Large Differences Between Consecutive Prime Numbers

By

Daniel Alan Goldston

A.B. (University of California) 1975
M.A. (University of California) 1978

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Mathematics

in the

GRADUATE DIVISION

OF THE

UNIVERSITY OF CALIFORNIA, BERKELEY

Approved:

R. Sherman Lehman .... May 13, 1981

Chairman Date

Andrew P. Ogg .... May 21, 1981

John V. Wehausen .... May 20, 1981

....

Dedicated to the memory of Caryl Hashitani

Table of Contents

Acknowledgments

I would like to thank my adviser Professor R.S. Lehman for many years of help and encouragement, as well as the difficult task of reading through and catching many mistakes in this thesis.

I also thank Professors H.L. Montgomery and D.R. Heath-Brown for reading an earlier version of my results, and making corrections and helpful suggestions.

Over the years I have benefited from the support of many people. I would especially like to thank my parents, and my friends Edwina Chan, Leslie Finta, Koji Maru, Suzanne Niki, and Angela Cheer, who also made helpful suggestions concerning my research.

Finally, I thank Valerie Heatlie for her excellent typing of this thesis.

1. Introduction

Let p_n denote the n-th prime number. In this thesis we study $p_{n+1} - p_n$, the difference between consecutive prime numbers. By the prime number theorem there are $\sim x/\log x$ primes in the interval $[2,x]$ and therefore $p_{n+1} - p_n$ is on average $\log x$ (or $\log p_n$) in this interval. We can express this by the relation

$$(1.1) \qquad \sum_{n=1}^{N} \frac{(p_{n+1} - p_n)}{\log p_n} \sim N \qquad \text{as } N \to \infty .$$

We next ask how $p_{n+1} - p_n$ is distributed around its average. On the one hand we expect $p_{n+1} - p_n = 2$ infinitely often (i.e. the twin prime conjecture), but this has never been proved. On the other hand, more is known about large differences. First, Rankin [24] has proved, for any $\varepsilon > 0$,

$$(1.2) \qquad p_{n+1} - p_n > \frac{(1-\varepsilon)e^C \log p_n \log\log p_n \log\log\log\log p_n}{(\log\log\log p_n)^2}$$

infinitely often, where C is Euler's constant. Second, it has recently been proved by Heath-Brown and Iwaniec [16] that, for any $\varepsilon > 0$,

$$(1.3) \qquad p_{n+1} - p_n \ll p_n^{11/20+\varepsilon} \, [1] .$$

---

[1] $f \ll g$ means there exists an absolute positive constant A such that $|f| \leqslant Ag$.

The conjecture is

$$(5) \qquad \sum_{0 < \gamma, \gamma' \leq T} x^{i(\gamma-\gamma')} \, w(\gamma-\gamma') \ll T \log T$$

uniformly for $T \leq x \leq T^{2+\delta}$, $\delta > 0$. Here $w(u) = \frac{4}{4 + u^2}$ and γ denotes the imaginary part of a zero of the zeta function. This conjecture is related to earlier work of H.L. Montgomery on the pair correlation of zeros of the zeta function.

The method of proof of (3) and (4) uses an improved form of Cramér's original method, together with the conjecture (5). While (5) remains unproved, we prove, assuming the Riemann Hypothesis, that (5) is true "on average". We also prove, again on the Riemann Hypothesis, that the conjecture holds in the range $T \leq x \leq T \log T$.

R. Sherman Lehman

LARGE DIFFERENCES BETWEEN CONSECUTIVE PRIME NUMBERS

by

Daniel Alan Goldston

Abstract

Let p_n denote the n-th prime number. We wish to examine $p_{n+1} - p_n$. The results we are concerned with here are, assuming the Riemann Hypothesis,

(1) $$p_{n+1} - p_n \ll p_n^{\frac{1}{2}} \log p_n \quad ,$$

and

(2) $$\sum_{\substack{p_n \leqslant x \\ p_{n+1}-p_n \geqslant H}} (p_{n+1} - p_n) \ll \frac{x \log^2 x}{H} \quad ,$$

uniformly for $H \geqslant 1$. These results are due to Cramér (1920) and Selberg (1943), respectively.

We prove, assuming the Riemann Hypothesis and an additional conjecture on the zeros of the zeta function,

(3) $$p_{n+1} - p_n \ll p_n^{\frac{1}{2}} \log^{\frac{1}{2}} p_n \quad ,$$

and

(4) $$\sum_{\substack{p_n \leqslant x \\ p_{n+1}-p_n \geqslant H}} (p_{n+1} - p_n) \ll \frac{x \log x}{H} \quad ,$$

uniformly for $H \geqslant 1$.

Probably (1.2) is closer to the truth than (1.3). Shanks [26] conjectured

$$\lim_{x\to\infty}\left[\sup_{p_n\leqslant x}\left(\frac{p_{n+1}-p_n}{\log^2 p_n}\right)\right] = 1 \tag{1.4}$$

which strengthens slightly an earlier conjecture of Cramér [4][5]. However this conjecture remains beyond anything so far proved, even on the Riemann hypothesis.

The first work on large differences between consecutive primes was done by Cramér. Most of his results assume the RH (the Riemann hypothesis). Cramér's starting point is the formula [1]

$$\begin{aligned}\sum_{\gamma>0} e^{\rho z} = {} & \frac{e^z}{2\pi i}\sum_{n=2}^{\infty}\frac{\Lambda(n)}{n}\left\{\frac{1}{z-\log n}+\frac{1}{\log n}\right\} + \\ & + \frac{1}{2\pi i}\sum_{n=2}^{\infty}\frac{\Lambda(n)}{n}\left\{\frac{1}{z+\log n}-\frac{1}{\log n}\right\} + \\ & + \left(\frac{1}{4}+\frac{C+\log 2\pi}{2\pi i}\right)\left(1+\frac{1}{z}\right) + \frac{1}{2\pi i}\frac{\Gamma'}{\Gamma}\left(\frac{z}{\pi i}\right) + \\ & + \frac{1}{2}e^z - \frac{z}{2\pi i}\int_0^1 e^{sz}\log|\zeta(s)|\,ds - \\ & - \frac{1}{2\pi i z}\int_0^\infty \frac{t}{e^t-1}\frac{dt}{t+z}, \qquad \text{for } \operatorname{Im}(z) > 0, \end{aligned} \tag{1.5}$$

where $\rho = \beta + i\gamma$ is a non-trivial zero of the Riemann zeta function, and $\Lambda(n)$ is $\log p$ if n is a prime power p^m and zero otherwise. We now take real parts of (1.5). Letting $z = -\log\tau + iy$, $0<y\leqslant 1$, and $\tau\to\infty$, it follows that (see [5])

$$(1.6) \qquad -2\pi \operatorname{Re} \sum_{\gamma>0} e^{\rho(-\log \tau + iy)} = \sum_{n=2}^{\infty} \frac{\Lambda(n)}{n} \left\{ \frac{y}{(\log \tau/n)^2 + y^2} \right\} - \pi + O\left(\frac{1}{\log \tau}\right) .$$

In section 2 we give a different proof of this formula. The sum on the right is over primes and prime powers, and as usual the prime powers can be shown to make an insignificant contribution to the sum. If we let $y \to 0$, the factor

$$\frac{y}{(\log \tau/n)^2 + y^2}$$

is large for n near τ but falls off rapidly to zero for n away from τ. Thus equation (1.6) relates primes in short intervals to zeros of $\zeta(s)$. While (1.5) and (1.6) are true unconditionally, they seem most useful when the RH is assumed. Cramér proved, assuming the RH (see [2],[5]),

$$(1.7) \qquad p_{n+1} - p_n \ll p_n^{\frac{1}{2}} \log p_n ,$$

and

$$(1.8) \qquad \sum_{\substack{p_n \leqslant x \\ p_{n+1} - p_n \geqslant H}} (p_{n+1} - p_n) \ll \frac{x \log^3 x}{H \log H} ,$$

where $H = x^{\alpha} \log^{\beta} x$, $0 \leqslant \alpha \leqslant 1$, $\beta > 0$. Equation (1.8) is a measure of the frequency of large differences between primes; we see it becomes non-trivial as soon as H increases faster than $\log^3 x/\log \log x$.

Selberg [25] improved on (1.8). He proved, assuming the RH,

$$(1.9) \qquad \sum_{\substack{p_n \leqslant x \\ p_{n+1} - p_n \geqslant \frac{p_n}{x} H}} (p_{n+1} - p_n) \ll \frac{x \log^2 x}{H} ,$$

uniformly for $H \geqslant 1$. This immediately shows

$$(1.10) \qquad \sum_{\substack{p_n \leqslant x \\ p_{n+1} - p_n \geqslant H}} (p_{n+1} - p_n) \ll \frac{x \log^2 x}{H} ,$$

which improves (1.8) when $H \leqslant x^{\varepsilon}$ $(\varepsilon > 0)$. Selberg's method is based on estimating the integral

$$(1.11) \qquad J(\beta,T) = \int_1^{T^\beta} (\psi(x + \tfrac{x}{T}) - \psi(x) - \tfrac{x}{T})^2 \frac{dx}{x^2} .$$

Here $\psi(x) = \sum_{n \leqslant x} \Lambda(n)$. Selberg proved, assuming the RH, for fixed β,

$$(1.12) \qquad J(\beta,T) \ll_\beta \frac{\log^2 T}{T} .$$

Recently Montgomery proved (see [9]), assuming the RH,

$$(1.13) \qquad J(\beta,T) \ll \beta \frac{\log^2 T}{T} .$$

For $0 < \beta \leqslant 1$ it is easy to prove unconditionally [9]

$$(1.14) \qquad J(\beta,T) \sim \frac{\beta^2}{2} \frac{\log^2 T}{T} \qquad \text{as} \quad T \to \infty .$$

Gallagher and Mueller [9] proved that an asymptotic relation also holds for $\beta > 1$, but this assumes the RH and also another conjecture concerning the zeros of $\zeta(s)$. Their result implies in particular, assuming the RH and conjecture B (to be described later), for $\beta \geqslant 1$,

$$(1.15) \qquad J(\beta,T) \sim (\beta - \tfrac{1}{2}) \frac{\log^2 T}{T} \qquad \text{as} \quad T \to \infty .$$

This result implies that (1.7) can be improved to give

$$p_{n+1} - p_n = o(p_n^{\frac{1}{2}} \log p_n) ,$$

and similarly for (1.10). Thus, the classical results of Cramér and Selberg can be improved by assuming additional information on the zeros of the zeta function. The purpose of this thesis is to improve on these results of Gallagher and Mueller. We must first formulate the conjectures needed.

