

## Lecture VI

I shall begin by describing the most general context in which the elementary approach to the prime number theorem works (at least at present). We consider the case of Beurling's generalized integers. Let us have a set of real numbers  $p_i$ ,  $i = 1, 2, 3, \dots$ ,

$$1 < p_1 \leq p_2 \leq \dots \leq p_i \leq \dots,$$

such that  $p_i \rightarrow \infty$  as  $i \rightarrow \infty$ . We form all possible finite products  $\prod p_i^{\alpha_i}$  and order them according to magnitude or size:

$$n_1 = 1, n_2 = p_1, \dots, n_i \leq n_{i+1}, \dots.$$

We denote by  $N(x)$  the number of  $n_i \leq x$  and assume we have an asymptotic law

$$N(x) = Ax + R(x),$$

where we shall assume  $A = 1$  and

$$R(x) = o\left(\frac{x}{\log^2 x}\right),$$

and shall try to establish the P.N.T., that is if we denote by  $\pi(x)$  the number of  $p_i \leq x$  we shall show that

$$\pi(x) = \frac{x}{\log x} + o\left(\frac{x}{\log x}\right),$$

in the form

$$\vartheta(x) = x + o(x),$$

where

$$\vartheta(x) = \sum_{p_i \leq x} \log p_i.$$

Similarly we use the notation

$$\psi(x) = \sum_{p_i^{\alpha} \leq x} \log p_i,$$

and

$$\Lambda(n_i) = \begin{cases} \log p_i & \text{if } n_i = p_i^{\alpha}, \alpha > 0, \\ 0 & \text{otherwise.} \end{cases}$$

For simplicity, I shall drop the indices and write  $n$ ,  $m$  or  $d$  for the generalized integers,  $p$ ,  $q$  and  $r$  for generalized primes, and use greek letters  $\mu$ ,  $\nu$  to denote ordinary integers, and  $x$ ,  $y$ ,  $t$ ,  $u$ ,  $v$  to denote real numbers. Constants will be denoted by capital latin letters.

We define  $\mu(n) = (-1)^\nu$  if  $n$  is the product of  $\nu$  distinct primes  $p$  and otherwise  $\mu(n) = 0$ . We also write

$$d|n \quad \text{if} \quad n = \prod p_i^{\alpha_i}, \quad d = \prod p_i^{\beta_i} \quad \text{and} \quad \alpha_i \geq \beta_i \quad \text{for all } i.$$

We have then

$$\sum_{d|n} \mu(d) = \begin{cases} 0 & \text{for } n \neq 1, \\ 1 & \text{for } n = 1. \end{cases}$$

We need some preliminary estimations. We first shall show

$$\sum_{d \leq x} \frac{\mu(d)}{d} = \mathcal{O}(1).$$

We have for  $x \geq 1$ ,

$$\begin{aligned} 1 &= \sum_{n \leq x} \sum_{d|n} \mu(d) = \sum_{d \leq x} \mu(d) N\left(\frac{x}{d}\right) \\ &= x \sum_{d \leq x} \frac{\mu(d)}{d} + \mathcal{O}\left(\sum_{d \leq x} \left|R\left(\frac{x}{d}\right)\right|\right). \end{aligned}$$

Here

$$\sum_{d \leq x} \left|R\left(\frac{x}{d}\right)\right| = x \sum_{d \leq x} \frac{\varepsilon\left(\frac{x}{d}\right)}{d(1 + \log \frac{x}{d})^2},$$

where we here and in the future denote by  $\varepsilon(x)$  a function that tends to zero as  $x \rightarrow \infty$ . Dividing the interval  $(1, x]$  into subintervals  $xe^{-\nu} \leq d \leq xe^{1-\nu}$  for  $\nu = 1, 2, \dots, \log x$ , we get

$$\begin{aligned} \sum_{d \leq x} \frac{\varepsilon\left(\frac{x}{d}\right)}{d(1 + \log \frac{x}{d})^2} &< \sum_{1 \leq \nu \leq \log x} \frac{\varepsilon(e^{\nu-1})N\left(\frac{x}{e^{\nu-1}}\right)}{xe^{-\nu}(1 + \nu - 1)^2} \\ &< A \sum_{\nu \geq 1} \frac{\varepsilon(e^{\nu-1})}{\nu^2} = \mathcal{O}(1). \end{aligned}$$

Thus

$$1 = x \sum_{d \leq x} \frac{\mu(d)}{d} + \mathcal{O}(x),$$

and

$$\sum_{d \leq x} \frac{\mu(d)}{d} = \mathcal{O}(1)$$

follows.

