A NOTE ON THE NUMBER OF PRIMES IN SHORT INTERVALS

D. A. GOLDSTON AND S. M. GONEK

(Communicated by William Adams)

Abstract. Let \( J(\beta, T) = \int_1^T (\sum_{p^k \leq x/T} \log p - x/T)^2 \frac{dx}{x^2} \), where the sum is over prime powers. H. L. Montgomery has shown that on the Riemann hypothesis, there is a positive constant \( C_0 \) such that for each \( \beta \geq 1 \), \( J(\beta, T) \leq C_0 \beta \log^2 T/T \), provided that \( T \) is sufficiently large. Here we prove a slightly stronger result from which we deduce a lower bound of the same order.

1. Introduction

In 1943 A. Selberg [7] proved that if the Riemann hypothesis (RH) is true, then
\[
J(\beta, T) = \int_1^T \left( \psi \left( x + \frac{x}{T} \right) - \psi(x) - \frac{x}{T} \right)^2 x^{-2} dx
\ll \beta \frac{\log^2 T}{T}
\]
for fixed \( \beta \geq 1 \) and \( T \geq 2 \); here \( \psi(x) = \sum_{n \leq x} \Lambda(n) \), where \( \Lambda(n) = \log p \) if \( n = p^m \) with \( p \) a prime number and \( m \geq 1 \), and \( \Lambda(n) = 0 \) otherwise. H. L. Montgomery (unpublished) later made the \( \beta \) dependence explicit by proving that on RH there exists an absolute constant \( C_0 \) such that, for each \( \beta \geq 1 \),
\[
(1) \quad J(\beta, T) \leq C_0 \frac{\beta \log^2 T}{T}
\]
as \( T \to \infty \). Proofs of this subsequently appeared in [1], [5], and [4]. Our object here is to prove a stronger result for \( J(\beta, T) \) on RH which immediately implies (1) and, moreover, shows that apart from constants (1) is best possible.

We shall use the standard symbols \( \ll, \gg, O, o, \) and \( \sim \) and, unless otherwise indicated, all implied constants will be absolute.

Theorem. Assume the Riemann hypothesis. Then there are absolute constants \( C_2 > C_1 > 0 \) such that for each \( \beta > 0 \),
\[
C_1 \frac{\log^2 T}{T} \leq J(\beta + 2, T) - J(\beta, T) \leq C_2 \frac{\log^2 T}{T}
\]
for all sufficiently large \( T \).
Corollary. Assume the Riemann hypothesis. Then there are absolute constants $D_2 > D_1 > 0$ such that, for each $\beta \geq 1$,

$$D_1 \frac{\beta \log^2 T}{T} \leq J(\beta, T) \leq D_2 \frac{\beta \log^2 T}{T}$$

for all sufficiently large $T$.

The Theorem should be compared with a result of Gallagher and Mueller [1] (also see [3]) which asserts that RH and the pair correlation conjecture together imply that for fixed $\beta_1 > \beta_0 \geq 1$,

$$J(\beta_1, T) - J(\beta_0, T) = ((\beta_1 - \beta_0) + o(1)) \frac{\log^2 T}{T} \quad (\text{as } T \to \infty).$$

Since for $0 < \beta \leq 1$ one also has (unconditionally) that

$$J(\beta, T) \sim \frac{\beta^2 \log^2 T}{2} \quad (\text{as } T \to \infty)$$

(see [1]), we see that on the above hypotheses

$$J(\beta + 2, T) - J(\beta, T) \sim \begin{cases} (3/2 + \beta - \beta^2/2) \frac{\log^2 T}{T} & \text{if } 0 < \beta \leq 1, \\ 2 \frac{\log^2 T}{T} & \text{if } \beta \geq 1. \end{cases}$$

Our proof will actually show that if $\beta > 0$, then

$$\frac{3 \log^2 T}{T} \leq J(\beta + 2, T) - J(\beta, T) \leq 21.65 \frac{\log^2 T}{T}$$

for all sufficiently large $T$. It is also possible by our method to show that

$$(\beta_1 - \beta_0) \frac{\log^2 T}{T} \ll J(\beta_1, T) - J(\beta_0, T) \ll (\beta_1 - \beta_0) \frac{\log^2 T}{T}$$

for $\beta_1 > \beta_0 > 0$ as long as $\beta_1 - \beta_0 > 6 - 2\sqrt{6} = 1.10102\ldots$. It is doubtful, however, whether one can obtain this for arbitrarily small differences $\beta_1 - \beta_0$ on RH alone.