Our starting point is the work of Montgomery [21] on the pair correlation of zeros of the zeta function. Montgomery defined

$$(1.16) \qquad F(\alpha) = F(\alpha,T) = \left(\frac{T \log T}{2\pi}\right)^{-1} \sum_{0<\gamma,\gamma'\leqslant T} T^{i\alpha(\gamma-\gamma')} w(\gamma-\gamma') ,$$

$$w(u) = \frac{4}{4+u^2} ,$$

where $T \geqslant 2$, α real. The function $F(\alpha)$ is an even function. Since

$$F(\alpha) = \frac{4}{T \log T} \int_{-\infty}^{\infty} \left| \sum_{0<\gamma\leqslant T} \frac{T^{i\alpha\gamma}}{1+(t-\gamma)^2} \right|^2 dt ,$$

we have

$$(1.17) \qquad F(\alpha) \geqslant 0 .$$

Montgomery proved, assuming the RH,

$$(1.18)\qquad F(\alpha) = (1+o(1))T^{-2\alpha}\log T + \alpha + o(1) \qquad \text{as} \quad T \to \infty ,$$

uniformly for $0 \leqslant \alpha \leqslant 1-\varepsilon$, for any $\varepsilon > 0$. This is in fact true for $0 \leqslant \alpha \leqslant 1$, see lemma B. Montgomery conjectured

Conjecture A. The relation

$$(1.19)\qquad F(\alpha) = 1 + o(1)$$

holds uniformly for $1 \leqslant \alpha \leqslant M$, M any fixed constant.

Let

$$(1.20)\qquad N(T,u) = \sum_{\substack{0<\gamma,\gamma'\leqslant T \\ 0<\gamma-\gamma'<u}} 1 \ .$$

Conjecture A implies

Conjecture B. We have

$$(1.21)\qquad N(T,u) = \frac{T}{2\pi}\log T \int_0^{\frac{u\log T}{2\pi}} \left(1 - \left(\frac{\sin \pi u}{\pi u}\right)^2 \right) du + o(T\log T)$$

uniformly for $0 < a \leqslant u/\log T \leqslant b < \infty$.

This conjecture has a number of important applications both to the zeta function and primes besides (1.15) (see [9]).

Our main theorem uses a conjecture considerably weaker than conjecture A. The conjecture is

Conjecture C. We have

$$(1.22)\qquad F(\alpha) \ll 1 \qquad \text{uniformly for} \quad 1 \leqslant \alpha \leqslant 2+\delta , \quad \delta > 0 .$$

Equivalently, this may be stated as

$$\sum_{0<\gamma,\gamma'\leqslant T} x^{i(\gamma-\gamma')} w(\gamma-\gamma') \ll T \log T \tag{1.23}$$

uniformly for $T \leqslant x \leqslant T^{2+\delta}$, $\delta > 0$.

<u>Theorem 1</u>. Assuming the Riemann Hypothesis and Conjecture C, we have

$$\sum_{\substack{p_n \leqslant x \\ p_{n+1} - p_n \geqslant H}} (p_{n+1} - p_n) \ll \frac{x \log x}{H} \tag{1.24}$$

uniformly for $H \geqslant 1$.

The proof of this theorem follows Cramér's approach, and starts from equation (1.6). The proof also involves a technical improvement (lemma 3) which shows how Cramér's method can be sharpened to give (1.10) instead of (1.8) even without Conjecture C. Theorem 1 is proved in section 3.

Corollary 1. Assuming the Riemann Hypothesis and Conjecture C, we have

$$p_{n+1} - p_n \ll p_n^{\frac{1}{2}} \log^{\frac{1}{2}} p_n \,. \tag{1.25}$$

To prove this, suppose $p_{n+1} - p_n = d$ is any given difference between consecutive primes. Taking $x = p_n$ and $H = d$ in (1.24) implies

$$d \ll \frac{p_n \log p_n}{d}$$

which gives $d \ll p_n^{\frac{1}{2}} \log^{\frac{1}{2}} p_n$.

Corollary 2. Assuming the Riemann Hypothesis and Conjecture C, we have

$$\sum_{p_n \leqslant x} (p_{n+1} - p_n)^2 \ll x \log^2 x . \tag{1.26}$$

To prove this, integrate (1.24) with respect to H from 1 to x. We have

$$\int_1^x \left\{ \sum_{\substack{p_n \leqslant x \\ p_{n+1} - p_n \geqslant H}} (p_{n+1} - p_n) \right\} dH \ll x \log^2 x ,$$

while the left side is

$$= \sum_{p_n \leqslant x} (p_{n+1} - p_n) \int_1^{p_{n+1}-p_n} dH = \sum_{p_n \leqslant x} (p_{n+1} - p_n)^2 + O(x) .$$

The result now follows.

Without Conjecture C the best result is Selberg's

$$\sum_{p_n \leqslant x} (p_{n+1} - p_n)^2 \ll x \log^3 x . \tag{1.27}$$

This follows in the same way by using (1.10); but if we integrate (1.9) with respect to H we obtain, assuming the RH,

$$\sum_{p_n \leqslant x} \frac{(p_{n+1} - p_n)^2}{p_n} \ll \log^3 x . \tag{1.28}$$

Equation (1.27) follows easily from (1.28), but the reverse argument loses a factor of log x, so (1.28) can be viewed as somewhat stronger. Equation (1.28) is also of interest because it applies to a question considered by Cramér (see [5], pg. 44). With regard to Corollary 2, Erdös [7] has conjectured

$$(1.29) \qquad \sum_{p_n \leqslant x} (p_{n+1} - p_n)^2 \ll x \log x \quad .$$

The prime number theorem and the Cauchy-Schwarz inequality give, for any $\varepsilon > 0$ and x sufficiently large,

$$(1.30) \qquad (1 - \varepsilon) x \log x \leqslant \Big(\sum_{p_n \leqslant x} 1 \Big)^{-1} \Big(\sum_{p_n \leqslant x} (p_{n+1} - p_n) \Big)^2$$

$$\leqslant \sum_{p_n \leqslant x} (p_{n+1} - p_n)^2 \qquad .$$

Corollary 3. Assuming the Riemann Hypothesis and Conjecture C, we have

$$(1.31) \qquad \sum_{\substack{p_n \leqslant x \\ p_{n+1} - p_n \geqslant H}} (p_{n+1} - p_n) = o(x) \qquad \text{if } \frac{H}{\log x} \to \infty \text{ as } x \to \infty .$$

This is immediate from Theorem 1. (In fact, from the proof of Theorem 1 it is clear that we only need Conjecture C to hold in the range $1 \leqslant \alpha \leqslant 1 + \delta$, $\delta > 0$; or more precisely that (1.23) holds for $T \leqslant x \leqslant T\phi(T) \log T$, where $\phi(T) \to \infty$ as $T \to \infty$.)

The question of when (1.31) holds was first considered by Cramér [4], and can be interpreted as follows. The sum over all prime intervals is

$$\sum_{p_n \leq x} (p_{n+1} - p_n) \sim x \quad .$$

From this sum we delete all "short" prime intervals $p_{n+1} - p_n < H$, and wish to know how small we can take H and have

$$\sum_{\substack{p_n \leq x \\ p_{n+1} - p_n \geq H}} (p_{n+1} - p_n) = o(x) \quad .$$

i.e., almost all positive integers are contained in prime intervals $p_{n+1} - p_n < H$. On the RH, Selberg's result (1.10) shows (1.31) is true if $H/\log^2 x \to \infty$. Corollary 3 shows this is true for $H/\log x \to \infty$, subject to the RH and Conjecture C. In view of (1.1) it is likely this is the best possible. If we examine this more closely, we see Theorem 1 implies

$$\sum_{\substack{p_n \leq x \\ p_{n+1} - p_n \geq \alpha \log x}} (p_{n+1} - p_n) \ll \frac{x}{\alpha} \quad .$$

Gallagher [8] has examined the distribution of primes in intervals $[n, n + \alpha \log n]$, assuming the Hardy-Littlewood r-tuple conjecture. Let $\pi_a(x)$ denote the number of $n \leq x$ for which $n+a_1, n+a_2, \ldots, n+a_r$ are all prime. Then the conjecture is

$$(1.32) \qquad \pi_a(x) \sim C_a \frac{x}{\log^r x} \qquad \text{as} \quad x \to \infty \quad ,$$

provided $C_a \neq 0$, where

$$c_a = \prod_p \left(\frac{p}{p-1}\right)^{r-1} \left(\frac{p - \nu_a(p)}{p-1}\right)$$

and $\nu_a(p)$ is the number of distinct residue classes mod p occupied by $a_1, a_2, \ldots, a_r$. Gallagher's theorem assumes (1.32) holds uniformly over all tuples $1 \leqslant a_1, \ldots, a_r \leqslant H$, for $H = \alpha \log x$. As a special case of this theorem, we have for fixed α, as $x \to \infty$,

$$(1.33) \qquad \sum_{\substack{p_n \leqslant x \\ p_{n+1} - p_n \geqslant \alpha \log x}} (p_{n+1} - p_n) \sim (1 + \alpha) e^{-\alpha} x$$

and

$$(1.34) \qquad \sum_{\substack{p_n \leqslant x \\ p_{n+1} - p_n \geqslant \alpha \log x}} 1 \sim e^{-\alpha} x \quad .$$

We thus see our theorem gives a far slower fall off in these sums as $\alpha \to \infty$. While these results are for fixed α, they may still indicate the behavior of these sums for $\alpha = \alpha(x) \to \infty$. If this is the case, we see that a result like (1.4) seems plausible. Also, if we formally integrate (1.33) with respect to α from 0 to ∞, we have

$$\sum_{p_n \leqslant x} \frac{(p_{n+1} - p_n)^2}{\log x} \sim x \int_0^\infty (1 + \alpha) e^{-\alpha} d\alpha \sim 2x \quad ,$$

which suggests we might strengthen the conjecture (1.29) to

$$(1.35) \qquad \sum_{p_n \leqslant x} (p_{n+1} - p_n)^2 \sim 2x \log x \quad .$$

Hooley [17] has already made a more general conjecture for

$$\sum_{p_n \leqslant x} (p_{n+1} - p_n)^{\gamma} \qquad (\gamma > 0) \quad .$$

We mention that recently the questions considered here have been treated unconditionally. We refer the reader to the articles of Heath-Brown [11],[12],[13], and Ivić [19].

