LECTURE NOTES 3 FOR CAMBRIDGE PART III COURSE ON "PROBABILISTIC NUMBER THEORY", MICHAELMAS 2015

ADAM J HARPER

Abstract. These are rough notes covering the third block of lectures in the "Probabilistic Number Theory" course. In these lectures we will see how the Euler product for the Riemann zeta function and the multiplicative independence of distinct primes provides some "almost independent" structure, for varying imaginary part. This ultimately leads to the Selberg central limit theorem for $\log|\zeta(1/2+it)|$, as t varies.

(No originality is claimed for any of the contents of these notes. In particular, they borrow substantially from the paper [1] of Radziwiłł and Soundararajan.)

13. "Almost independence" for the zeta function

Back at the start of the course, we observed that if $f(n)$ is an additive function and if $n \leq N$ then we can write

$$f(n) = \sum_{p \leq N} \left(\sum_{k=1}^{\infty} f(p^k) \mathbf{1}_{p^k || n} \right) =: \sum_{p \leq N} f_p(n).$$

Since the functions $f_p(n)$ behave "almost independently" as $n \leq N$ varies, this ultimately leads to the Erdős–Kac central limit theorem for additive functions.

In this chapter we will see that we can obtain an analogous decomposition for $\log \zeta(s)$ or $\log|\zeta(s)| = \Re \log \zeta(s)$. The easiest setting to see this is when $\Re(s) > 1$.

Lemma 13.1. *For any $\sigma > 1$ and any $t \in \mathbb{R}$, we have*

$$\log \zeta(\sigma+it) = \sum_p \sum_{k=1}^{\infty} \frac{1}{k} \frac{1}{p^{k(\sigma+it)}}, \quad \textit{and} \quad \log|\zeta(\sigma+it)| = \sum_p \sum_{k=1}^{\infty} \frac{1}{k} \frac{\cos(kt\log p)}{p^{k\sigma}}.$$

The double series on the right are absolutely convergent.

Proof of Lemma 13.1. In view of the Euler product expression for the zeta function (Lemma 10.2), we have

$$\log \zeta(\sigma+it) = -\sum_p \log\left(1 - \frac{1}{p^{\sigma+it}}\right), \quad \sigma > 1,\ t \in \mathbb{R}.$$

Date: 8th January 2016.

The lemma follows on inserting the Taylor series expansion of $\log\left(1 - \frac{1}{p^{\sigma+it}}\right)$, and (for the statement about $\log|\zeta(\sigma+it)|$) on taking real parts and noting that $\Re p^{-ikt} = \cos(-kt\log p) = \cos(kt\log p)$. □

In the case of additive functions, the summands $f_p(n)$ were "almost independent" because divisibility by distinct primes are "almost independent" events, by the Chinese Remainder Theorem. In the case of the zeta function we will think of the imaginary part t as varying over some interval $[T, 2T]$, and the almost independence arises because the values $\log p$ are linearly independent over $\mathbb{Q}$ (this is just a restatement of the uniqueness of prime factorisations), so the terms $\cos(kt\log p)$ vary almost independently for distinct primes p. To explore this, we first prove a simple integral estimate.

Lemma 13.2. *Let $A(s) = \sum_{n\leq X} \frac{a_n}{n^s}$ and $B(s) = \sum_{n\leq X} \frac{b_n}{n^s}$, where the a_n, b_n are arbitrary complex numbers. Then for any $T \geq 0$ and any $\sigma \in \mathbb{R}$ we have*

$$\int_T^{2T} A(\sigma+it)\overline{B(\sigma+it)}dt = T\sum_{n\leq X} \frac{a_n\overline{b_n}}{n^{2\sigma}} + O\left(X \sum_{\substack{m,n\leq X, \\ m\neq n}} \frac{|a_n|}{n^\sigma}\frac{|b_m|}{m^\sigma}\right).$$

Proof of Lemma 13.2. If we simply insert the definitions of $A(\sigma+it)$ and $B(\sigma+it)$, we find

$$\int_T^{2T} A(\sigma+it)\overline{B(\sigma+it)}dt = \sum_{n\leq X}\frac{a_n}{n^\sigma}\sum_{m\leq X}\frac{\overline{b_m}}{m^\sigma}\int_T^{2T}\frac{1}{n^{it}}\frac{1}{m^{-it}}dt.$$

The terms where $m = n$ give the first sum in the statement of the Lemma, whilst if $m \neq n$ we have

$$\int_T^{2T}\frac{1}{n^{it}}\frac{1}{m^{-it}}dt = \int_T^{2T} e^{it\log(m/n)} \ll \frac{1}{\log(m/n)} \ll X,$$

since if $1 \leq n < m \leq X$ then $\log(m/n) = \log(1 + \frac{m-n}{n}) \geq \log(1+\frac{1}{n}) \gg 1/X$ (similarly if $1 \leq m < n \leq X$). This gives the "big Oh" term claimed in the Lemma. □

Notice that the "big Oh" term here does not grow with T, so we can expect the first sum typically to dominate once T is large enough.