2. A LEMMA

We prove the Theorem by relating $J(\beta, T)$ to averages of the function

$$F(\alpha, T) = \left(\frac{T}{2\pi} \log T\right)^{-1} \sum_{0 < \gamma, \gamma' \leq T} T^{i \alpha (\gamma - \gamma')} w(\gamma - \gamma')$$

introduced by Montgomery [6]; here $\alpha$ is real, $T \geq 2$, $w(u) = 4/4 + u^2$, and $\gamma, \gamma'$ denote the imaginary parts of zeros of the Riemann zeta-function. We shall then require the following result which generalizes and strengthens Lemma A of [2].
Lemma. Assume the Riemann hypothesis and let

\[ G(\alpha , T) = \left( \frac{T}{2\pi \log T} \right)^{-1} \sum_{0 < \gamma , \gamma' \leq T} \left( \frac{\sin \frac{a}{2}(\gamma - \gamma')}{\frac{a}{2}(\gamma - \gamma')} \log T \right)^2 w(\gamma - \gamma') . \]

Then for \( a > 0 \), \( \beta \) real, and \( T \geq 2 \),

\[ a \left( 1 - \frac{1}{2}G \left( \frac{a}{2}, T \right) \right) \leq \int_{\beta}^{\beta + a} F(\alpha , T) \, d\alpha \leq a \left( G(a , T) + \frac{1}{2}G \left( \frac{a}{2}, T \right) \right) . \]

Proof. The proof of the lower bound in Lemma A (which corresponds to \( a = 2 \) here) extends in a straightforward way to give the lower bound in (3).

On the other hand, the upper bound in Lemma A generalizes to \( 2aG(a , T) \) which is not as good as the bound in (3).

To obtain the present upper bound define \( K_b(u) = \max(1 - |u|/b, 0), b > 0 \), and consider the function

\[ R_a(u) = K_a(u) + \frac{1}{2}K_a/2(u - a/2) + \frac{1}{2}K_a/2(u + a/2) . \]

Defining the Fourier transform of \( f(x) \) by

\[ \hat{f}(\omega) = \int_{-\infty}^{\infty} f(x)e(-x\omega) \, dx , \]

where \( e(u) = e^{2\pi iu} \), we have that

\[ \hat{K}_b(\omega) = b \left( \frac{\sin \pi b\omega}{\pi b\omega} \right)^2 . \]

Thus

\[ \hat{R}_a(\omega) = a \left\{ \left( \frac{\sin \pi a\omega}{\pi a\omega} \right)^2 + \frac{1}{2} \cos \pi a\omega \left( \frac{\sin \pi a\omega/2}{\pi a\omega/2} \right)^2 \right\} . \]

Now clearly \( R_a(u) \geq 0 \) for all \( u \), and \( R_a(u) = 1 \) for \( |u| \leq a/2 \). Furthermore (see [5]), \( F(\alpha , T) \geq 0 \). Hence

\[ \int_{\beta}^{\beta + a} F(\alpha , T) \, d\alpha = \int_{(\beta + a/2) - a/2}^{(\beta + a/2) + a/2} F(\alpha , T) \, d\alpha \]

\[ \leq \int_{-\infty}^{\infty} F(\alpha , T)R_a(\alpha - (\beta + a/2)) \, d\alpha \]

\[ = \left( \frac{T}{2\pi \log T} \right)^{-1} \sum_{0 < \gamma , \gamma' \leq T} T^{i(\beta + a/2)(\gamma - \gamma')} \hat{R}_a \left( \frac{(\gamma - \gamma') \log T}{2\pi} \right) \times w(\gamma - \gamma') \]

\[ \leq a \left( G(a , T) + \frac{1}{2}G(a/2 , T) \right) . \]

