UNIVERSITY OF CAMBRIDGE

MATHEMATICAL TRIPOS Part III

Monday, 6 June, 2022 1:30 pm to 4:30 pm

PAPER 117

FIVE WAYS TO THINK ABOUT PRIMES

Before you begin please read these instructions carefully

Candidates have THREE HOURS to complete the written examination.

Attempt no more than **THREE** questions.
There are **FOUR** questions in total.
The questions carry equal weight.

STATIONERY REQUIREMENTS
Cover sheet
Treasury tag
Script paper
Rough paper

SPECIAL REQUIREMENTS
None

You may not start to read the questions printed on the subsequent pages until instructed to do so by the Invigilator.

1 Let $\mathcal{P}$ be a set of primes, $z \geqslant 3$, and

$$P_z := \prod_{\substack{p \in \mathcal{P} \\ p < z}} p.$$

Let $\mathcal{A} = (a_n)_{n=1}^{\infty}$ be a sequence of non-negative real numbers.

(a) Define $|\mathcal{A}_d|$. What does it mean for $\mathcal{A}$ to have distribution $(X, P_z, g, (r_d))$?

(b) Define $S(\mathcal{A}, \mathcal{P}, z)$ in terms of the values a_n. State the general Selberg sieve upper bound for $S(\mathcal{A}, \mathcal{P}, z)$ when the sequence $\mathcal{A}$ has distribution $(X, P_z, g, (r_d))$, using sieve weights with level of support D. [You need not provide the explicit form of the sieve weights themselves.]

(c) Suppose

$$a_n = \begin{cases} 1 & \text{if } n = m(m+2)(m+6) \text{ for some } m \in \mathbb{N} \text{ with } m \leqslant X \\ 0 & \text{otherwise.} \end{cases}.$$

Find a suitable g for which $\mathcal{A}$ has distribution $(X, P_z, g, (r_d))$, and provide an upper bound on the remainders r_d. Your bound should involve the quantity $\omega(d)$, which denotes the number of distinct prime divisors of d counted without multiplicity.

(d) Prove that there is an absolute constant $c > 0$ such that, for any parameters z and D satisfying $3 \leqslant z \leqslant D$,

$$\sum_{p \leqslant z} \frac{1}{p(\log p)^3} \exp\Big(-\frac{1}{2}\frac{\log D}{\log p}\Big) \ll \frac{1}{(\log D)^3} \exp\Big(-c\frac{\log D}{\log z}\Big).$$

[You may assume Chebyshev's estimates or the Prime Number Theorem without proof.]

(e) By using $\mathcal{A}$ from part (c), choosing a suitable set of primes $\mathcal{P}$, choosing suitable values for z and D, and using a form of Buchstab's identity, use the Selberg sieve to prove that there is an absolute constant C for which there are infinitely many $m \in \mathbb{N}$ with

$$\omega(m) + \omega(m+2) + \omega(m+6) \leqslant C.$$

[You may assume Mertens' estimates without proof, and the bound $k^{\omega(n)} \ll_{k,\varepsilon} n^{\varepsilon}$, valid for all $k \in \mathbb{N}$ and $\varepsilon > 0$.]

(f) Describe very briefly any changes that are required in order to adapt your method from the previous parts to prove the following: there is an absolute constant C for which there are infinitely many $m \in \mathbb{N}$ with

$$\omega(m) + \omega(m+1) + \omega(m+2) \leqslant C.$$

2 [Throughout this question, you do not need to formally justify any valid application of Fubini's theorem.]

(a) Let $F : \mathbb{R}_{>0} \longrightarrow \mathbb{R}$ be a smooth compactly supported function.

Define the *Mellin transform* $\widetilde{F}$, giving the range where $\widetilde{F}$ is well-defined. State the *Mellin inversion formula* for F and its range of validity.

Suppose in addition that $|\widetilde{F}(\sigma + it)| = O(\exp(-|t|^{1/2}))$ for all $\sigma \in (0,2)$ and for all $t \in \mathbb{R}$. Sketch a proof that there is some absolute constant $c > 0$ for which, for all $X \geqslant 3$,

$$\sum_{n \geqslant 1} \Lambda(n) F\Big(\frac{n}{X}\Big) = C_{F,X} + O(X^{1-\frac{c}{\log\log X}}),$$

where $C_{F,X}$ is some explicit quantity depending on X and F (which you should determine).

[Any results from the course on the size of $|\zeta(s)|$, $|\zeta'(s)|$, $|\zeta(s)|^{-1}$, on properties of Dirichlet series, and about the location of the zeros and poles of $\zeta(s)$ may be assumed without proof, provided they are clearly stated when used. You may not assume the Prime Number Theorem.]