Lemma 13.3 (Truncated Moments for log zeta)**.** *For any $X \geq 2$, any $T \geq 0$, any $\sigma \in \mathbb{R}$ and even $j \in \mathbb{N}$ we have*

$$\int_T^{2T}\left(\sum_{p\leq X}\frac{\cos(t\log p)}{p^\sigma}\right)^j dt \;=\; T\frac{j!}{2^j(j/2)!}\left(\sum_{p\leq X}\frac{1}{p^{2\sigma}}\right)^{j/2}\left(1 + O_j\left(\frac{\sum_{p\leq X}\frac{1}{p^{4\sigma}}}{\left(\sum_{p\leq X}\frac{1}{p^{2\sigma}}\right)^2}\right)\right) + $$
$$+O_j\left(X^j\left(\sum_{n\leq X^j}\frac{1}{n^\sigma}\right)^2\right),$$

whilst for any odd $j \in \mathbb{N}$ we have

$$\int_T^{2T} \left(\sum_{p \le X} \frac{\cos(t \log p)}{p^\sigma} \right)^j dt \ll_j X^j \left(\sum_{n \le X^j} \frac{1}{n^\sigma} \right)^2 .$$

Proof of Lemma 13.3. Note first that we can write $\cos(t \log p) = \Re p^{it} = (1/2)(p^{it}+p^{-it})$, and therefore we always have

$$\begin{aligned} \int_T^{2T} \left(\sum_{p \le X} \frac{\cos(t \log p)}{p^\sigma} \right)^j dt &= \frac{1}{2^j} \int_T^{2T} \left(\sum_{p \le X} \frac{1}{p^{\sigma+it}} + \sum_{p \le X} \frac{1}{p^{\sigma-it}} \right)^j dt \\ &= \frac{1}{2^j} \sum_{k=0}^{j} \binom{j}{m} \int_T^{2T} \left(\sum_{p \le X} \frac{1}{p^{\sigma+it}} \right)^k \left(\sum_{p \le X} \frac{1}{p^{\sigma-it}} \right)^{j-k} dt. \end{aligned}$$

Furthermore, if we expand out we find that $\left(\sum_{p \le X} \frac{1}{p^{\sigma+it}} \right)^k = \sum_{n \le X^k} \frac{a_n(k)}{n^{\sigma+it}} =: A_k(\sigma+it)$, say, where $a_n(k)$ denotes the number of ways of writing n as a product of k primes less than X (with primes counted with multiplicity, and different orderings counted as distinct), and with $a_n(k) = 0$ if there is no such representation of n.

Using this notation, and applying Lemma 13.2, we get that $\int_T^{2T} \left(\sum_{p \le X} \frac{\cos(t \log p)}{p^\sigma} \right)^j dt$ is

$$\begin{aligned} &= \frac{1}{2^j} \sum_{k=0}^{j} \binom{j}{m} \int_T^{2T} A_k(\sigma+it) \overline{A_{j-k}(\sigma+it)} dt \\ &= \frac{1}{2^j} \sum_{k=0}^{j} \binom{j}{m} \left(T \sum_{n \le X^j} \frac{a_n(k) a_n(j-k)}{n^{2\sigma}} + O\left(X^j \sum_{\substack{m,n \le X^j, \\ m \ne n}} \frac{a_n(k)}{n^\sigma} \frac{a_m(j-k)}{m^\sigma} \right) \right). \end{aligned}$$

We always have $a_n(k) \le k! \le j!$, similarly for $a_m(j-k)$, so the total contribution from the "big Oh" terms may be bounded as stated in the lemma.

In the other sums, note that if $a_n(k) \ne 0$ and $a_n(j-k) \ne 0$ then n must be a product of k prime factors and also a product of $j-k$ prime factors. This can only happen if $k = j-k$, in other words if j is even and if $k = j/2$. If j is even, the total contribution from all those sums becomes

$$T \frac{\binom{j}{j/2}}{2^j} \sum_{n \le X^j} \frac{a_n(j/2)^2}{n^{2\sigma}} = T \frac{\binom{j}{j/2}}{2^j} \left(\sum_{\substack{n \le X^j, \\ n \text{ product of } j/2 \text{ distinct } p \le X}} \frac{a_n(j/2)^2}{n^{2\sigma}} + O_j(\sum_{\substack{n \le X^j, \\ n \text{ prod. of } j/2 \text{ non-distinct } p \le X}} \frac{1}{n^{2\sigma}}) \right)$$

Finally, if n is a product of $j/2$ *distinct* primes less than X then we have $a_n(j/2) = (j/2)!$, so we can rewrite the above as

$$T\frac{\binom{j}{j/2}}{2^j}(j/2)!\left(\sum_{n\leq X^j}\frac{a_n(j/2)}{n^{2\sigma}} + O_j(\sum_{\substack{n\leq X^j, \\ n \text{ prod. of } j/2 \text{ non-distinct } p\leq X}}\frac{1}{n^{2\sigma}})\right).$$