We next show

$$\sum_{d \leq x} \frac{\mu(d)}{d} \log \frac{x}{d} = o(\log \log x).$$

Consider first

$$\begin{aligned} \sum_{n \leq y} \frac{1}{n} &= \int_1^y \frac{1}{t} dN(t) = \frac{N(y)}{y} + \int_1^y \frac{N(t)}{t^2} dt \\ &= 1 + o\left(\frac{1}{\log^2 y}\right) + \log y + \int_1^y \frac{\varepsilon(t)}{t(1 + \log t)^2} dt \\ &= \log y + 1 + o\left(\frac{1}{\log^2 y}\right) + \int_1^\infty \frac{\varepsilon(t) dt}{t(1 + \log t)^2} \\ &\quad + o\left(\int_y^\infty \frac{dt}{t(1 + \log t)^2}\right) \\ &= \log y + C + o\left(\frac{1}{1 + \log y}\right). \end{aligned}$$

So we can write

$$\log \frac{x}{d} = \sum_{n \leq \frac{x}{d}} \frac{1}{n} - C + \frac{\varepsilon(\frac{x}{d})}{1 + \log \frac{x}{d}},$$

and so

$$\begin{aligned} \sum_{d \leq x} \frac{\mu(d)}{d} \log \frac{x}{d} &= \sum_{d \leq x} \frac{\mu(d)}{d} \sum_{n \leq \frac{x}{d}} \frac{1}{n} - C \sum_{d \leq x} \frac{\mu(d)}{d} \\ &\quad + \mathcal{O}\left(\sum_{d \leq x} \frac{\varepsilon(\frac{x}{d})}{d(1 + \log \frac{x}{d})}\right). \end{aligned}$$

Here

$$\begin{aligned} \sum_{d \leq x} \frac{\mu(d)}{d} \sum_{n \leq \frac{x}{d}} \frac{1}{n} &= \sum_{nd \leq x} \frac{\mu(d)}{nd} \\ &= \sum_{m \leq x} \frac{1}{m} \sum_{d|m} \mu(d) = 1, \end{aligned}$$

also

$$\sum_{d \leq x} \frac{\mu(d)}{d} = \mathcal{O}(1),$$

and

$$\begin{aligned} \sum_{d \leq x} \frac{\varepsilon\left(\frac{x}{d}\right)}{d(1 + \log \frac{x}{d})} &< \sum_{1 \leq \nu \leq \log x} \frac{\varepsilon(e^{\nu-1})e^{\nu} N(xe^{1-\nu})}{x(1 + \nu - 1)} \\ &< A \sum_{1 \leq \nu \leq \log x} \frac{\varepsilon(e^{\nu-1})}{\nu} = o(\log \log x). \end{aligned}$$

This gives

$$\sum_{d \leq x} \frac{\mu(d)}{d} \log \frac{x}{d} = o(\log \log x).$$

We finally show that

$$\begin{aligned} \sum_{d \leq x} \frac{\mu(d)}{d} \log^2 \frac{x}{d} &= \log x + \sum_{n \leq x} \frac{\Lambda(n)}{n} + o(\log x) \\ &= \log x + \sum_{p \leq x} \frac{\log p}{p} + o(\log x). \end{aligned}$$

We have

$$\begin{aligned} \sum_{d \leq x} \frac{\mu(d)}{d} \log^2 \frac{x}{d} &= \sum_{d \leq x} \frac{\mu(d)}{d} \log \frac{x}{d} \sum_{n \leq \frac{x}{d}} \frac{1}{n} - C \sum_{d \leq x} \frac{\mu(d)}{d} \log \frac{x}{d} \\ &\quad + \mathcal{O} \left( \sum_{d \leq x} \frac{\varepsilon\left(\frac{x}{d}\right) \log \frac{x}{d}}{d(1 + \log \frac{x}{d})} \right). \end{aligned}$$