In section 4 we prove two results on $F(\alpha)$. The trivial estimate for $F(\alpha)$ is $F(\alpha) \leqslant F(0) = (1 + o(1))\log T$ by (1.18). Hence

(1.36) $$F(\alpha) \leqslant (1 + o(1))\log T \qquad \text{as } T \to \infty \quad .$$

We first prove

Lemma A. Assuming the Riemann Hypothesis, we have for any $\varepsilon > 0$ and T sufficiently large,

(1.37) $$\int_x^{x+1} F(\alpha)\,d\alpha \leqslant \frac{8}{3} + \varepsilon \quad ,$$

for any x independent of T.

This shows Conjecture C is true on average. Hence, given any $\varepsilon > 0$, there exists a set $A = A(\varepsilon)$ such that the measure $|A| < \varepsilon$ and

(1.38) $$F(\alpha) \ll 1 \qquad \text{for } \alpha \varepsilon [1, 2+\delta] \setminus A \quad .$$

By Dirichlet's theorem, we know there are large $\alpha = \alpha(T)$ where $F(\alpha)$ is as close to $F(0) \sim \log T$ as we wish. Equation (1.36) shows these large values of $F(\alpha)$ occur on intervals of measure $\ll \frac{1}{\log T}$ at most. For $\alpha \varepsilon [1, 2+\delta]$ we need to prove these spikes do not occur. Our second result shows this is true for a small interval to the right of $\alpha = 1$. The result is

Lemma B. Assuming the Riemann Hypothesis,

$$F(\alpha) = \alpha + O(T^{\alpha-1}(\log T)^{-1}) + O(T^{\frac{\alpha-1}{2}} \log^{-\frac{1}{2}} T) + O(\log^{-\frac{1}{2}} T) \tag{1.39}$$

for $\alpha \geqslant \varepsilon$, or equivalently, for $x \geqslant T^{\varepsilon}$,

$$\sum_{0<\gamma,\gamma'\leqslant T} x^{i(\gamma-\gamma')} w(\gamma-\gamma') = \frac{T \log x}{2\pi} + O(x) + O(\sqrt{Tx \log x}\,) + O(T\sqrt{\log x}\,) . \tag{1.40}$$

This shows (1.18) holds uniformly in $0 \leqslant \alpha \leqslant 1$, which is stated by Montgomery in his paper. Equation (1.40) is nontrivial for $x \leqslant T \log^2 T$; for larger x the trivial estimate (1.36) is stronger. If we could replace the term $O(x)$ by $o(x)$ in (1.40), then this result would be strong enough to replace the use of Conjecture C in Corollary 3.

After I had obtained the results in the first four sections of this thesis, I learned that some of the results have already been obtained. First, J. Mueller [23] has obtained a somewhat weaker form of Corollary 1. Second, D.R. Heath-Brown [15] has obtained all my corollaries, as well as a result on small differences between consecutive primes. His method is different. H.L. Montgomery pointed out to me another method in a letter. This method is quite flexible, and makes it possible to avoid most of the technical difficulties in my proof of Theorem 1. In section 5 I present this method and use it to again prove Theorem 1. I also prove Theorem 2, which is a result of Heath-Brown similar to Corollary 3. Finally, I prove a slightly stronger form of Heath-Brown's small difference result:

Theorem 3. (Assume the Riemann Hypothesis) There exist absolute constants C and K such that, if for $x \geqslant 4$,

$$(1.41) \qquad \left| \sum_{0<\gamma,\gamma'\leqslant T} x^{i(\gamma-\gamma')} w(\gamma-\gamma') - \frac{T\log T}{2\pi} \right| \leqslant \varepsilon(x) T \log T$$

holds uniformly in

$$\frac{x}{K\log x} \leqslant T \leqslant \frac{x}{[\varepsilon(x)]^{4/3}\log x} \quad ,$$

where $\varepsilon(x)$ is a positive function such that $\frac{1}{\log x} \leqslant \varepsilon(x) \leqslant 1$, then

$$(1.42) \qquad \min_{\frac{x}{C}\leqslant p_n \leqslant Cx} \left(\frac{p_{n+1} - p_n}{\log p_n} \right) \leqslant K\varepsilon(x) \quad .$$

We thus see that if we could replace the right hand side of (1.40) by

$$= \frac{T\log x}{2\pi} (1 + o(1)) + o(x)$$

for $T \leqslant x \leqslant KT\log T$, then we would have (with RH)

$$\underline{\lim}_{n\to\infty} \left(\frac{p_{n+1} - p_n}{\log p_n} \right) = 0 \quad .$$

2. Proof of Cramér's Formula

In this section the Riemann Hypothesis is not assumed. Cramér proved (1.5) in [1] by working with the multivalued function $\log \zeta(s)$. The branches of $\log \zeta(s)$ cause complications; we avoid this by working with $\frac{\zeta'}{\zeta}(s)$ and then use a limiting case to obtain (1.6).

Consider

$$\int_C e^{sz} \frac{\zeta'}{\zeta}(s)ds \quad ,$$

where C is the rectangle with vertices at 2, $2+iT_n$, $-\delta+iT_n$, and $-\delta$, and along the bottom horizontal side we indent the contour to avoid the pole at $s = 1$ with a semicircle of radius $\frac{1}{2}$ centered at $s = 1$. Here $0 < \delta < \frac{1}{2}$, and T_n is picked, as usual (see [18], pg. 71), such that $n \leqslant T_n < n+1$ and

$$\frac{\zeta'}{\zeta}(\sigma + iT_n) \ll \log^2 T_n \qquad (-1 \leqslant \sigma \leqslant 2) \quad .$$

Letting $n \to \infty$, we see for $\mathrm{Im}(z) > 0$ the integral along the top of the contour goes to zero. By the residue theorem we conclude, for $\mathrm{Im}(z) = y > 0$,

$$(2.1) \qquad 2\pi i \sum_{\gamma > 0} e^{\rho z} = \int_0^\infty e^{(2+it)z} \frac{\zeta'}{\zeta}(2+it)i\,dt - \int_0^\infty e^{(-\delta+it)z} \frac{\zeta'}{\zeta}(-\delta+it)i\,dt + \int_L e^{sz} \frac{\zeta'}{\zeta}(s)ds \quad ,$$

where L is along the real axis from $s = -\delta$ to $s = 2$, indented around $s = 1$ as above. Now let $z = -\log \tau + iy$, and consider (2.1) as $\tau \to \infty$ and $0 < y \leqslant 1$. From now on all error estimates are absolute in all the variables. First, on L $\frac{\zeta'}{\zeta}(s)$ is regular and hence bounded so

$$\int_L e^{sz} \frac{\zeta'}{\zeta}(s)ds \ll \int_{-\delta}^{\frac{1}{2}} \tau^{-\sigma} d\sigma + O(\tau^{-\frac{1}{2}}) \ll \frac{\tau^{\delta}}{\log \tau} ,$$

$$\text{for} \quad 0 \leqslant \delta \leqslant \tfrac{1}{2} .$$

Next, by the absolute convergence of the Dirichlet series for $\frac{\zeta'}{\zeta}(2+it)$, we have

$$\int_0^{\infty} e^{(2+it)z} \frac{\zeta'}{\zeta}(2+it)dt = -e^{2z} \sum_{n=2}^{\infty} \frac{\Lambda(n)}{n^2} \int_0^{\infty} e^{it(z-\log n)} idt$$

$$= e^{2z} \sum_{n=2}^{\infty} \frac{\Lambda(n)}{n^2} \frac{1}{z - \log n} \ll \tau^{-2} \sum_{n=2}^{\infty} \frac{\Lambda(n)}{n^2} \ll \tau^{-2} .$$

Hence we conclude, for $0 < \delta \leqslant \frac{1}{2}$ and $0 < y \leqslant 1$,

$$(2.2) \qquad 2\pi i \sum_{\gamma > 0} e^{\rho z} = -i \int_0^{\infty} e^{(-\delta+it)z} \frac{\zeta'}{\zeta}(-\delta+it)dt + O\left(\frac{\tau^{\delta}}{\log \tau}\right).$$

By the functional equation, the integral on the right is

$$= -i\int_0^\infty e^{(-\delta+it)z}\log 2\pi dt - i\int_0^\infty e^{(-\delta+it)z}\frac{\pi}{2}\cot\left(\frac{\pi}{2}(-\delta+it)\right)dt$$

$$+ i\int_0^\infty e^{(-\delta+it)z}\frac{\Gamma'}{\Gamma}(1+\delta-it)dt + i\int_0^\infty e^{(-\delta+it)z}\frac{\zeta'}{\zeta}(1+\delta-it)dt$$

$$= I_1 + I_2 + I_3 + I_4 \quad .$$

As before we have

$$I_1 = \frac{\log 2\pi e^{-\delta z}}{z} \ll \frac{\tau^\delta}{\log \tau} \quad ,$$

and

$$I_4 = -i\sum_{n=2}^\infty \frac{\Lambda(n)}{n^{1+\delta}} e^{-\delta z}\int_0^\infty e^{it(z+\log n)}dt$$

$$= e^{-\delta z}\sum_{n=2}^\infty \frac{\Lambda(n)}{n^{1+\delta}} \frac{1}{z+\log n} \quad .$$

For I_3 we use the well-known formula of Gauss:

$$(2.3) \qquad \frac{\Gamma'}{\Gamma}(w) = \int_0^\infty \left(\frac{e^{-u}}{u} - \frac{e^{-wu}}{1-e^u}\right)du \quad , \quad \text{for} \quad \operatorname{Re} w > 0 \quad .$$