Recalling the definition of $a_n(j/2)$, we see $\sum_{n\leq X^j}\frac{a_n(j/2)}{n^{2\sigma}} = \left(\sum_{p\leq X}\frac{1}{p^{2\sigma}}\right)^{j/2}$, which gives the first term in the statement of the lemma (for j even). Meanwhile we have

$$\sum_{\substack{n\leq X^j, \\ n \text{ prod. of } j/2 \text{ non-distinct } p\leq X}}\frac{1}{n^{2\sigma}} \leq \sum_{p\leq X}\frac{1}{p^{4\sigma}}\sum_{n: \text{ prod. of } j/2-2 \text{ primes } p\leq X}\frac{1}{n^{2\sigma}} \leq \left(\sum_{p\leq X}\frac{1}{p^{4\sigma}}\right)\left(\sum_{p\leq X}\frac{1}{p^{2\sigma}}\right)^{j/2-2},$$

which gives the multiplicative error term in the statement of the Lemma. □

Equipped with the Truncated Moments Lemma for log zeta, we can think about whether it is possible to deduce a central limit type theorem. To do this we first want the "big Oh" terms to be of smaller order than the supposed main terms $T\frac{j!}{2^j(j/2)!}\left(\sum_{p\leq X}\frac{1}{p^{2\sigma}}\right)^{j/2}$. Similarly as in our proof of the Erdős-Kac theorem (Theorem 5.1), we will arrange this by taking $X = T^{1/\phi(T)}$, where $\phi(T)$ is some function that tends to infinity with T.

Having made this restriction, let $\mathbb{P}_T$ denote the continuous uniform probability measure on the interval $[T, 2T]$ (where $T > 0$), and let $\mathbb{E}_T$ denote the expectation induced by $\mathbb{P}_T$. (This clashes slightly with the notation $\mathbb{P}_N$ that we used in an analogous discrete setting in Chapter 1, but there should be no confusion in practice.) *If* it turns out that the "big Oh" terms are all small, then Lemma 13.3 will imply that

$$\mathbb{E}_T\sum_{p\leq T^{1/\phi(T)}}\frac{\cos(t\log p)}{p^\sigma} = \frac{1}{T}\int_T^{2T}\sum_{p\leq T^{1/\phi(T)}}\frac{\cos(t\log p)}{p^\sigma}dt \to 0 \quad \text{as } T\to\infty,$$

and also

$$\mathbb{E}_T\left(\sum_{p\leq T^{1/\phi(T)}}\frac{\cos(t\log p)}{p^\sigma}\right)^2 = \frac{1}{T}\int_T^{2T}\left(\sum_{p\leq T^{1/\phi(T)}}\frac{\cos(t\log p)}{p^\sigma}\right)^2 dt \sim \frac{1}{2}\sum_{p\leq T^{1/\phi(T)}}\frac{1}{p^{2\sigma}} \quad \text{as } T\to\infty.$$

Back in Remark 5.2, we noted that we cannot hope to prove a normal limit theorem unless the variance $\frac{1}{2}\sum_{p\leq T^{1/\phi(T)}}\frac{1}{p^{2\sigma}}$ tends to infinity as $T\to\infty$, which can only happen if neither σ nor $\phi(T)$ is too large.

Bearing in mind the above discussion, we formulate and prove the following result.

Proposition 13.4. *Let $\phi(T)$ be any function that satisfies $\phi(T)\to\infty$ as $T\to\infty$, but also $\log\phi(T) = o(\log\log T)$ as $T\to\infty$. Further, let $W(T)$ be any function that satisfies $0\leq W(T)\leq\phi(T)/2$ for all large T.*

Then under the probability measure $\mathbb{P}_T$, we have

$$\frac{\sum_{p \leq T^{1/\phi(T)}} \frac{\cos(t \log p)}{p^{1/2+W(T)/\log T}}}{\sqrt{(1/2)\log\log T}} \xrightarrow{d} N(0,1) \quad \text{as } T \to \infty.$$

Proof of Proposition 13.4. In view of the Method of Moments (Corollary 4.6), it will suffice to prove that for each fixed $j \in \mathbb{N}$ we have

$$\mathbb{E}_T \left(\frac{\sum_{p \leq T^{1/\phi(T)}} \frac{\cos(t \log p)}{p^{1/2+W(T)/\log T}}}{\sqrt{(1/2)\log\log T}} \right)^j \to m_j \quad \text{as } T \to \infty,$$

where m_j denote the standard normal moments. We recall from the Normal Moments Lemma (Lemma 4.5) this means that $m_j = 0$ when j is odd, and $m_j = \frac{j!}{2^{j/2}(j/2)!}$ when j is even.