(b) Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be a smooth function supported on $[-1,1]$, with $f(0) = 1$, and let

$$g(t) := \int_{-\infty}^{\infty} e^x f(x) e(-tx)\, dx.$$

Let $Y \geqslant X \geqslant 2$, and fix $D = X^{1/10}$.

Show that

$$\sum_{Y < n \leqslant Y+X} \Big(\sum_{d|n} \mu(d) f\Big(\frac{\log d}{\log D}\Big)\Big)^2 \geqslant \pi(Y+X) - \pi(Y).$$

By expanding the square and using the Fourier inversion formula, show that

$$\sum_{Y < n \leqslant Y+X} \Big(\sum_{d|n} \mu(d) f\Big(\frac{\log d}{\log D}\Big)\Big)^2 = X \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(t_1) g(t_2) H(t_1,t_2,D)\, dt_1\, dt_2 + O(D^2), \quad (\dagger)$$

where $H(t_1,t_2,D)$ is an absolutely uniformly convergent infinite product over primes that you should define. [You need not explain why this infinite product converges.]

Assuming that for all $t_1, t_2 \in \mathbb{R}$

$$|H(t_1,t_2,D)| \ll \Big|\zeta\Big(1 + \frac{1-2\pi i t_1}{\log D}\Big)\Big|^{-1} \Big|\zeta\Big(1 + \frac{1-2\pi i t_2}{\log D}\Big)\Big|^{-1} \Big|\zeta\Big(1 + \frac{2-2\pi i t_1 - 2\pi i t_2}{\log D}\Big)\Big|,$$

by truncating the range of integration in $(\dagger)$ deduce that $\pi(Y+X) - \pi(Y) \ll \frac{X}{\log X}$. [You may not appeal to the usual Selberg sieve, but may use any results from the course on $\zeta(s)$ and on compactly supported smooth functions, provided they are clearly stated when used.]

3 (a) State Vaughan's identity for the von Mangoldt function $\Lambda(n)$ with thresholds U and V. If $UV \leqslant N$, give the corresponding Vaughan's identity expansion of the sum

$$\sum_{n \leqslant N} \Lambda(n) e(\alpha n^2).$$

For the rest of this question, let $\alpha \in \mathbb{R}$, and suppose $a, q \in \mathbb{N}$ with $q \geqslant 2$, $a \leqslant q$, $\gcd(a,q) = 1$, and $|\alpha - \frac{a}{q}| \leqslant \frac{1}{q^2}$.

(b) Show that for all $M, R \geqslant 2$ and $m_0 \in \mathbb{R}$,

$$\sum_{m_0 \leqslant m < m_0 + M} \min(R, \|\alpha m\|^{-1}) \ll (\log q)\Big(\frac{MR}{q} + M + R + q\Big),$$

where $\|\alpha m\|$ denotes the distance from αm to the nearest integer.

(c) Let $R, T \geqslant 2$, and let $(b_t)_{T < t \leqslant 2T}$ and $(c_r)_{R < r \leqslant 2R}$ be arbitrary complex coefficients. Show that

$$\Big| \sum_{\substack{T < t \leqslant 2T \\ R < r \leqslant 2R}} b_t c_r e(\alpha r^2 t^2) \Big| \ll \|b\|_2 \|c\|_2 \Big(\sum_{\substack{T < t_1, t_2 \leqslant 2T \\ R < r_1, r_2 \leqslant 2R}} e(\alpha (t_1^2 - t_2^2)(r_1^2 - r_2^2)) \Big)^{1/4}, \qquad (*)$$

where $\|b\|_2 := \Big(\sum_{T < t \leqslant 2T} |b_t|^2 \Big)^{1/2}$ and $\|c\|_2 := \Big(\sum_{R < r \leqslant 2R} |c_r|^2 \Big)^{1/2}$.

(d) Make the substitution $s_1 = r_1 - r_2$, $s_2 = r_1 + r_2$, $u_1 = t_1 - t_2$, $u_2 = t_1 + t_2$. By considering an inner sum over s_1, or otherwise, establish that there is some absolute constant $C > 0$ for which the left-hand side of $(*)$ is

$$\ll \|b\|_2 \|c\|_2 \Big(R^{1/4} T^{1/2} + R^{1/2} T^{1/4} + \Big(\sum_{m \leqslant CRT^2} \tau_4(m) \min(R, \|\alpha m\|^{-1}) \Big)^{1/4} \Big),$$

where $\tau_4 = 1 \star 1 \star 1 \star 1$. [You may use the standard upper bound on the absolute value of sums of the form $\sum_{n \in I} e(\beta n)$ for an interval I without proof, provided this bound is clearly stated when used.]