Thus

$$I_3 = i\int_0^\infty e^{(-\delta+it)z}\left\{\int_0^\infty\left(\frac{e^{-u}}{u} - \frac{e^{-u(1+\delta-it)}}{1-e^{-u}}\right)du\right\}dt$$

$$= ie^{-\delta z}\int_0^\infty\int_0^\infty\left(\frac{e^{-u+itz}}{u} - \frac{e^{-u(1+\delta)+it(z+u)}}{1-e^{-u}}\right)dtdu$$

$$= -e^{-\delta z}\int_0^\infty\left(\frac{e^{-u}}{zu} - \frac{e^{-u(1+\delta)}}{(z+u)(1-e^{-u})}\right)du \quad .$$

The interchange of the order of integration is justified by Fubini's theorem. Since

$$e^{-u}\left(\frac{1}{zu} - \frac{e^{-\delta u}}{(z+u)(1-e^{-u})}\right) \ll \begin{cases} \dfrac{1}{\log \tau} & \text{if} \quad 0 \leqslant u \leqslant 1 \\[2ex] \dfrac{e^{-u}}{\log \tau} & \text{if} \quad u > 1 \end{cases},$$

we have

$$I_3 \ll \frac{\tau^\delta}{\log \tau} \quad .$$

We substitute these results into (2.2), and on taking imaginary parts obtain

$$(2.4) \qquad 2\pi \operatorname{Re}\sum_{\gamma>0} e^{\rho z} = \operatorname{Im} e^{-\delta z}\sum_{n=2}^{\infty}\frac{\Lambda(n)}{n^{1+\delta}}\frac{1}{z+\log n} + \operatorname{Im}(I_2) + + O\left(\frac{\tau^\delta}{\log \tau}\right),$$

where

$$\mathrm{Im}(I_2) = -\frac{\pi}{2}\,\mathrm{Re}\int_0^\infty e^{(-\delta+it)z}\cot\left(\frac{\pi}{2}(-\delta+it)\right)dt \quad .$$

Equation (2.4) has been derived for $\delta > 0$, and we now take the limit as $\delta \to 0$, holding z fixed. The error term which holds uniformly in δ, approaches $O\left(\frac{1}{\log \tau}\right)$. The sum is

$$\sum_{n=2}^{\infty}\frac{\Lambda(n)}{n^{1+\delta}}\left\{\frac{-\tau^{\delta}y\cos\delta y+\tau^{\delta}\log(\tau/n)\sin\delta y}{(\log\tau/n)^2+y^2}\right\}=$$

$$= -\tau^{\delta}\cos\delta y\sum_{n=2}^{\infty}\frac{\Lambda(n)}{n^{1+\delta}}\,\frac{y}{(\log\tau/n)^2+y^2} +$$

$$+\tau^{\delta}\sin\delta y\sum_{n=2}^{\infty}\frac{\Lambda(n)}{n^{1+\delta}}\,\frac{\log\tau/n}{(\log\tau/n)^2+y^2} \quad .$$

Letting $\delta \to 0$, the first term is

$$= -\sum_{n=2}^{\infty}\frac{\Lambda(n)}{n}\,\frac{y}{(\log\tau/n)^2+y^2} \quad ,$$

where for any fixed z the sum is

$$\ll_z \sum_{n=2}^{\infty}\frac{\Lambda(n)}{n\log^2 n} \ll_z 1 \quad .$$

Thus the sum converges to a well-defined function of z. The second term is

$$= \tau^{\delta}\sin \delta y \left\{ \sum_{n \leqslant \tau/2} + \sum_{\tau/2 < n \leqslant 2\tau} + \sum_{n > 2\tau} \right\}$$

$$= \tau^{\delta}\sin \delta y \left\{ \Sigma_1 + \Sigma_2 + \Sigma_3 \right\}.$$

Now

$$\Sigma_1 \ll \sum_{n \leqslant \tau/2} \frac{\Lambda(n)}{n^{1+\delta}(\log \tau - \log n)}$$

$$\ll \int_1^{\tau/2} \frac{du}{u^{1+\delta}(\log \tau - \log u)}$$

$$\ll \tau^{-\delta} \int_{\log 2}^{\log \tau} \frac{e^{\delta v}}{v}\, dv \ll \int_{\log 2}^{\log \tau} \frac{dv}{v}$$

$$\ll \log \log \tau \quad .$$

Next,

$$\Sigma_2 \ll \frac{1}{y} \sum_{\tau/2 \leqslant n \leqslant 2\tau} \frac{\Lambda(n)}{n^{1+\delta}} \ll \frac{\tau^{-\delta}}{y} \quad ,$$

and

$$\Sigma_3 \ll \sum_{n \geqslant 2\tau} \frac{\Lambda(n)}{n^{1+\delta}(\log n - \log \tau)}$$

$$\ll \int_{2\tau}^{\infty} \frac{du}{u^{1+\delta}(\log u - \log \tau)}$$

$$\ll \tau^{-\delta} \int_{\log 2}^{\infty} e^{-\delta v}\, \frac{dv}{v}$$

$$\ll \tau^{-\delta} \int_{\delta \log 2}^{\infty} e^{-u} \frac{du}{u}$$

$$\ll \tau^{-\delta} \left\{ \int_{\delta \log 2}^{1} \frac{du}{u} + \int_{1}^{\infty} e^{-u} du \right\}$$

$$\ll \tau^{-\delta} |\log \delta| \quad .$$

We conclude the second term is

$$= O\left(\tau^{\delta} |\sin \delta y| \left(\log \log \tau + \frac{\tau^{-\delta}}{y} + \tau^{-\delta} \log \frac{1}{\delta} \right)\right)$$

$$= O(\delta y \tau^{\delta} \log \log \tau) + O(\delta) + O\left(y\delta \log \left(\frac{1}{\delta}\right) \right) \quad .$$

Hence it goes to zero as $\delta \to 0$.

From these results and (2.4), we see that (1.6) will be proved if

$$\text{(2.5)} \qquad \lim_{\delta \to 0} \operatorname{Im}(I_2) = \pi + O\left(\frac{1}{\log \tau} \right) \quad .$$

To prove this, we consider

$$- \frac{\pi}{2} \int_{C} e^{sz} \cot \left(\frac{\pi}{2} s \right) ds \quad ,$$

where C is formed from the rectangle with vertices at 0, iT_n, $-\delta+iT$, and $-\delta$ by indenting a semicircle of radius ε around 0 from $-\varepsilon$ to $i\varepsilon$. Here T_n is chosen as before; on letting $n \to \infty$ the integral along the top segment goes to zero. By Cauchy's theorem, letting $\varepsilon \to 0$, we have

$$\mathrm{Im}(I_2) = -\frac{\pi}{2} \lim_{\varepsilon \to 0} \mathrm{Re} \int_{\varepsilon}^{\infty} e^{itz} \cot\left(\frac{i\pi t}{2}\right) dt$$

$$- \frac{\pi}{2} \lim_{\varepsilon \to 0} \mathrm{Im} \int_{-\delta}^{-\varepsilon} e^{\sigma z} \cot\left(\frac{\pi}{2}\sigma\right) d\sigma$$

$$- \frac{\pi}{2} \lim_{\varepsilon \to 0} \mathrm{Im} \int_{s} e^{sz} \cot\left(\frac{\pi}{2} s\right) ds$$

$$= J_1 + J_2 + J_3 \quad ,$$

where s is the semicircle described above.

Now

$$J_1 = \lim_{\varepsilon \to 0} \frac{\pi}{2} \mathrm{Im} \int_{\varepsilon}^{\infty} e^{itz} \left(1 - \frac{2}{1 - e^{-\pi t}} \right) dt$$

$$= \lim_{\varepsilon \to 0} \pi \int_{\varepsilon}^{\infty} \frac{e^{-yt} \sin(t \log \tau)}{1 - e^{-\pi t}} dt + O\left(\frac{1}{\log \tau}\right) .$$

This integral converges as $\varepsilon \to 0$, and can be evaluated using (2.3). Taking imaginary parts of (2.3), we have, for $w = u + iv$, $u > 0$,

$$\mathrm{Im} \frac{\Gamma'}{\Gamma}(w) = \int_{0}^{\infty} \frac{e^{-ut} \sin(tv)}{1 - e^{-t}} dt \quad ,$$

so a change of variables shows

$$J_1 = \operatorname{Im} \frac{\Gamma'}{\Gamma}\left(\frac{-iz}{\pi}\right) + o\left(\frac{1}{\log \tau}\right) .$$

By Stirling's formula, for $\operatorname{Re} w > 0$,

$$\frac{\Gamma'}{\Gamma}(w) = \log w + O\left(\frac{1}{|w|}\right) . \tag{2.6}$$

Hence we have

$$J_1 = \arg\left(\frac{-iz}{\pi}\right) + o\left(\frac{1}{\log \tau}\right)$$

$$= \frac{\pi}{2} + o\left(\frac{1}{\log \tau}\right) .$$

Next,

$$J_2 = \lim_{\varepsilon \to 0} - \frac{\pi}{2} \int_{-\delta}^{-\varepsilon} \tau^{-\sigma} \sin \sigma y \cot\left(\frac{\pi}{2}\sigma\right) d\sigma$$

$$= - \frac{\pi}{2} \int_{-\delta}^{0} \tau^{-\sigma} \sin(\sigma y) \cot\left(\frac{\pi}{2}\sigma\right) d\sigma \quad ,$$

since the integrand has a removable singularity at $\sigma = 0$. Hence

$$\lim_{\varepsilon \to 0} J_2 = 0 .$$

Finally, since $s = 0$ is a simple pole of $e^{sz} \cot \frac{\pi}{2} s$, we have as $\varepsilon \to 0$,

$$J_3 = - \frac{\pi}{2} \operatorname{Im} \left\{ - \frac{1}{4} \cdot 2\pi i \operatorname*{Res}_{s=0} \left(e^{sz} \cot \frac{\pi}{2} s\right) \right\}$$

$$= - \frac{\pi}{2} \operatorname{Im} \left\{ - \frac{\pi i}{2} \cdot \frac{2}{\pi} \right\}$$

$$= \frac{\pi}{2} .$$

Combining these results proves (2.5), and thus the proof of Cramér's formula is complete.