When j is odd, Lemma 13.3 implies that

$$\mathbb{E}_T \left(\frac{\sum_{p \leq T^{1/\phi(T)}} \frac{\cos(t \log p)}{p^{1/2+W(T)/\log T}}}{\sqrt{(1/2)\log\log T}} \right)^j \ll_j \frac{1}{T(\log\log T)^{j/2}} T^{j/\phi(T)} \left(\sum_{n \leq T^{j/\phi(T)}} \frac{1}{n^{1/2+W(T)/\log T}} \right)^2 .$$

We can upper bound the right hand side extremely crudely by $\frac{1}{T} T^{3j/\phi(T)}$, and this does tend to 0 as $T \to \infty$ since we assume that $\phi(T) \to \infty$.

When j is even, Lemma 13.3 implies that

$$\mathbb{E}_T \left(\frac{\sum_{p \leq T^{1/\phi(T)}} \frac{\cos(t \log p)}{p^{1/2+W(T)/\log T}}}{\sqrt{(1/2)\log\log T}} \right)^j$$

$$= \frac{1}{((1/2)\log\log T)^{j/2}} \frac{j!}{2^j (j/2)!} \left(\sum_{p \leq T^{1/\phi(T)}} \frac{1}{p^{1+2W(T)/\log T}} \right)^{\frac{j}{2}} \left(1 + O_j \left(\frac{\sum_{p \leq T^{1/\phi(T)}} \frac{1}{p^{2+4W(T)/\log T}}}{\left(\sum_{p \leq T^{1/\phi(T)}} \frac{1}{p^{1+2W(T)/\log T}} \right)^2} \right) \right.$$

$$\left. + O_j \left(\frac{1}{T(\log\log T)^{j/2}} T^{j/\phi(T)} \left(\sum_{n \leq T^{j/\phi(T)}} \frac{1}{n^{1/2+W(T)/\log T}} \right)^2 \right) \right).$$

Now notice that we always have $\sum_{p \leq X} \frac{1}{p^{1+w}} = \sum_{p \leq X} \frac{1}{p} e^{-w \log p}$, and if $0 \leq w \log p \leq 1$ we can write this as $\sum_{p \leq X} \frac{1}{p}(1 + O(w \log p))$. In the special case above we have $\frac{2W(T)\log p}{\log T} \leq \frac{2W(T)}{\phi(T)} \leq 1$ for all large T, by assumption about $W(T)$, and so

$$\sum_{p \leq T^{1/\phi(T)}} \frac{1}{p^{1+2W(T)/\log T}} = \sum_{p \leq T^{1/\phi(T)}} \frac{1}{p} + O\left(\frac{W(T)}{\log T} \sum_{p \leq T^{1/\phi(T)}} \frac{\log p}{p} \right) = \log\log T - \log\phi(T) + O(1),$$

using the estimates of Mertens (Fact 2 from Chapter 0). In the Proposition we assume that $\log \phi(T) = o(\log\log T)$, so overall the above is $(1 + o(1)) \log\log T$.

Putting everything together, when j is even we see $\mathbb{E}_T \left(\frac{\sum_{p \leq T^{1/\phi(T)}} \frac{\cos(t \log p)}{p^{1/2+W(T)/\log T}}}{\sqrt{(1/2)\log\log T}} \right)^j$ is

$$= (1+o_j(1)) \frac{1}{(1/2)^{j/2}} \frac{j!}{2^j (j/2)!} (1 + O_j(\frac{1}{(\log\log T)^2})) + O_j(\frac{1}{T} T^{3j/\phi(T)}).$$

Here we estimated the final "big Oh" term as we did for odd j. The above expression $\to \frac{j!}{2^{j/2}(j/2)!} = m_j$ as $T \to \infty$, which proves the Proposition. □

Comparing Proposition 13.4 with Lemma 13.1, we have established a central limit theorem for the partial sums over primes when σ is very close to 1/2, but we only know that these partial sums approximate $\log|\zeta(\sigma+it)|$ on the different range $\sigma > 1$. It turns out that (for a suitable choice of $\phi(T)$ and $W(T)$) these partial sums do approximate $\log|\zeta(\sigma+it)|$, *on average over* $T \leq t \leq 2T$, for σ very close to 1/2, but unlike in the Erdős–Kac theorem it requires quite a lot of work to show this. This will be the subject of the next section.

14. Selberg's Central Limit Theorem

To finish the course, we will say as much as we can about the proof of the following result, which is one of the most classical and fundamental probabilistic results on the zeta function.

Theorem 14.1 (Selberg's Central Limit Theorem, Selberg, 1946)**.** *Under the probability measure* $\mathbb{P}_T$*, we have*

$$\frac{\log|\zeta(1/2+it)|}{\sqrt{(1/2)\log\log T}} \stackrel{d}{\to} N(0,1) \quad \text{as } T \to \infty.$$

Remark 14.2. Although $\zeta(s)$ has infinitely many zeros on the line $\Re(s) = 1/2$, it only has finitely many in any interval $T \leq t \leq 2T$, so the continuous uniform measure $\mathbb{P}_T$ doesn't see them.