This proves the result.
We require the lower bound of the lemma for \( a = 2 - \eta \) and the upper bound for \( a = 2 + \eta \), where \( \eta > 0 \). Montgomery [6] has shown that on RH,
\[
G(\alpha, T) \sim \left( \frac{1}{\alpha} + \frac{\alpha}{3} \right)
\]
for \( 0 < \alpha \leq 1 \) as \( T \to \infty \). Hence
\[
\int_\beta^{\beta+2-\eta} F(\alpha, T) d\alpha \geq (1 + o(1)) \left( \frac{2}{3} - c_1(\eta) \right) \quad \text{(as } T \to \infty \text{)},
\]
where \( c_1(\eta) \to 0 \) as \( \eta \to 0^+ \). For the upper bound we use the inequality
\[
G(j + \eta, T) \leq 4/3 + c_2(\eta) + o(1)
\]
as \( T \to \infty \) (\( j = 1 \) or 2), where \( c_2(\eta) \to 0 \) as \( \eta \to 0^+ \); these follow from Lemma 7 of [2]. We then obtain
\[
\int_\beta^{\beta+2+\eta} F(\alpha, T) d\alpha \leq (1 + o(1))(4 + c_3(\eta)) \quad \text{(as } T \to \infty \text{)},
\]
where \( c_3(\eta) \to 0 \) as \( \eta \to 0^+ \).

3. PROOF OF THE THEOREM AND COROLLARY

We begin by quoting two results from Goldston [3]. We remind the reader that the Riemann hypothesis is assumed throughout this section.

Let \( g(x) \) be a complex valued function such that \( g(x) \ll (1+x^2)^{-1} \), \( \tilde{g}(\omega) \ll (1 + \omega^2)^{-1} \), and \( \tilde{g}(\omega) = 0 \) for \( \omega \leq 0 \), and define
\[
H_{\pm}(\mu, U) = \int_0^U \left| \sum_{\gamma} g(\pm(t-\gamma)\mu) \right|^2 dt,
\]
where the sum is over the ordinates of the nontrivial zeros of \( \zeta(s) \). Then by Equations (5.2) and (5.3) of [3],
\[
H_{\pm}(\mu, U) = U \int_1^\infty F(\alpha, U)|\tilde{g}(\alpha)|^2 d\alpha + o(U)
\]
for \( |\mu - 1/2\pi \log U| \leq C \log \log U \), where \( C \) is an arbitrary positive constant. Furthermore, by Lemma 5 of [3]
\[
\int_0^\infty (\psi(e^{u+\delta}) - \psi(e^u) - (e^{\delta} - 1)e^u)^2 e^{-u} \left| \tilde{g} \left( \frac{u}{2\pi \nu} \right) \right|^2 du
\]
\[
= 8\pi \nu^2 \int_{-\infty}^{\infty} \left( \frac{\sin \frac{\nu}{2} t}{t} \right)^2 \left| \sum_{\gamma} g((t-\gamma)\nu) \right|^2 dt
\]
\[
+ O(\delta) + O(\nu^2 \delta^{3/2} \log 1/\delta)
\]
\[1\] Lemma 5 contains two typographical errors: " \( T > 2 \) " should be removed and the factor \( (\psi(e^{u+\delta}) - \psi(e^u) - (e^{\delta} - 1)e^u) \) in (4.2) should be squared.
uniformly for \( \nu \geq 1 \) and \( 0 < \delta \leq 1/4 \). If we set \( e^\delta = 1 + 1/T \), \( \nu = 1/2\pi \log T \), and \( x = e^\nu \), we may rewrite this as

\[
\int_1^\infty \left( \psi \left( x + \frac{x}{T} \right) - \psi(x) - \frac{x}{T} \right)^2 \left| \frac{\log x}{\log T} \right|^2 \frac{dx}{x^2} = \frac{2}{\pi} \log^2 T \int_0^\infty \left( \sin \frac{\delta t}{t} \right)^2 \left( \sum g \left( (t - \gamma) \frac{\log T}{2\pi} \right) \right)^2 + \sum g \left( (\gamma - t) \frac{\log T}{2\pi} \right)^2 \right) dt + O \left( \frac{1}{T} \right).
\]