3. Proof of Theorem 1

Starting from Cramér's formula (1.6), we have, for $x \geqslant 10$,

$$(3.1)\qquad \int_{\frac{x}{e}}^{ex} \left| \sum_{n=2}^{\infty} \frac{\Lambda(n)}{n} \frac{y}{(\log \tau/n)^2 + y^2} - \pi + O\left(\frac{1}{\log \tau}\right) \right|^2 d\tau =$$

$$= 4\pi^2 \int_{\frac{x}{e}}^{ex} \left| \operatorname{Re} \sum_{\gamma > 0} e^{\rho(-\log\tau + iy)} \right|^2 d\tau \ .$$

We shall throughout this section take $y = \frac{1}{T}$, $x \geqslant T$, and $T \to \infty$. We obtain Theorem 1 by estimating each side of (3.1). Our first estimate is

Lemma 1. Assuming the Riemann Hypothesis and Conjecture C, we have

$$(3.2)\qquad \int_{\frac{x}{e}}^{ex} \left| \operatorname{Re} \sum_{\gamma > 0} e^{\rho(-\log\tau + iy)} \right|^2 d\tau \ll T \log T$$

for $T \leqslant x \leqslant T^{2+\delta}$.

Without Conjecture C, the right side of (3.2) becomes $\ll T \log^2 T$, but this holds uniformly for all x.

Proof. We have

$$\int_{\frac{x}{e}}^{ex} \left|\mathrm{Re} \sum_{\gamma>0} e^{\rho(-\log\tau + iy)}\right|^2 d\tau =$$

$$= \int_{\frac{x}{e}}^{ex} \left|\mathrm{Re}\ e^{\frac{iy}{2}} \sum_{\gamma>0} e^{-\gamma y} e^{-i\gamma\log\tau}\right|^2 \frac{d\tau}{\tau}$$

$$= \int_{-1+\log x}^{1+\log x} \left|\mathrm{Re}\ e^{\frac{iy}{2}} \sum_{\gamma>0} e^{-\gamma y} e^{-i\gamma v}\right|^2 dv$$

$$\leqslant \int_{-1+\log x}^{1+\log x} \left|\sum_{\gamma>0} e^{-\gamma y} e^{-i\gamma v}\right|^2 dv$$

$$\leqslant \int_{-1}^{1} \left|\sum_{\gamma>0} e^{-\gamma y} x^{i\gamma} e^{i\gamma u}\right|^2 du$$

$$\leqslant e^2 \int_{-1}^{1} \left|\sum_{\gamma>0} e^{-\gamma y} x^{i\gamma} e^{i\gamma u}\right|^2 e^{-2|u|} du$$

$$\leqslant e^2 \int_{-\infty}^{\infty} \left|\sum_{\gamma>0} e^{-\gamma y} x^{i\gamma} e^{i\gamma u}\right|^2 e^{-2|u|} du$$

$$\leqslant e^2 \sum_{\gamma>0} \sum_{\gamma'>0} e^{-(\gamma+\gamma')y} x^{i(\gamma-\gamma')} \int_{-\infty}^{\infty} e^{i(\gamma-\gamma')u} e^{-2|u|} du$$

$$\leqslant e^2 \sum_{\gamma,\gamma'>0} e^{-(\gamma+\gamma')y} x^{i(\gamma-\gamma')} w(\gamma-\gamma') \qquad ,$$

where $w(u) = \dfrac{4}{4+u^2}$ as in (1.16). The interchange of the order of summation and integration is justified by absolute convergence of the series proved below (see (3.5)). We conclude

$$(3.3) \qquad \int_{\frac{x}{e}}^{ex} \left| \operatorname{Re} \sum_{\gamma>0} e^{\rho(-\log\tau + iy)} \right|^2 d\tau \ll S(x,y) \quad ,$$

where

$$(3.4) \qquad S(x,y) = \sum_{\gamma,\gamma'>0} e^{-(\gamma+\gamma')y} x^{i(\gamma-\gamma')} w(\gamma-\gamma') \quad .$$

The function $S(x,y)$ behaves like $F(\alpha)$ where $x = T$, but we need to make some preliminary manipulations before we can apply Conjecture C.

First, since there are $O(\log\tau)$ zeros with ordinate between τ and $\tau+1$ (see [27]), we have, for $\omega \geqslant 2$, $0 < y < 1$,

$$\sum_{\omega<\gamma} e^{-\gamma y} \ll \sum_{n=0}^{\infty} \sum_{\omega+n<\gamma\leqslant\omega+n+1} e^{-\gamma y}$$

$$\ll \sum_{n=0}^{\infty} \log(\omega+n) e^{-\omega y} e^{-ny} \ll \int_{\omega}^{\infty} e^{-yu} \log u \, du$$

$$\ll \frac{e^{-\omega y}\log\omega}{y} + \frac{1}{y}\int_{\omega}^{\infty} \frac{e^{-uy}}{u} du \quad .$$

This last integral is equal to $\displaystyle\int_{\omega y}^{\infty} \frac{e^{-v}}{v} dv$, and this is

$$\ll \begin{cases} \log \frac{1}{\omega y} \ll \log\left(\frac{\omega}{y}\right) & \text{if } \omega y < 1 \\ e^{-\omega y} & \text{if } \omega y \geq 1 \end{cases} .$$

Thus, for $\omega \geq 2$, $0 < y < 1$,

$$(3.5) \qquad \sum_{\omega < \gamma} e^{-\gamma y} \ll \frac{e^{-\omega y} \log\left(\frac{\omega}{y}\right)}{y} .$$

The well known result (see [6])

$$(3.6) \qquad \sum_{\gamma} \frac{1}{1 + (\omega-\gamma)^2} \ll \log(|\omega| + 2)$$

can be proved similarly. By (3.5), since $y = \frac{1}{T}$,

$$\sum_{\substack{\gamma' > 0 \\ \gamma > 2T \log T}} e^{-(\gamma+\gamma')y} x^{i(\gamma-\gamma')} w(\gamma-\gamma') \ll$$

$$\ll \sum_{\gamma' > 0} e^{-\gamma' y} \sum_{\gamma > 2T \log T} e^{-\gamma y}$$

$$\ll \frac{\log\left(\frac{1}{y}\right)}{y} \cdot \frac{e^{-2 \log T} \log T}{y}$$

$$\ll \log^2 T .$$

By (3.6),

$$\sum_{\substack{\gamma'>0 \\ 0<\gamma<\frac{T}{\log^2 T}}} e^{-(\gamma+\gamma')y} x^{i(\gamma-\gamma')} w(\gamma-\gamma') \ll \sum_{0<\gamma<\frac{T}{\log^2 T}} \left(\sum_{\gamma'} \frac{1}{1+(\gamma-\gamma')^2} \right)$$

$$\ll \sum_{0<\gamma<\frac{T}{\log^2 T}} \log \gamma$$

$$\ll T \quad .$$

These last two estimates give

$$(3.7) \qquad S(x,y) = \sum_{\frac{T}{\log^2 T} \leqslant \gamma,\, \gamma' \leqslant 2T\log T} e^{-(\gamma+\gamma')y} x^{i(\gamma-\gamma')} w(\gamma-\gamma') + O(T) \, .$$

We now note terms with $|\gamma-\gamma'| > \log^3 T$ can be ignored with an error $O(T)$. To see this, we use the result in [9], for $1 \leqslant u \leqslant T$,

$$(3.8) \qquad N(T,u) \sim \frac{Tu \log^2 T}{4\pi^2} \quad .$$

Thus

$$\sum_{\substack{\frac{T}{\log^2 T} \leqslant \gamma,\, \gamma' \leqslant 2T\log T \\ |\gamma-\gamma'| > \log^3 T}} e^{-(\gamma+\gamma')y} x^{i(\gamma-\gamma')} w(\gamma-\gamma') \ll$$

$$\ll \sum_{0 \leqslant k \leqslant O(\log T)} \left\{ \sum_{\substack{\frac{T}{\log^2 T} \leqslant \gamma,\, \gamma' \leqslant 2T\log T \\ 2^k\log^3 T \leqslant |\gamma-\gamma'| \leqslant 2^{k+1}\log^3 T}} (\gamma-\gamma')^2 \right\}$$

$$\ll \sum_{0 \leqslant k \leqslant O(\log T)} \frac{N(2T\log T,\, 2^{k+1}\log^3 T)}{(2^k \log^3 T)^2} \ll T \quad .$$

For terms with $|\gamma-\gamma'| \leqslant \log^3 T$ we can replace $e^{-(\gamma+\gamma')y}$ in (3.7) by $e^{-2\gamma y}$ with an error

$$\ll \sum_{\substack{\frac{T}{\log^2 T} \leqslant \gamma\,,\, \gamma' \leqslant 2T\log T \\ |\gamma-\gamma'| \leqslant \log^3 T}} (e^{-(\gamma+\gamma')y} - e^{-2\gamma y})\, w(\gamma-\gamma')$$

$$\ll (e^{\log^3 T/T} - 1) \sum_{0<\gamma\,,\,\gamma' \leqslant 2T\log T} w(\gamma-\gamma') \ll \log^6 T \quad .$$

The last sum was estimated using (3.6). We have now shown

$$S(x,y) = \sum_{\substack{\frac{T}{\log^2 T} \leqslant \gamma\,,\, \gamma' \leqslant 2T\log T \\ |\gamma-\gamma'| < \log^3 T}} e^{-2\gamma y} x^{i(\gamma-\gamma')} w(\gamma-\gamma') + O(T) \quad . \tag{3.9}$$

Next, the sum on the right is

$$= \sum_{\frac{T}{\log^2 T} \leqslant \gamma \leqslant T} e^{-2\gamma y} \left\{ \sum_{\substack{\gamma' \\ |\gamma-\gamma'| < \log^3 T}} x^{i(\gamma-\gamma')} w(\gamma-\gamma') \right\} +$$

$$+ \sum_{0 \leqslant k \leqslant O(\log T)} \left(\sum_{kT < \gamma \leqslant (k+1)T} e^{-\gamma y} \left\{ \sum_{\substack{\gamma' \\ |\gamma-\gamma'| < \log^3 T}} x^{i(\gamma-\gamma')} w(\gamma-\gamma') \right\} \right) +$$

$$+ O(\log^8 T) \, .$$

The error term comes from the terms with $\frac{T}{\log^2 T} - \log^3 T < \gamma' < \frac{T}{\log^2 T}$ and