(e) By splitting into cases according to whether $\tau_4(n) < X$ or $\tau_4(n) \geqslant X$, show that for all $X \geqslant 1$ the left-hand side of $(*)$ is

$$\ll \|b\|_2 \|c\|_2 (\log q)(\log RT)^{O(1)} R^{1/2} T^{1/2} \Big(\frac{1}{X^{1/4}} + \frac{X^{1/4}}{q^{1/4}} + \frac{X^{1/4}}{R^{1/4}} + \frac{X^{1/4} q^{1/4}}{R^{1/2} T^{1/2}} + \frac{1}{T^{1/4}} \Big). \quad (\dagger)$$

[You may use without proof the estimate $\sum_{n \leqslant N} \tau_4(n)^2 \ll N (\log N)^{O(1)}$, valid for all $N \geqslant 2$.]

(f) Suppose $RT = N$. Discuss the ranges of R, T, q, X for which $(\dagger)$ is smaller than the trivial bound $\|b\|_2 \|c\|_2 N^{1/2}$ by a large power of $\log N$.

4 Let $N \in \mathbb{N}$, and let $\mathcal{H}_N$ be the set of functions $g : [N] \longrightarrow \mathbb{R}$.

(a) Let $\varepsilon > 0$. Suppose that $g_1, g_2 \in \mathcal{H}_N$ with

$$\sup_{\alpha \in [0,1]} \Big| \sum_{n \leqslant N} g_1(n) e(\alpha n) - \sum_{n \leqslant N} g_2(n) e(\alpha n) \Big| \leqslant \varepsilon N.$$

Suppose moreover that for $i = 1, 2$,

$$\int_0^1 \Big| \sum_{n \leqslant N} g_i(n) e(\alpha n) \Big|^3 d\alpha \ll N^2.$$

Defining

$$T(g_i, g_i, g_i, g_i) := \sum_{\substack{n_1, n_2, n_3, n_4 \leqslant N \\ n_1 + 2n_2 + 3n_3 = 4n_4}} g_i(n_1) g_i(n_2) g_i(n_3) g_i(n_4),$$

prove that $|T(g_1, g_1, g_1, g_1) - T(g_2, g_2, g_2, g_2)| = O(\varepsilon N^3)$. [If your proof splits into several similar cases, you need only present the details of one case. Any standard inequalities may be used without proof.]

Let $\mathcal{F} \subset \mathcal{H}_N$ be a set of functions of the form $f : [N] \longrightarrow [-1, 1]$.

(b) For $g_1, g_2 \in \mathcal{H}_N$ define the inner-product $\langle g_1, g_2 \rangle$. For $g \in \mathcal{H}_N$, define the quantities $\|g\|_{\mathcal{F}}$ and $\|g\|_{\mathcal{F}}^*$.

(c) Suppose that $\|g_1 g_2\|_{\mathcal{F}}^* \leqslant \|g_1\|_{\mathcal{F}}^* \|g_2\|_{\mathcal{F}}^*$ for all $g_1, g_2 \in \mathcal{H}_N$. Suppose further that $\mathcal{F}$ is closed and convex, the constant 1 function is in $\mathcal{F}$, and that $f \in \mathcal{F}$ if and only if $-f \in \mathcal{F}$.

Prove that there is some absolute constant C (independent of N) for which the following holds: for all $\varepsilon \in (0, 1/2]$, and for all $\nu \in \mathcal{H}_N$ with $\nu \geqslant 0$, $\sum_{n \leqslant N} \nu(n) \leqslant N$, and

$$\|\nu - 1\|_{\mathcal{F}} \leqslant \exp(-\varepsilon^{-C}),$$

if $g_1 \in \mathcal{H}_N$ with $0 \leqslant g_1 \leqslant \nu$, then there exists $g_2 \in \mathcal{H}_N$ with $0 \leqslant g_2 \leqslant 1$ and

$$\|g_1 - g_2\|_{\mathcal{F}} \leqslant \varepsilon.$$

[You may assume that there is some absolute constant $C > 0$ with the following property: for all $\varepsilon \in [0, 1/2)$, there exists a polynomial $P_\varepsilon(T) := \sum_{i=0}^m a_i T_i$ with real coefficients such that

$$(m+1)\varepsilon^{-m} \max_i |a_i| \leqslant \exp(\varepsilon^{-C})$$

and

$$\sup_{t \in [-\frac{10}{\varepsilon}, \frac{10}{\varepsilon}]} |P_\varepsilon(t) - \max(0, t)| \leqslant \frac{\varepsilon}{100}.$$

You may also assume that if $A, B \subset \mathbb{R}^N$ then $A + B$ is closed and convex provided A is closed, bounded, and convex, and B is closed and convex. If you appeal to a version of the separating hyperplane theorem, you should clearly state the theorem. If you wish to use the dense model theorem, however, you should first prove it.]

END OF PAPER