The first term on the right-hand side equals

$$\sum_{nd \leq x} \frac{\mu(d) \log \frac{x}{d}}{nd} = \sum_{m \leq x} \frac{1}{m} \sum_{d|m} \mu(d) \log \frac{x}{d},$$

but

$$\sum_{d|m} \mu(d) \log \frac{x}{d} = \begin{cases} \log x, & \text{for } m = 1, \\ \Lambda(m), & \text{for } m > 1. \end{cases}$$

So

$$\begin{aligned} \sum_{d \leq x} \frac{\mu(d)}{d} \log \frac{x}{d} \sum_{n \leq \frac{x}{d}} \frac{1}{n} &= \log x + \sum_{n \leq x} \frac{\Lambda(n)}{n} \\ &= \log x + \sum_{p \leq x} \frac{\log p}{p} + \mathcal{O}(1). \end{aligned}$$

Also

$$\begin{aligned} \sum_{d \leq x} \frac{\varepsilon\left(\frac{x}{d}\right) \log \frac{x}{d}}{d(1 + \log \frac{x}{d})} &< \sum_{d \leq x} \frac{\varepsilon\left(\frac{x}{d}\right)}{d} < \sum_{1 \leq \nu \leq \log x} \frac{\varepsilon(e^{\nu-1})e^{\nu}}{x} N\left(\frac{x}{e^{\nu-1}}\right) \\ &< A \sum_{1 \leq \nu \leq \log x} \varepsilon(e^{\nu-1}) = o(\log x). \end{aligned}$$

Thus altogether we get:

$$\sum_{d \leq x} \frac{\mu(d)}{d} \log^2 \frac{x}{d} = \log x + \sum_{p \leq x} \frac{\log p}{p} + o(\log x).$$

Now consider

$$\begin{aligned} \sum_{n \leq x} \sum_{d|n} \mu(d) \log^2 \frac{x}{d} &= \sum_{d \leq x} \mu(d) \log^2 \frac{x}{d} N\left(\frac{x}{d}\right) \\ &= x \sum_{d \leq x} \frac{\mu(d)}{d} \log^2 \frac{x}{d} + \mathcal{O}\left(x \sum_{d \leq x} \frac{\varepsilon\left(\frac{x}{d}\right) \log^2 \frac{x}{d}}{d(1 + \log \frac{x}{d})^2}\right). \end{aligned}$$

The last term is

$$\mathcal{O}\left(x \sum_{d \leq x} \frac{\varepsilon\left(\frac{x}{d}\right)}{d}\right) = o(x \log x),$$

so we get altogether

$$\sum_{n \leq x} \sum_{d|n} \mu(d) \log^2 \frac{x}{d} = x \log x + x \sum_{p \leq x} \frac{\log p}{p} + o(x \log x).$$

On the other hand, it is easily verified that for  $n \leq x$

$$\sum_{d|n} \mu(d) \log^2 \frac{x}{d} = \begin{cases} \log^2 x, & \text{for } n = 1, \\ \log \frac{x^2}{p} \log p, & \text{for } n = p^{\nu}, \nu > 0, \\ 2 \log p \log q, & \text{for } n = p^{\nu} q^{\mu}, p \neq q, \\ 0, & \text{if } n \text{ has 3 or more prime factors.} \end{cases}$$

Seen either by induction, or by noting that it is the second derivative of

$$x^y \sum_{d|n} \mu(d) d^{-y} = x^y \prod_{p|n} (1 - p^{-y})$$

at  $y = 0$ . So we have also

$$\begin{aligned} \sum_{n \leq x} \sum_{d|n} \mu(d) \log^2 \frac{x}{d} &= \log^2 x + \sum_{p^{\nu} \leq x} \log \frac{x^2}{p} \log p \\ &\quad + \sum_{\substack{p^{\nu} q^{\mu} \leq x \\ p \neq q}} \log p \log q. \end{aligned}$$

Comparing results we get

$$\begin{aligned} \sum_{p \leq x} \log \frac{x^2}{p} \log p + \sum_{pq \leq x} \log p \log q \\ = x \log x + x \sum_{p \leq x} \frac{\log p}{p} + o(x \log x). \end{aligned}$$