We now choose a pair of functions \( g \), \( \hat{g} \) for which the above conditions hold and such that \( \hat{g} \) approximates the characteristic function of the interval \([\beta, \beta + b]\) from below, with \( \beta > 0 \). More specifically, we take \( \hat{g} = 0 \) off of \([\beta, \beta + b]\), \( \hat{g} = 1 \) on \([\beta + \eta, \beta + b - \eta]\), and \( |\hat{g}| \leq 1 \) otherwise, where \( 0 < \eta < b/2 \) is fixed. (Such a pair may be constructed explicitly by a linear change of variable from the pair defined in (3.6) of [3].) With this choice of \( \hat{g} \) in (7) we immediately obtain

\[
U \int_{\beta + \eta}^{\beta + b - \eta} \int F(\alpha, U) d\alpha + o(U) \leq H(\mu, U) \leq U \int_{\beta}^{\beta + b} F(\alpha, U) d\alpha + o(U)
\]

for \( |\mu - 1/2\pi \log U| \leq C \log \log U \). The same choice in the left-hand side of (8) leads to the inequalities

\[
J(\beta + b - \eta, T) - J(\beta + \eta, T) \leq \int_1^\infty \left( \psi \left( x \right) - \psi(x) - \frac{x}{T} \right)^2 \left| \frac{\log x}{\log T} \right|^2 \frac{dx}{x^2} \leq J(\beta + b, T) - J(\beta, T).
\]

We now obtain the lower bound of the theorem. Taking \( b = 2 \) in (10) and using (8), we see that

\[
J(\beta + 2, T) - J(\beta, T) \geq \frac{2}{\pi} \log^2 T \int_0^{\theta T} \left( \sin \frac{\delta t}{t} \right)^2 \left\{ \sum g \left( (t - \gamma) \frac{\log T}{2\pi} \right) \right\} dt + O \left( \frac{1}{T} \right),
\]

where \( 0 < \theta < \pi \). Now \((\sin(\delta t/2)/t)^2\) is monotone decreasing for \( 0 \leq t \leq \theta T \), so by (6) this is

\[
\geq \frac{2}{\pi} \log^2 T \left( \sin \frac{\delta \theta T}{\theta T} \right)^2 \left\{ H_+ \left( \frac{\log T}{2\pi}, \theta T \right) + H_- \left( \frac{\log T}{2\pi}, \theta T \right) \right\} + O \left( \frac{1}{T} \right).
\]
Next, using the lower bound in (9), we find that this is
\[ \geq \frac{4}{\pi} \frac{\sin \frac{\delta T}{2}}{\theta^2} \log^2 \frac{T}{\theta} \int_{\beta + \eta}^{\beta + 2 - \eta} F(\alpha, \theta T) \, d\alpha + o \left( \frac{\log^2 T}{T} \right). \]

Finally, by (4) we have that
\[ J(\beta + 2, T) - J(\beta, T) \geq \frac{4}{\pi} \frac{\sin \frac{\theta}{2}}{\theta^2} \log^2 \frac{T}{\theta} (2/3 - c_1(\eta)) + o \left( \frac{\log^2 T}{T} \right). \]

The optimal choice of \( \theta \) is the unique solution (on \( (0, \pi) \)) of the equation \( \tan \theta/2 = \theta \), namely \( \theta = 2.33112 \ldots \). Using this and taking \( \eta \) sufficiently small, we obtain
\[ J(\beta + 2, T) - J(\beta, T) \geq (0.307 + o(1)) \frac{\log^2 T}{T}, \]

for \( \beta > 0 \) and all \( T \) sufficiently large.

To obtain the upper bound we again take \( g \) and \( \hat{g} \) as above (although \( b \) will be different). By the growth condition on \( g \) and the estimate \( \sum_{\nu - 1 < \gamma \leq \nu} 1 \ll \log(|\nu| + 2) \), we easily obtain the bound
\[ \left| \sum_{\gamma} g \left( \pm (t - \gamma) \frac{\log T}{2\pi} \right) \right| \ll \log(|t| + 2). \]