In spite of the above remark, it is difficult to work with $\log|\zeta(\sigma+it)|$ when $\sigma \leq 1$ since it does blow up at any zero of the zeta function. Thus it is difficult even to start directly trying to show that $\log|\zeta(\sigma+it)| \approx \sum_{p \leq T^{1/\phi(T)}} \frac{\cos(t \log p)}{p^\sigma}$ for most t, which is our approximate strategy for deducing Selberg's Central Limit Theorem from Proposition 13.4. To get around this, an obvious idea is to take exponentials and try instead to show that $\zeta(\sigma+it) \approx \exp\{\sum_{p \leq T^{1/\phi(T)}} \frac{1}{p^{\sigma+it}}\}$, or more-or-less equivalently to try to show that $\zeta(\sigma+it)\exp\{-\sum_{p \leq T^{1/\phi(T)}} \frac{1}{p^{\sigma+it}}\} \approx 1$.

We know how to approximate $\zeta(\sigma+it)$ (using Definition 10.3), so our difficulties have now switched to approximating $\exp\{-\sum_{p \leq T^{1/\phi(T)}} \frac{1}{p^{\sigma+it}}\}$, for most t, by something simple enough that we can compare it with $\zeta(\sigma+it)$. The following lemma is the key to doing this.

Lemma 14.3. *Let T be large, and suppose that $\sigma = \sigma(T) \geq 1/2$. For a complex number s, define $P_1(s) := \sum_{p^k \leq T^{1/(\log\log T)^2}} \frac{1}{kp^{ks}}$ and $P_2(s) := \sum_{T^{1/(\log\log T)^2} < p^k \leq T^{1/(\log\log\log T)^2}} \frac{1}{kp^{ks}}$. Then*

$$\mathbb{P}_T(|P_1(\sigma+it)| > \log\log T) \to 0 \quad \text{as } T \to \infty, \quad \text{and} \quad \mathbb{P}_T(|P_2(\sigma+it)| > \log\log\log T) \to 0 \quad \text{as } T \to \infty.$$

Moreover, for any $T \leq t \leq 2T$ such that $|P_1(\sigma+it)| \leq \log\log T$ we have

$$\exp\{P_1(\sigma+it)\} \left(\sum_{0 \leq k \leq 100\log\log T} \frac{(-1)^k}{k!} P_1(\sigma+it)^k \right) = 1 + O(\frac{1}{\log^{99} T}),$$

and for any $T \leq t \leq 2T$ such that $|P_2(\sigma+it)| \leq \log\log\log T$ we have

$$\exp\{P_2(\sigma+it)\} \left(\sum_{0 \leq k \leq 100\log\log\log T} \frac{(-1)^k}{k!} P_2(\sigma+it)^k \right) = 1 + O(\frac{1}{(\log\log T)^{99}}).$$

Proof of Lemma 14.3. Using Chebychev's inequality we have $\mathbb{P}_T(|P_1(\sigma+it)| > \log\log T) \leq \frac{1}{(\log\log T)^2}\mathbb{E}_T|P_1(\sigma+it)|^2$, and using Lemma 13.2 with $A(s) = B(s) = P_1(s)$ the right hand side is

$$= \frac{1}{(\log\log T)^2} \left(\sum_{p^k \leq T^{1/(\log\log T)^2}} \frac{1}{k^2 p^{2k\sigma}} + O\left(\frac{T^{1/(\log\log T)^2}}{T} \left(\sum_{p^k \leq T^{1/(\log\log T)^2}} \frac{1}{kp^{k\sigma}} \right)^2 \right) \right).$$

The "big Oh" term here obviously tends to zero as $T \to \infty$, and since $\sigma \geq 1/2$ the first term is

$$\leq \frac{1}{(\log\log T)^2} \left(\sum_{p \leq T^{1/(\log\log T)^2}} \frac{1}{p} + O(1) \right) \leq \frac{1}{\log\log T},$$

using the estimate of Mertens (Fact 2 from Chapter 0). The proof that $\mathbb{P}_T(|P_2(\sigma+it)| > \log\log\log T) \to 0$ is exactly similar, the key point being that

$$\frac{1}{(\log\log\log T)^2} \sum_{T^{1/(\log\log T)^2} < p \leq T^{1/(\log\log\log T)^2}} \frac{1}{p} = \frac{2\log\log\log T - 2\log\log\log\log T + O(1)}{(\log\log\log T)^2},$$

again by Mertens' estimate.