$2T \log T < \gamma' < 2T \log T + \log^3 T$ being added to the sum. We now apply partial summation since $e^{-2\gamma y}$ is non-increasing as γ increases. Hence

$$\sum_{kT<\gamma\leqslant(k+1)T} e^{-\gamma y} \left\{ \sum_{\substack{\gamma' \\ |\gamma-\gamma'|<\log^3 T}} x^{i(\gamma-\gamma')} w(\gamma-\gamma') \right\} \ll$$

$$\ll e^{-k} \max_{u_k \leqslant (k+1)T} \left[\sum_{kT<\gamma\leqslant u_k} \left\{ \sum_{\substack{\gamma' \\ |\gamma-\gamma'|<\log^3 T}} x^{i(\gamma-\gamma')} w(\gamma-\gamma') \right\} \right],$$

with a similar result for the sum with $\frac{T}{\log^2 T} \leqslant \gamma \leqslant T$. We thus have

$$S(x,y) \ll \max_{u\leqslant T} \left[\sum_{\substack{\frac{T}{\log^2 T}\leqslant\gamma,\gamma'\leqslant u \\ |\gamma-\gamma'|<\log^3 T}} x^{i(\gamma-\gamma')} w(\gamma-\gamma') \right] +$$

$$+ \sum_{1\leqslant k\leqslant O(\log T)} e^{-k} \max_{u_k\leqslant(k+1)T} \left[\sum_{\substack{kT<\gamma,\gamma'\leqslant u_k \\ |\gamma-\gamma'|<\log^3 T}} x^{i(\gamma-\gamma')} w(\gamma-\gamma') \right] +$$

$$+ O(T) \quad .$$

We can now apply Montgomery's theorem and Conjecture C to the sums in the brackets above. First,

$$\sum_{\substack{kT<\gamma,\,\gamma'\leq u_k \\ |\gamma-\gamma'|<\log^3 T}} = \sum_{\substack{0<\gamma,\,\gamma'\leq u_k \\ |\gamma-\gamma'|<\log^3 T}} - \sum_{\substack{0<\gamma,\,\gamma'\leq kT \\ |\gamma-\gamma'|<\log^3 T}} + O(\log^8 T) .$$

A similar result holds for the sum in the first bracket. Next, we can drop the condition $|\gamma-\gamma'| < \log^3 T$ with an error $O(T)$, as before. Applying (1.23) with T replaced by u_k or kT, we have for $T \leqslant x \leqslant T^{2+\delta'}$, $(\delta' > 0)$

$$S(x,y) \ll T \log T + \sum_{1\leqslant k\leqslant O(\log T)} e^{-k}(u_k \log u_k + kT \log T)$$

$$\ll T \log T \quad .$$

Combining this with (3.3), the result follows.

Lemma 2. Let $1 \leqslant H \leqslant x^{3/4}$, $\frac{1}{y} = T = \frac{x}{H}$. If $\psi(u+H) - \psi(u) > \beta H$ for some u in $[\frac{x}{3}, \frac{3}{2}x]$, then, for all τ in $[u,u+H]$,

$$(3.10) \quad \sum_{n=2}^{\infty} \frac{\Lambda(n)}{n} \frac{y}{(\log \tau/n)^2 + y^2} > \frac{1}{20}\beta ,$$

as $x \to \infty$.

Proof. For all τ and n in $[u,u+H]$,

$$(\log \frac{\tau}{n})^2 \leqslant \log^2(\frac{u+H}{u}) \leqslant (\frac{H}{u})^2 \leqslant 9(\frac{H}{x})^2 = 9y^2 .$$

Hence $$\sum_{n=2}^{\infty} \frac{\Lambda(n)}{n} \frac{y}{(\log \tau/n)^2 + y^2} \geqslant \sum_{u \leqslant n \leqslant u+H} \frac{\Lambda(n)}{n} \frac{y}{(\log \tau/n)^2 + y^2}$$

$$\geqslant \frac{1}{10y(u+H)} \sum_{u \leqslant n \leqslant u+H} \Lambda(n)$$

$$> \frac{1}{20xy} \cdot \beta H = \frac{\beta}{20} .$$

Lemma 3. Let $1 \leqslant H \leqslant x^{3/4}$, $\frac{1}{y} = T = \frac{x}{H}$, and ω be a large absolute constant. Then, as $x \to \infty$, if $1 \leqslant H \leqslant \log^3 x$,

$$\sum_{\substack{\frac{x}{2} \leqslant p_n \leqslant x \\ 2\omega H \geqslant p_{n+1} - p_n \geqslant \omega H}} (p_{n+1} - p_n) <$$

$$< \omega \int_{\frac{x}{e}}^{ex} \left| \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n} \frac{y}{(\log \tau/n)^2 + y^2} - \pi + O\left(\frac{1}{\log \tau}\right) \right|^2 d\tau + 2\omega H x^{\frac{1}{2}} .$$

If $\log^3 x < H \leqslant x^{3/4}$,

$$\sum_{\substack{\frac{x}{2} \leqslant p_n \leqslant x \\ 2\omega H \geqslant p_{n+1} - p_n \geqslant \omega H}} (p_{n+1} - p_n) <$$

$$< \omega \int_{\frac{x}{e}}^{ex} \left| \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n} \frac{y}{(\log \tau/n)^2 + y^2} - \pi + O\left(\frac{1}{\log \tau}\right) \right|^2 d\tau .$$

It is convenient to prove this in the form

$$(3.11)\qquad \sum_{\substack{\frac{x}{2}\leqslant p_n \leqslant x \\ 2\omega H \geqslant p_{n+1}-p_n \geqslant \omega H}} (p_{n+1} - p_n) <$$

$$< \omega \int_{\frac{x}{e}}^{ex} \left| \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n} \frac{y}{(\log \tau/n)^2 + y^2} - \pi + c\left(\frac{1}{\log \tau}\right) \right|^2 d\tau + $$

$$+ 2\omega E_H(x)$$

where

$$(3.12)\qquad E_H(x) = \begin{cases} Hx^{\frac{1}{2}} & \text{if } 1 \leqslant H \leqslant \log^3 x \\ 0 & \text{if } \quad H > \log^3 x \end{cases} .$$

We first describe the idea of the proof. Suppose $\frac{x}{2} \leqslant p_k \leqslant x$, $2\omega H \geqslant p_{k+1} - p_k \geqslant \omega H$ and $T = \frac{x}{H}$. Then for $\tau \in [\, p_k + \frac{\omega-1}{2} H\,,\, p_{k+1} - \frac{\omega-1}{2} H\,]$ and ω sufficiently large, we expect

$$(3.13)\qquad \sum_{n=2}^{\infty} \frac{\Lambda(n)}{n} \frac{y}{(\log \tau/n)^2 + y^2} \leqslant 1 \quad .$$

If this is true, then

$$\int_{p_k+\frac{\omega-1}{2}H}^{p_{k+1}-\frac{\omega-1}{2}H} \left| \sum_{n=2}^{\infty} \frac{\Lambda(n)}{n} \frac{y}{(\log \tau/n)^2 + y^2} - \pi + O\left(\frac{1}{\log \tau}\right) \right|^2 d\tau \geqslant$$

$$\geqslant \int_{p_k+\frac{\omega-1}{2}H}^{p_{k+1}-\frac{\omega-1}{2}H} 4\,d\tau \geqslant 4H \geqslant \frac{2}{\omega}(p_{k+1} - p_k) \quad .$$

Summing over k, for $\frac{x}{2} \leqslant p_k \leqslant x$, we obtain Lemma 3 since the intervals being integrated over are disjoint. The difficulty with this approach is (3.13) may not be true when H is small, i.e. $H \leqslant \log^3 x$. This is because the sum in (3.13) does not chop off terms outside the interval fast enough. To guarantee (3.13) holds, we need to take longer intervals, i.e. $p_{k+1} - p_k \geqslant \omega H \log x/\log H$. This is Cramer's method and the result is (1.8). To overcome this problem we use the following idea: the reason (3.13) fails is because there is at least one interval of length H near $[p_k, p_{k+1}]$ with an exceptionally large number of primes in it. Hence, if we consider τ in this interval, then the integral in (3.11), integrated over this interval will be $\gg H$ by Lemma 2. Thus, the prime intervals $[p_k, p_{k+1}]$ where (3.13) fails will make a contribution to the integral indirectly from a nearby interval with a large number of primes. We still need to show one interval with a large number of primes does not cause (3.13) to fail for too many different prime intervals $[p_k, p_{k+1}]$. It turns out that the more failures induced by such an interval also induces a contribution to the integral, so this is no problem. Finally, prime powers can effect (3.13) if $H \leqslant \log^3 x$. However, since there are $O(x^{\frac{1}{2}})$ prime powers in $[\frac{x}{2}, x]$, we can ignore all interval $[p_k, p_{k+1}]$ which contain any prime powers, with a negligible error.

Proof of Lemma 3. Take $\frac{1}{y} = T = \frac{x}{H}$, $x \to \infty$, and

(3.14) $$\frac{x}{2} \leqslant p_k \leqslant x, \quad 1 \leqslant H \leqslant x^{3/4}, \quad \text{and} \quad 2\omega H \geqslant p_{k+1} - p_k \geqslant \omega H,$$

where ω is a large absolute constant to be picked later, but $\omega \geqslant 10$. We have, taking $\tau \in [p_k + \frac{\omega-1}{2}, p_{k+1} - \frac{\omega-1}{2}]$,

$$(3.15)\quad \sum_{n=2}^{\infty} \frac{\Lambda(n)}{n} \frac{y}{(\log \tau/n)^2 + y^2} = \sum_{n \leqslant \tau/2} + \sum_{\tau/2 < n \leqslant p_k} + \sum_{p_k < n < p_{k+1}}$$

$$+ \sum_{p_{k+1} \leqslant n < 3\tau/2} + \sum_{n \geqslant 3\tau/2}$$

$$= \Sigma_1 + \Sigma_2 + \Sigma_3 + \Sigma_4 + \Sigma_5 \ .$$

Using the elementary result (see [27])

$$(3.16)\quad \sum_{n \leqslant x} \frac{\Lambda(n)}{n} = \log x + O(1) \quad ,$$

we have

$$\sum\nolimits_1 \ll y \sum_{n \leqslant \tau/2} \frac{\Lambda(n)}{n} \ll y \log \tau \ll \frac{H \log x}{x} \ll \frac{\log x}{x^{\frac{1}{4}}} = o(1) \quad ,$$

and

$$\sum\nolimits_5 \ll y \sum_{j=0}^{\infty} \frac{1}{(j+1)^2} \sum_{2^j(\frac{3\tau}{2}) < n \leqslant 2^{j+1}(\frac{3\tau}{2})} \frac{\Lambda(n)}{n}$$

$$\ll y \sum_{j=0}^{\infty} \frac{1}{(j+1)^2} \left\{ \log(2^{j+1}(\tfrac{3\tau}{2})) - \log(2^j(\tfrac{3\tau}{2})) + O(1) \right\}$$

$$\ll y \sum_{j=0}^{\infty} \frac{1}{(j+1)^2} (\log 2 + O(1)) \ll y = o(1) \quad .$$