Using this, we find that
\[ \int_0^\infty \left( \sin \frac{4t}{\log^2 \frac{T}{2\pi}} \right)^2 \left| \sum_{\gamma} g \left( \pm (t - \gamma) \frac{\log T}{2\pi} \right) \right|^2 \, dt \]
\[ = \int_0^{T \log^3 T} \left( \sin \frac{4t}{\log^2 \frac{T}{2\pi}} \right)^2 \left| \sum_{\gamma} g \left( \pm (t - \gamma) \frac{\log T}{2\pi} \right) \right|^2 \, dt \]
\[ + o \left( \int_0^{\infty} \frac{\log^2 t}{t^2} \, dt \right) \]
\[ \leq \frac{\delta^2}{4} \int_0^T \left| \sum_{\gamma} g \left( \pm (t - \gamma) \frac{\log T}{2\pi} \right) \right|^2 \, dt \]
\[ + \sum_{k=1}^{[5 \log \log T]} \int_{2^{k-1} T}^{2^k T} \left| \sum_{\gamma} g \left( \pm (t - \gamma) \frac{\log T}{2\pi} \right) \right|^2 t^{-2} \, dt + o \left( \frac{1}{T} \right) \]
\[ \leq \frac{\delta^2}{4} H_\pm \left( \frac{\log T}{2\pi}, T \right) + T^{-2} \sum_{k=1}^{[5 \log \log T]} 2^{2-2k} H_\pm \left( \frac{\log T}{2\pi}, 2^k T \right) \]
\[ + o \left( \frac{1}{T} \right). \]
The bound for $H_\pm$ from (9) is applicable in this range of $k$ and leads to
\[
(1 + o(1)) \left\{ \frac{1}{4T} \int_\beta^{\beta+b} F(\alpha, T) \, d\alpha + \frac{4}{T} \sum_{k=1}^{[\log \log T]} 2^{-k} \int_\beta^{\beta+b} F(\alpha, 2^k T) \, d\alpha \right\} + o \left( \frac{1}{T} \right).
\]

If we now set $b = 2 + 2\eta$ and use (5) we find that this is
\[
\leq (1 + o(1)) \left\{ \frac{4 + c_3(2\eta)}{4T} + \frac{4(4 + c_3(2\eta))}{T} \right\} + o \left( \frac{1}{T} \right)
\leq \frac{17 + 5c_3(2\eta) + o(1)}{T}
\]
as $T \to \infty$. Thus, we have shown that
\[
\int_0^\infty \left( \frac{\sin \frac{\theta}{2} t}{t} \right)^2 \left| \sum_{\gamma} g \left( \pm(t - \gamma) \frac{\log T}{2\pi} \right) \right|^2 \, dt
\leq \frac{17 + 5c_3(2\eta) + o(1)}{T}
\]
as $T \to \infty$.

We now combine (8), (10) (with $b = 2 + 2\eta$), and (11) to obtain
\[
J(\beta + 2 + \eta, T) - J(\beta + \eta, T)
\leq \frac{2}{\pi} \log^2 T \int_0^\infty \left( \frac{\sin \frac{\theta}{2} t}{t} \right)^2
\times \left\{ \left| \sum_{\gamma} g \left( (t - \gamma) \frac{\log T}{2\pi} \right) \right|^2 + \left| \sum_{\gamma} g \left( (\gamma - t) \frac{\log T}{2\pi} \right) \right|^2 \right\} \, dt
\leq \frac{4}{\pi} (17 + 5c_3(2\eta) + o(1)) \frac{\log^2 T}{T}
\leq (21.646 + 7c_3(2\eta) + o(1)) \frac{\log^2 T}{T}
\]
as $T \to \infty$. Taking $\eta$ sufficiently small, we see that
\[
J(\beta + 2 + \eta, T) - J(\beta + \eta, T) \leq 21.65 \frac{\log^2 T}{T}
\]
for $\beta > 0$ and all sufficiently large $T$. This gives the upper bound and completes the proof of the theorem.

We now prove the corollary. For $\beta > 2$ the corollary follows immediately from the theorem. Suppose then that $1 \leq \beta \leq 2$. By (2) and the fact that $J(\beta, T)$ is an increasing function of $\beta$ we have
\[
\frac{1}{2} \log^2 T \sim J(1, T) \leq J(\beta, T) \leq J(3, T).
\]
Thus,

\[ J(\beta, T) \gg \frac{\log^2 T}{T} \gg \beta \frac{\log^2 T}{T} \]

and, by the theorem,

\[ J(\beta, T) \leq (J(3, T) - J(1, T)) + J(1, T) \leq \frac{\log^2 T}{T} \ll \beta \frac{\log^2 T}{T}. \]

REFERENCES


DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE, SAN JOSE STATE UNIVERSITY, SAN JOSE, CALIFORNIA 95192

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ROCHESTER, ROCHESTER, NEW YORK 14627