To prove the final statements, we just note that if $|P| \leq K$ is any real number then

$$\begin{aligned} e^{-P} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} P^k &= \sum_{0 \leq k \leq 100K} \frac{(-1)^k}{k!} P^k + O(\frac{1}{(\lfloor 100K \rfloor + 1)!}|P|^{(\lfloor 100K \rfloor + 1)}) \\ &= \sum_{0 \leq k \leq 100K} \frac{(-1)^k}{k!} P^k + O(e^{-100K}), \end{aligned}$$

where the final step uses the general inequality $|P|^k/k! \leq (|P|e/k)^k$. □

Using Lemma 14.3, we can obtain our desired "simple" approximation to $\exp\{-\sum_{p \leq T^{1/\phi(T)}} \frac{1}{p^{\sigma+it}}\}$.

Proposition 14.4. *Let T be large, and suppose that $\sigma = \sigma(T) \geq 1/2$. For a complex number s, define $M(s) := \sum_n \frac{\mu(n)a(n)}{n^s}$, where $\mu(n)$ denotes the Möbius function, and where $a(n)$ denotes the characteristic function of the set of natural numbers n having at most $100 \log\log T$ prime factors less than $T^{1/(\log\log T)^2}$, and at most $100 \log\log\log T$ prime factors between $T^{1/(\log\log T)^2}$ and $T^{1/(\log\log\log T)^2}$, and no other prime factors.*

Then

$$\mathbb{P}_T\left(|M(\sigma+it)\exp\{\sum_{p^k \leq T^{1/(\log\log\log T)^2}} \frac{1}{kp^{k(\sigma+it)}}\} - 1| \geq \frac{1}{(\log\log T)^{20}} \right) \to 0 \quad \text{as } T \to \infty.$$

Proof of Proposition 14.4. Note first that since the function $\frac{\mu(n)}{n^s}$ is multiplicative, we can write $M(s) = \left(\sum_n \frac{\mu(n)a_1(n)}{n^s}\right)\left(\sum_n \frac{\mu(n)a_2(n)}{n^s}\right)$, where $a_1(n)$ is the characteristic function of those numbers having at most $100\log\log T$ prime factors less than $T^{1/(\log\log T)^2}$, and no other prime factors, and where $a_2(n)$ is the characteristic function of numbers having at most $100 \log\log\log T$ prime factors between $T^{1/(\log\log T)^2}$ and $T^{1/(\log\log\log T)^2}$, and no other prime factors. We will show that for most $T \leq t \leq 2T$ we have

$$\sum_n \frac{\mu(n)a_1(n)}{n^{\sigma+it}} \approx \sum_{0 \leq k \leq 100\log\log T} \frac{(-1)^k}{k!} P_1(\sigma+it)^k, \quad \sum_n \frac{\mu(n)a_2(n)}{n^{\sigma+it}} \approx \sum_{0 \leq k \leq 100\log\log\log T} \frac{(-1)^k}{k!} P_2(\sigma+it)^k,$$

which in view of Lemma 14.3 will prove the proposition.

Indeed, if we expand out $\sum_{0 \leq k \leq 100\log\log T} \frac{(-1)^k}{k!} P_1(\sigma+it)^k$ we will obtain a finite sum of the form $\sum_{n \leq T^{100/\log\log T}} \frac{b(n)}{n^{\sigma+it}}$, for certain coefficients $b(n)$. Note that $b(n)$ is a combinatorial coefficient that does not depend on σ. If n has at most $100\log\log T$ prime factors (counted with multiplicity), *and if all of its prime and prime power factors are* $\leq T^{1/(\log\log T)^2}$, then the coefficient $b(n)$ must be the same as the coefficient of $\frac{1}{n^s}$ in the infinite series (for $\Re(s) > 1$, say)

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{k!}\left(\sum_{p^j} \frac{1}{jp^{js}}\right)^k = \exp\{-\sum_{p^j} \frac{1}{jp^{js}}\} = \exp\{\sum_p \log(1-\frac{1}{p^s})\},$$

since none of the terms with $p^j > T^{1/(\log\log T)^2}$ or $k > 100\log\log T$ can possibly contribute to the coefficient of $\frac{1}{n^s}$ for such n. But in view of the Euler product expansion (Lemma 10.2) and also Lemma 12.2, we already know that for $\Re(s) > 1$ this series is $\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}$. Therefore we must have $b(n) = \mu(n)$ for such n. Let us also note, for use a bit later, than for all n we know that $|b(n)|$ is at most the coefficient of $\frac{1}{n^s}$ in the infinite series (for $\Re(s) > 1$, say)

$$\sum_{k=0}^{\infty} \frac{1}{k!}\left(\sum_{p^j} \frac{1}{jp^{js}}\right)^k = \exp\{\sum_{p^j} \frac{1}{jp^{js}}\} = \exp\{-\sum_p \log(1-\frac{1}{p^s})\} = \zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s},$$

so we always have $|b(n)| \leq 1$.