We still need to estimate the sum

$$\sum_{p \leq x} \frac{\log p}{p}.$$

We consider the sum

$$\begin{aligned} \sum_{n \leq x} \log n &= \int_1^x \log t \, dN(t) \\ &= N(x) \log x - \int_1^x \frac{N(t)}{t} \, dt \\ &= x(\log x - 1) + o\left(\frac{x}{\log x}\right). \end{aligned}$$

Also since

$$\log n = \sum_{d|n} \Lambda(d),$$

we have, using  $N(x) > \frac{1}{A}x$  for  $x \geq 1$ , that

$$\sum_{n \leq x} \log n = \sum_{d \leq x} \Lambda(d) N\left(\frac{x}{d}\right) > \frac{1}{A}x \sum_{d \leq x} \frac{\Lambda(d)}{d}.$$

Comparing this with the result above we get at first

$$\sum_{d \leq x} \frac{\Lambda(d)}{d} = \mathcal{O}(\log x).$$

Inserting this bound in our earlier inequality we get at first

$$\vartheta(x) = \sum_{p \leq x} \log p = \mathcal{O}(x),$$

or

$$\psi(x) = \sum_{d \leq x} \Lambda(d) = \mathcal{O}(x).$$

We use this to estimate

$$\sum_{d \leq x} \Lambda(d) N\left(\frac{x}{d}\right)$$

better. We get

$$\sum_{d \leq x} \Lambda(d) N\left(\frac{x}{d}\right) = x \sum_{d \leq x} \frac{\Lambda(d)}{d} + x \sum_{d \leq x} \frac{\Lambda(d)}{d} \frac{\varepsilon\left(\frac{x}{d}\right)}{\left(1 + \log \frac{x}{d}\right)^2}.$$

Here

$$\begin{aligned} x \sum_{d \leq x} \frac{\Lambda(d)}{d} \frac{\varepsilon\left(\frac{x}{d}\right)}{\left(1 + \log \frac{x}{d}\right)^2} &\leq \sum_{1 \leq \nu \leq \log x} \psi\left(xe^{1-\nu}\right) \frac{e^\nu \varepsilon\left(e^{\nu-1}\right)}{\nu^2} \\ &< Ax \sum_{\nu > 1} \frac{\varepsilon\left(e^{\nu-1}\right)}{\nu^2} = \mathcal{O}(x). \end{aligned}$$

Thus

$$\sum_{d \leq x} \Lambda(d) N\left(\frac{x}{d}\right) = x \sum_{d \leq x} \frac{\Lambda(d)}{d} + \mathcal{O}(x).$$

But since also

$$\sum_{d \leq x} \Lambda(d) N\left(\frac{x}{d}\right) = x(\log x - 1) + o\left(\frac{x}{\log x}\right)$$

we get by comparison

$$\sum_{d \leq x} \frac{\Lambda(d)}{d} = \log x + \mathcal{O}(1)$$

or

$$\sum_{p \leq x} \frac{\log p}{p} = \log x + \mathcal{O}(1),$$

and so

$$\begin{aligned} \sum_{p \leq x} \log \frac{x^2}{p} \log p + \sum_{pq \leq x} \log p \log q \\ = 2x \log x + o(x \log x). \end{aligned}$$

Since from  $\vartheta(x) = \mathcal{O}(x)$

$$\sum_{p \leq x} \log \frac{x}{p} \log p = \int_1^x \log \frac{x}{t} d\vartheta(t) = \int_1^x \frac{\vartheta(t)}{t} dt = \mathcal{O}(x),$$

we may rewrite the asymptotic formula above as

$$\sum_{p \leq x} \log^2 p + \sum_{pq \leq x} \log p \log q = 2x \log x + o(x \log x)$$

or

$$\log x \vartheta(x) + \sum_{p \leq x} \log p \vartheta\left(\frac{x}{p}\right) = 2x \log x + o(x \log x).$$

