt}{t \log t} + \int_3^x O \left( \frac{1}{t \log t \log \log t} \right) dt + O(1) \]
\[ = \log \log x + \int_3^x O \left( \frac{1}{t \log t \log \log t} \right) dt + O(1) \]

For the last term there is some constant \( M \) such that
\[ \left| \int_3^x O \left( \frac{1}{t \log t \log \log t} \right) \frac{dt}{t \log t \log \log t} \right| \leq M \int_3^x \frac{dt}{t \log t \log \log t} \]
\[ = M \log \log \log x + O(1). \]

It follows that
\[ \sum_{p \leq x} \frac{1}{p} = \log \log x + O \left( \log \log \log x \right). \]