It follows that

$$\sum_n \frac{\mu(n)a_1(n)}{n^{\sigma+it}} - \sum_{0 \leq k \leq 100\log\log T} \frac{(-1)^k}{k!} P_1(\sigma+it)^k = \sum_{\substack{n \leq T^{100/\log\log T}, \\ n \text{ has } >100\log\log T \text{ prime factors}, \\ \text{or } n \text{ has a prime power factor } >T^{1/(\log\log T)^2}}} \frac{\mu(n)a_1(n) - b(n)}{n^{\sigma+it}}.$$

Using Chebychev's inequality and Lemma 13.2 as before, and the fact that $|\mu(n)a_1(n) - b(n)| \leq 1 + |b(n)| \leq 2$, we then get

$$\mathbb{P}_T\left(\left|\sum_n \frac{\mu(n)a_1(n)}{n^{\sigma+it}} - \sum_{0 \leq k \leq 100\log\log T} \frac{(-1)^k}{k!} P_1(\sigma+it)^k\right| \geq \frac{1}{\log^{25} T}\right)$$

$$\leq (\log^{50} T)\mathbb{E}_T\left|\sum_{\substack{n \leq T^{100/\log\log T}, \\ n \text{ has } >100\log\log T \text{ prime factors}, \\ \text{or } n \text{ has a prime power factor } >T^{1/(\log\log T)^2}}} \frac{\mu(n)a_1(n) - b(n)}{n^{\sigma+it}}\right|^2$$

$$\ll \log^{50} T\left(\sum_{\substack{n \leq T^{100/\log\log T}, \\ n \text{ has all prime factors } \leq T^{1/(\log\log T)^2}, \\ n \text{ has } >100\log\log T \text{ prime factors}, \\ \text{or } n \text{ has a prime power factor } >T^{1/(\log\log T)^2}}} \frac{1}{n^{2\sigma}} + O\left(\frac{T^{300/\log\log T}}{T}\right)\right).$$

The contribution from the "big Oh" term obviously tends to zero, and since $\sigma \geq 1/2$ the contribution from the first term is

$$\leq \log^{50} T\left(\sum_{p \leq T^{1/(\log\log T)^2}} \sum_{k: p^k > T^{1/(\log\log T)^2}} \frac{\log T}{p^k} + \sum_{\substack{n \leq T^{100/\log\log T}, \\ n \text{ has } >100\log\log T \text{ prime factors}}} \frac{1}{n}\right)$$

$$\ll \log^{50} T\left(\frac{\log T}{T^{1/(2(\log\log T)^2)}} + \sum_{\substack{n \leq T^{100/\log\log T}, \\ n \text{ has } >100\log\log T \text{ prime factors}}} \frac{1}{n}\right).$$

Again, the first term here clearly tends to zero. To estimate the contribution from the second sum, we note that if $r > 1$, and if $\Omega(n)$ denotes the total number of prime factors (counted with multiplicity) of n, then

$$\sum_{\substack{n \leq T, \\ n \text{ has } >100\log\log T \text{ prime factors}}} \frac{1}{n} \leq \frac{1}{r^{100\log\log T}} \sum_{n \leq T} \frac{r^{\Omega(n)}}{n} \leq \frac{1}{r^{100\log\log T}} \prod_{p \leq T}\left(1 + \frac{r}{p} + \sum_{2 \leq k \leq (\log T)/\log p} \frac{r^k}{p^k}\right).$$

Provided that $r \leq 1.99$, say, the sum over proper prime powers ($k \geq 2$) is always $O(1/p^2)$, and so by taking logarithms we find $\prod_{p \leq T}\left(1 + \frac{r}{p} + \sum_{2 \leq k \leq (\log T)/\log p} \frac{r^k}{p^k}\right) =$

$\exp\{\sum_{p \le T}(\frac{r}{p} + O(\frac{1}{p^2}))\} = \exp\{r \log\log T + O(1)\} \ll \log^r T$. This implies overall that

$$\log^{50} T \sum_{\substack{n \le T^{100/\log\log T}, \\ n \text{ has } > 100 \log\log T \text{ prime factors}}} \frac{1}{n} \ll \log^{50} T \frac{\log^r T}{r^{100 \log\log T}} = \log^{50} T \frac{\log^r T}{\log^{100 \log r} T}.$$

Finally, if we choose $r = 1.99$ then the right hand side is $\ll \frac{1}{\log^{10} T}$, say, which in particular tends to zero as $T \to \infty$.

We have just shown that

$$\mathbb{P}_T\left(\left|\sum_n \frac{\mu(n)a_1(n)}{n^{\sigma+it}} - \sum_{0 \le k \le 100 \log\log T} \frac{(-1)^k}{k!} P_1(\sigma+it)^k\right| \ge \frac{1}{\log^{25} T}\right) \to 0 \quad \text{as } T \to \infty,$$

and an exactly similar argument (applying Chebychev's inequality, and bounding the contribution from numbers with more than $100 \log\log\log T$ prime factors larger than $T^{1/(\log\log T)^2}$ and no other prime factors, and from numbers with a prime factor smaller than $T^{1/(\log\log T)^2}$ or a prime power factor larger than $T^{1/(\log\log\log T)^2}$) shows that

$$\mathbb{P}_T\left(\left|\sum_n \frac{\mu(n)a_2(n)}{n^{\sigma+it}} - \sum_{0 \le k \le 100 \log\log\log T} \frac{(-1)^k}{k!} P_2(\sigma+it)^k\right| \ge \frac{1}{(\log\log T)^{25}}\right) \to 0 \quad \text{as } T \to \infty.$$