We now define A to be the set of all prime intervals $d_k = [p_k, p_{k+1}]$ satisfying (3.14). Next, if B denotes the subset of A such that d_k contains a prime power, we define

$$(3.17)\quad A_1 = A_1(H) = \begin{cases} A \setminus B & \text{if } H \leq \log^3 x \\ A & \text{if } H > \log^3 x \end{cases} .$$

Since the number of prime powers in $[\frac{x}{2}, x]$ is, for x sufficiently large,

$$< \frac{x^{1/2}}{\log x} + \frac{x^{1/3}}{\log x} + \cdots \qquad < x^{1/2} \quad ,$$

we see $|B| < x^{1/2}$ and

$$(3.18)\quad \sum_{d_k \in A} (p_{k+1} - p_k) \leq \sum_{d_k \in A_1} (p_{k+1} - p_k) + 2\omega E_H \quad ,$$

where E_H is defined by (3.12). From now on we consider only $d_k \in A_1$, and use (3.18) to obtain the final result.

First consider Σ_3. If $H \leq \log^3 x$, we have

$$\sum_{d_k \in A_1} {}_3 = 0 \ .$$

If $H \geq \log^3 x$, we have

$$\sum_3 \ll \sum_{2 \leq m \leq O(\log x)} \left\{ \sum_{p_k < p^m < p_{k+1}} \Lambda(n) \frac{1}{xy} \right\}$$

$$\ll \frac{\log x}{xy} \sum_{2 \leq m \leq O(\log x)} \left\{ \sum_{p_k^{1/m} < p < p_k^{1/m} + \frac{2\omega H}{p_k^{1-1/m}}} 1 \right\}.$$

Now if $m \geq 5$, then $\frac{2\omega H}{p_k^{1-1/m}} < 1$ for x large enough. Hence the inner sum has at most one term in it. Thus, for $H \geq \log^3 x$,

$$\sum_3 \ll \frac{\log x}{H} \left(\sum_{m=2}^{4} Hx^{-1+1/m} + \sum_{5 \leq m \leq O(\log x)} 1 \right)$$

$$\ll \frac{\log x}{x^{1/2}} + \frac{\log^2 x}{H}$$

$$\ll \frac{1}{\log x} = o(1) \quad .$$

We conclude $\sum_{d_k \in A_1} {}_3 = o(1)$ for all H.

Next, for $p_k-(j+1)H \leq n \leq p_k-jH$,

$$\left(\log \frac{T}{n}\right)^2 \geq \log^2\left(\frac{p_k + \frac{\omega}{4} H}{p_k - jH}\right) \geq \frac{1}{4}\left(\frac{\frac{\omega}{4} + j\ H}{p_k - jH}\right)^2 \geq \frac{H^2\left(\frac{\omega}{4} + j\right)^2}{4x^2} \quad .$$

Hence

$$\sum_2 < \sum_{0 \leq j < \frac{p_k}{2H}} \left\{ \sum_{p_k-(j+1)H \leq n < p_k - jH} \frac{\Lambda(n)}{n} \frac{y}{(\log \tau/n)^2 + y^2} \right\}$$

$$< \sum_{0 \leq j < \frac{p_k}{2H}} \left\{ \sum_{p_k-(j+1)H \leq n < p_k - jH} \frac{\Lambda(n)}{n} \left(\frac{4x^2 y}{H^2 (\frac{\omega}{4} + j)^2} \right) \right\}$$

$$\leq \frac{16}{H} \sum_{0 \leq j \leq \frac{p_k}{2H}} \frac{1}{(\frac{\omega}{4} + j)^2} \left\{ \sum_{p_k-(j+1)H \leq n < p_k - jH} \Lambda(n) \right\} .$$

Now suppose

$$\text{(3.19)} \qquad \sum_{p_k-(j+1)H \leq n \leq p_k - jH} \Lambda(n) \leq H \left(\frac{\omega}{4} + j \right)^{\frac{1}{2}} , \quad \text{for } j = 0,1,2,\dots .$$

Then we have

$$\sum_2 < \frac{16}{H} \sum_{j=0}^{\infty} H \left(\frac{\omega}{4} + j \right)^{-3/2} < 16 \; \frac{8}{\omega^{3/2}} + \int_0^{\infty} \frac{du}{\left(\frac{\omega}{4} + u\right)^{3/2}}$$

$$\leq 16 \left(\frac{8}{\omega^{3/2}} + \frac{4}{\omega^{1/2}} \right) < 200 \omega^{-1/2} .$$

Next suppose

$$\text{(3.20)} \qquad \sum_{p_{k+1} + jH \leq n \leq p_{k+1} + (j+1)H} \Lambda(n) < H \left(\frac{\omega}{4} + j \right)^{\frac{1}{2}} , \text{ for } j = 0,1,2,\dots .$$

Then we obtain similarly

$$\sum_4 < 200\omega^{-\frac{1}{2}} .$$

We conclude from our estimates and from (3.15), for $d_k \in A_1$, $\tau \in [\, p_k + \frac{\omega-1}{2} H \, , \, p_{k+1} - \frac{\omega-1}{2} H \,]$, and subject to (3.19) and (3.20),

$$\sum_{n=2}^{\infty} \frac{\Lambda(n)}{n} \frac{y}{(\log \tau/n)^2 + y^2} < 400\omega^{-\frac{1}{2}} .$$

Choosing $\omega = (400)^2 = 160{,}000$, we see this sum is less than 1. Hence, for these d_k,

$$(3.21) \quad \int_{p_k + \frac{\omega-1}{2}H}^{p_{k+1} - \frac{\omega-1}{2}H} \left| \sum_{n=2}^{\infty} \frac{\Lambda(n)}{n} \frac{y}{(\log \tau/n)^2 + y^2} - \pi + O\left(\frac{1}{\log x}\right) \right|^2 d\tau$$

$$\geqslant \int_{p_k + \frac{\omega-1}{2}H}^{p_{k+1} - \frac{\omega-1}{2}H} 4d\tau \geqslant 4(p_{k+1} - p_k - \omega H + H) \geqslant 4H .$$

The next step is to divide A_1 into two sets A_2 and A_3, where A_2 is the subset of A_1 composed of d_k's which satisfy (3.19) and (3.20), and A_3 is the set of d_k for which (3.19) or (3.20) fail. If $d_k \in A_3$, then there exists a j such that

$$\psi(p_k - jH) - \psi(p_k - (j+1)H) > H\left(\frac{\omega}{4} + j\right)^{\frac{1}{2}} ,$$

or

$$\psi(p_{k+1} + (j+1)H) - \psi(p_{k+1} + jH) > H\left(\frac{\omega}{4} + j\right)^{\frac{1}{2}} ,$$

or both. Suppose for instance the first holds. Then, for

$$\tau \in [p_k - (j+1)H , p_k - jH] ,$$

we have by Lemma 2,

$$\int_{p_k-(j+1)H}^{p_k-jH} \left| \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n} \frac{y}{(\log \tau/n)^2 + y^2} - \pi + O\left(\frac{1}{\log \tau}\right) \right|^2 d\tau$$

$$\geqslant \int_{p_k-(j+1)H}^{p_k-jH} \left| \frac{1}{20}\left(\frac{\omega}{4} + j\right)^{\frac{1}{2}} - \pi - \varepsilon \right|^2 d\tau \qquad (\varepsilon > 0)$$

$$\geqslant \frac{H}{1600}\left(\frac{\omega}{4} + j\right) \qquad (\geqslant 4H) .$$

The same result holds in the second case with the integration over $p_{k+1} + jH \leqslant \tau \leqslant p_{k+1} + (j+1)H$. Clearly these intervals are disjoint from the $d_k \in A_2$. However we must consider the possibility that these intervals are not disjoint for different $d_k \in A_3$.

For each $d_k \in A_3$ there could be many exceptions to (3.19) or (3.20). We begin by picking, for this d_k, only the smallest j where (3.19) is false, and the smallest j' where (3.20) is false. At least one of these occurs. Call these intervals $\underline{c}_k$ and $\overline{c}_k$ respectively, and let L be the collection of these intervals corresponding to all the $d_k \in A_3$. Hence for each $d_k \in A_3$ we get either one or two c_k's. We comment that by the way we chose them, the c_k's need not be ordered by k, i.e. we can have $k_1 < k_2$ but $\underline{c}_{k_2}$ to the left of $\underline{c}_{k_1}$.