We also can rewrite

$$\sum_{p \leq x} \frac{\log p}{p} = \log x + \mathcal{O}(1)$$

as

$$\int_1^x \frac{\vartheta(t)}{t^2} dt = \log x + \mathcal{O}(1).$$

By partial summation we get from

$$\sum_{p \leq x} \log^2 p + \sum_{pq \leq x} \log p \log q = 2x \log x + o(x \log x)$$

that

$$\sum_{p \leq x} \log p + \sum_{pq \leq x} \frac{\log p \log q}{\log pq} = 2x + o(x).$$

From this

$$\begin{aligned} \sum_{pq \leq x} \log p \log q &= \sum_{p \leq x} \log p \sum_{q \leq \frac{x}{p}} \log q \\ &= 2x \sum_{p \leq x} \frac{\log p}{p} - \sum_{p \leq x} \log p \sum_{qr \leq x} \frac{\log q \log r}{\log qr} \\ &\quad + \mathcal{O}\left(x \sum_{p \leq x} \varepsilon\left(\frac{x}{p}\right) \frac{\log p}{p}\right) \\ &= 2x \log x - \sum_{qr \leq x} \frac{\log q \log r}{\log qr} \vartheta\left(\frac{x}{qr}\right) + o(x \log x). \end{aligned}$$

Combining this with

$$\log x \vartheta(x) + \sum_{pq \leq x} \log p \log q = 2x \log x + o(x \log x),$$

we get

$$\log x \vartheta(x) = \sum_{pq \leq x} \frac{\log p \log q}{\log pq} \vartheta\left(\frac{x}{pq}\right) + o(x \log x).$$



If we write  $\vartheta(x) = x + \rho(x)$  the two last equations give

$$\log x \rho(x) = - \sum_{p \leq x} \log p \rho\left(\frac{x}{p}\right) + o(x \log x)$$

and

$$\log x \rho(x) = \sum_{pq \leq x} \frac{\log p \log q}{\log pq} \rho\left(\frac{x}{pq}\right) + o(x \log x).$$

From this

$$2 \log x |\rho(x)| \leq \sum_{p \leq x} \log p \left| \rho\left(\frac{x}{p}\right) \right| + \sum_{pq \leq x} \frac{\log p \log q}{\log pq} \left| \rho\left(\frac{x}{pq}\right) \right| + o(x \log x).$$

If we write

$$s(y) = \sum_{p \leq y} \log p + \sum_{pq \leq y} \frac{\log p \log q}{\log pq} = 2y + o(y),$$

we may rewrite the above inequality as

$$2 \log x |\rho(x)| \leq \int_1^x \left| \rho\left(\frac{x}{t}\right) \right| ds(t) + o(x \log x).$$

With some manipulation and using

$$|\rho(y) - \rho(y')| \leq y - y' + o(y)$$

for  $y > y'$ , we get

$$\begin{aligned} |\rho(x)| &\leq \frac{1}{\log x} \int_1^x \left| \rho\left(\frac{x}{t}\right) \right| dt + o(x) \\ &= \frac{x}{\log x} \int_1^x \frac{|\rho(t)|}{t^2} dt + o(x). \end{aligned}$$

Assume now that we have shown  $|\rho(x)| \leq \alpha x$  for  $x > x_0$ , and some constant  $\alpha$ . From

$$\int_1^x \frac{\rho(t)}{t^2} dt = \mathcal{O}(1)$$

we see that

$$\left| \int_y^{ye^\lambda} \frac{\rho(t)}{t^2} dt \right| \leq A,$$

so either there is a  $t_0$  in  $y, ye^\lambda$  with  $|\rho(t_0)| < \log t_0$  or we have

$$\int_y^{ye^\lambda} \frac{|\rho(t)|}{t^2} dt \leq A.$$

In the second case, choosing  $\lambda$  so large that  $\lambda\alpha = 2A$ , we have

$$\int_y^{ye^\lambda} \frac{|\rho(t)|}{t^2} dt \leq \frac{1}{2}\alpha \int_y^{ye^\lambda} \frac{dt}{t}.$$

In the first case we have

$$|\rho(t)| \leq \log t_0 + |t - t_0| + \epsilon t \leq \frac{\alpha}{2}t$$

in an interval of length  $t_0, t_0 + \frac{\alpha}{3}t_0$ .