Therefore, with probability tending to 1 as $T \to \infty$, we have

$$\begin{aligned} & M(\sigma+it) \exp\{\sum_{p^k \le T^{1/(\log\log\log T)^2}} \frac{1}{kp^{k(\sigma+it)}}\} \\ = & \left(\sum_n \frac{\mu(n)a_1(n)}{n^{\sigma+it}}\right) \exp\{P_1(\sigma+it)\} \left(\sum_n \frac{\mu(n)a_2(n)}{n^{\sigma+it}}\right) \exp\{P_2(\sigma+it)\} \\ = & \left(\sum_{0 \le k \le 100 \log\log T} \frac{(-1)^k}{k!} P_1(\sigma+it)^k + O\left(\frac{1}{\log^{25} T}\right)\right) \exp\{P_1(\sigma+it)\} \cdot \\ & \cdot \left(\sum_{0 \le k \le 100 \log\log\log T} \frac{(-1)^k}{k!} P_2(\sigma+it)^k + O\left(\frac{1}{(\log\log T)^{25}}\right)\right) \exp\{P_2(\sigma+it)\}. \end{aligned}$$

Using Lemma 14.3, we also know that with probability tending to 1 as $T \to \infty$ we have $|P_1(\sigma+it)| \le \log\log T$ and $|P_2(\sigma+it)| \le \log\log\log T$. Then $1/\log T \le |\exp\{P_1(\sigma+it)\}| \le \log T$ and $1/\log\log T \le |\exp\{P_2(\sigma+it)\}| \le \log\log T$, and then by Lemma 14.3 the above is

$$\begin{aligned} & \left(\exp\{-P_1(\sigma+it)\} + O\left(\frac{1}{\log^{25} T}\right)\right) \exp\{P_1(\sigma+it)\} \cdot \\ & \cdot \left(\exp\{-P_2(\sigma+it)\} + O\left(\frac{1}{(\log\log T)^{25}}\right)\right) \exp\{P_2(\sigma+it)\} = 1 + O\left(\frac{1}{(\log\log T)^{24}}\right). \end{aligned}$$

This finishes the proof of the proposition. □

[[We ran out of time to cover more of the Selberg central limit theorem in the course.]]

We have shown that for most $T \le t \le 2T$ we have $M(\sigma+it)\exp\{\sum_{p^k \le T^{1/(\log\log\log T)^2}} \frac{1}{kp^{k(\sigma+it)}}\} \approx 1$, and we already know from Proposition 13.4 that $\Re(\sum_{p^k \le T^{1/(\log\log\log T)^2}} \frac{1}{kp^{k(\sigma+it)}})$ has an approximately Gaussian distribution as $T \le t \le 2T$ varies (apart from the proper prime power contribution, but this is negligible). There are two remaining steps in proving the Selberg central limit theorem:

- Show that for most $T \le t \le 2T$ we have $M(\sigma+it)\zeta(\sigma+it) \approx 1$, and therefore for most $T \le t \le 2T$ we must have $\zeta(\sigma+it) \approx \exp\{\sum_{p^k \le T^{1/(\log\log\log T)^2}} \frac{1}{kp^{k(\sigma+it)}}\}$. *This says that the Euler product expression for the zeta function remains approximately valid on a wide range of σ and t.* Note that the coefficients of $M(\sigma+it)$ are close to the Möbius function (on a certain range), so the fact that $M(\sigma+it)\zeta(\sigma+it) \approx 1$ is not so surprising. However, *one can only show that $M(\sigma+it)\zeta(\sigma+it) \approx 1$ when σ is a bit larger than 1/2.*
- Show the "continuity statement" that knowing the Selberg central limit theorem for suitable $\sigma = 1/2 + o(1)$ implies it when $\sigma = 1/2$.

The first step can be performed using another Chebychev type calculation, somewhat like the second moment calculation for the zeta function that you will perform in question 6 on Example Sheet 3. The second step is different, and requires a bit of complex analysis trickery and some (general and standard) information about the zeros of the zeta function. Notice that this is the only step in the proof that involves such manipulations, everything else having been done on the level of sums over primes that were (hopefully) somewhat intuitive. This nice approach to proving Selberg's central limit theorem is due to Radziwiłł and Soundararajan [1], whose paper may be consulted for further details.

References

[1] M. Radziwiłł, K. Soundararajan. Selberg's central limit theorem for $\log|\zeta(1/2+it)|$. *Preprint*, available online at `http://arxiv.org/abs/1509.06827`

Jesus College, Cambridge, CB5 8BL
E-mail address: `A.J.Harper@dpmms.cam.ac.uk`