Suppose r of the c_k's intersect each other, in the sense that their union is connected (an interval). Then all these c_k's must lie between two of the $d_k \in A_3$, say between d_i and d_{i+1}. Two of them may be $\bar{c}_i$ and $\underline{c}_{i+1}$, but the rest must come from some other d_k's. Hence for at least one of these c_k's, we must have $j \geqslant \frac{r-2}{2}\omega$, since each d_k is at least of length ωH. Hence, if $r > 2$, we have by (3.22), for this c_k, say c^*,

$$(3.23) \quad \int_{c^*} \left| \sum_{n=2}^{\infty} \frac{\Lambda(n)}{n} \frac{y}{(\log \tau/n)^2 + y^2} - \pi + O\left(\frac{1}{\log \tau}\right) \right|^2 d\tau$$

$$\geqslant \frac{H}{1600}\left(\frac{\omega}{4} + \frac{r-2}{2}\cdot\omega\right) \geqslant \frac{H}{1600}\left(\frac{r}{2} - \frac{3}{4}\right)\omega \geqslant \frac{Hr\omega}{6400} \geqslant 2rH .$$

We now define a set L^* of $c_k \in L$ as follows. From L delete intersecting c_k's, if $r = 2$ just delete either one of them, and if $r > 2$ delete all the c_k except for one with $j \geqslant \frac{r-2}{2}\omega$. What remains is L^*. Now $|L| \geqslant |A_3|$, and if $r(i)$ denotes the number of $c_k \in L$ connected to c_i in the sense used above,

$$(3.24) \quad \sum_{c_i \in L^*} r(i) = |L| \geqslant |A_3| .$$

The $c_k \in L^*$ are disjoint, as are the $d_k \in A_2$, so combining (3.18) and (3.21), we have

$$\sum_{\substack{\frac{x}{2} \leqslant p_n \leqslant x \\ 2\omega H \geqslant p_{n+1}-p_n \geqslant \omega H}} (p_{n+1} - p_n) = \sum_{d_n \in A} (p_{n+1} - p_n)$$

$$\leqslant \sum_{d_n \in A_1} 2\omega H + 2\omega E_H$$

$$\leqslant \sum_{d_n \in A_2} 2\omega H + \sum_{d_n \in A_3} 2\omega H + 2\omega E_H$$

$$\leqslant \omega \sum_{d_n \in A_2} \int_{p_n + \frac{\omega-1}{2}H}^{p_{n+1} - \frac{\omega-1}{2}H} \left| \sum_{n=2}^{\infty} \frac{\Lambda(n)}{n} \frac{y}{(\log \tau/n)^2 + y^2} - \pi + O\left(\frac{1}{\log \tau}\right) \right|^2 d\tau$$

$$+ 2\omega H |A_3| + 2\omega E_H .$$

Lemma 3 now follows by noting, by (3.22), (3.23), and (3.24) (use (3.22) when $r = 1$ or 2)

$$2\omega H |A_3| \leqslant 2\omega H \sum_{c_i \in L^*} r(i)$$

$$\leqslant \omega \sum_{c_i \in L^*} \int_{c_i} \left| \sum_{n=2}^{\infty} \frac{\Lambda(n)}{n} \frac{y}{(\log \tau/n)^2 + y^2} - \pi + O\left(\frac{1}{\log \tau}\right) \right|^2 d\tau .$$

Proof of Theorem 1. Combining Lemma 1 and Lemma 3, we have, by (3.1),

$$\sum_{\substack{\frac{x}{2} \leqslant p_n \leqslant x \\ 2\omega H \geqslant p_{n+1}-p_n \geqslant \omega H}} (p_{n+1} - p_n) \ll T \log T + E_H \ll \frac{x \log x}{H} ,$$

for $1 \leqslant H \leqslant x^{\frac{1}{2}+\delta}$. The result is trivial for $0 < H \leqslant 1$, and for $H \geqslant x^{\frac{1}{2}+\delta}$ it is true by (1.7). Hence the result holds uniformly for $H > 0$. Replace ωH by $H, 2H, 4H, \ldots,$ successively. Adding, we obtain, for $H > 0$,

$$\sum_{\substack{\frac{x}{2} \leqslant p_n \leqslant x \\ p_{n+1}-p_n \geqslant H}} (p_{n+1} - p_n) \ll \frac{x \log x}{H} .$$

Theorem 1 follows on replacing x by $\frac{x}{2}, \frac{x}{4}, \ldots,$ successively, and adding.

4. Results on $F(\alpha)$

We first prove Lemma A. Let x be any real number. Then using (1.16) and (1.17), we have

$$\int_{x-\frac{1}{2}}^{x+\frac{1}{2}} F(\alpha)\,d\alpha \leqslant 2\int_{x-1}^{x+1} F(\alpha)\,(1-|\alpha-x|)\,d\alpha \leqslant$$

$$\leqslant \frac{4\pi}{T\log T}\left|\sum_{0<\gamma,\gamma'\leqslant T} w(\gamma-\gamma')\int_{x-1}^{x+1} e^{i\alpha(\gamma-\gamma')\log T}(1-|\alpha-x|)\,d\alpha\right|$$

$$\leqslant \frac{4\pi}{T\log T}\left|\sum_{0<\gamma,\gamma'\leqslant T} e^{ix(\gamma-\gamma')\log T}\left(\frac{\sin\frac{\gamma-\gamma'}{2}\log T}{\frac{\gamma-\gamma'}{2}\log T}\right)^2 w(\gamma-\gamma')\right|$$

$$\leqslant \frac{4\pi}{T\log T}\sum_{0<\gamma,\gamma'\leqslant T}\left(\frac{\sin\frac{\gamma-\gamma'}{2}\log T}{\frac{\gamma-\gamma'}{2}\log T}\right)^2 w(\gamma-\gamma') \quad .$$

Montgomery proved, assuming the RH, as $T\to\infty$,

$$(4.1)\qquad \sum_{0<\gamma,\gamma'\leqslant T}\left(\frac{\sin\frac{\alpha(\gamma-\gamma')}{2}\log T}{\frac{\alpha(\gamma-\gamma')}{2}\log T}\right)^2 w(\gamma-\gamma') \sim \left(\frac{1}{\alpha}+\frac{\alpha}{3}\right)\frac{T\log T}{2\pi}$$

for fixed α, $0<\alpha<1$. This also holds for $\alpha=1$ from the corresponding result $\alpha=1$ for (1.18) (see Lemma B proved below). Hence, taking $\alpha=1$ in (4.1) we have, for $\varepsilon>0$,

$$\int_{x-\frac{1}{2}}^{x+\frac{1}{2}} F(\alpha)\,d\alpha \leqslant \frac{8}{3}+\varepsilon \quad .$$

This proves Lemma A.

To prove Lemma B, we need an explicit formula due to Montgomery ([21], pg. 185). Suppose $x \geqslant 1$. Then, assuming the RH,

$$(4.2) \quad 2x^{\frac{1}{2}} \sum_{\gamma} \frac{x^{i(\gamma-t)}}{1+(t-\gamma)^2} = -\sum_{n=1}^{\infty} \frac{\Lambda(n)a_n(x)}{n^{it}} + \frac{2x^{1-it}}{\left(\frac{1}{2}+it\right)\left(\frac{3}{2}-it\right)}$$

$$x^{-\frac{1}{2}}(\log(|t|+2)+O(1)) + O(x^{-\frac{3}{2}}(|t|+2)^{-1}) ,$$

where $a_n(x) = \min((n/x)^{\frac{1}{2}}, (x/n)^{\frac{3}{2}})$. For $x \geqslant 1$, $T \geqslant 2$, we denote

$$(4.3) \quad F(x,T) = \sum_{0<\gamma,\gamma'\leqslant T} x^{i(\gamma-\gamma')} w(\gamma-\gamma') \quad , \quad w(u) = \frac{4}{4+u^2} .$$

Thus $F(\alpha) = \left(\frac{T}{2\pi}\log T\right)^{-1} F(T^{\alpha},T)$. Montgomery proved (unconditionally)

$$(4.4) \quad F(x,T) = \frac{2}{\pi}\int_{-\infty}^{\infty} \left| \sum_{0<\gamma\leqslant T} \frac{x^{i\gamma}}{1+(t-\gamma)^2} \right|^2 dt =$$

$$= \frac{2}{\pi}\int_{0}^{T} \left| \sum_{\gamma} \frac{x^{i\gamma}}{1+(t-\gamma)^2} \right|^2 dt + O(\log^3 T) .$$

We now compute

$$\int_{T^{\frac{1}{2}}}^{T} \left| \sum_{n=1}^{\infty} \frac{\Lambda(n)a_n(x)}{n^{it}} \right|^2 dt .$$

First,

$$\int_{T^{1/2}}^{T} \left| 2x^{1/2} \sum_{\gamma} \frac{x^{i(\gamma-t)}}{1+(t-\gamma)^2} \right|^2 dt \ll xT \log^2 T$$

by (3.6), and

$$\int_{T^{1/2}}^{T} \left| \frac{2x^{1-it}}{(1/2+it)(3/2-it)} \right|^2 dt \ll \frac{x^2}{T^{3/2}} \quad ,$$

$$\int_{T^{1/2}}^{T} \left| x^{-1/2}(\log|t|+2)+O(1)) + O(x^{-3/2}(|t|+2)^{-1} \right|^2 dt \ll \frac{T \log^2 T}{x} \quad .$$

We conclude from (4.2) by the Cauchy-Schwarz inequality, for $T^{\varepsilon} \leqslant x \leqslant T^2$,

$$\int_{T^{1/2}}^{T} \left| \sum_{n=1}^{\infty} \frac{\Lambda(n)a_n(x)}{n^{it}} \right|^2 dt = 4x \int_{T^{1/2}}^{T} \left| \sum_{\gamma} \frac{x^{i\gamma}}{1+(\gamma-t)^2} \right|^2 dt + O(Tx) \quad .$$

From (4.4) we conclude, using the estimate (1.36), assuming RH,

$$(4.5) \quad F(x,T) = \frac{1}{2\pi x} \int_{T^{1/2}}^{T} \left| \sum_{n=1}^{\infty} \frac{\Lambda(n)a_n(x)}{n^{it}} \right|^2 dt + O(T) \quad ,$$

for $T^{\varepsilon} \leqslant x \leqslant T^2$. Denote

$$(4.6) \quad A(x,T) = \frac{1}{x} \int_{0}^{T} \left| \sum_{n=1}^{\infty} \frac{\Lambda(n)a_n(x)}{n^{it}} \right|^2 dt \quad ,$$

and

$$(4.7)\quad B(x,T) = \frac{1}{x}\int_{-T}^{T}\left(1-\frac{|t|}{T}\right)\left|\sum_{n=1}^{\infty}\frac{\Lambda(n)a_n(x)}{n^{it}}\right|^2 dt .$$

$B(x,T)$ is an average of $A(x,T)$. We first estimate $B(x,T)$, then use this result to estimate $A(x,T)$. The sum in (4.7) is absolutely convergent, thus we can square the sum and integrate term by term. Hence

$$B(x,T) = \frac{1}{x}\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\Lambda(n)\Lambda(m)a_n(x)a_m(x)\int_{-T}^{T}\left(1-\frac{|t|}{T}\right)e^{it\log m/n}\,dt$$

$$= \frac{T}{x}\sum_{n=1}^{\infty}\Lambda^2(n)a_n^2(x) + \frac{T}{x}\sum_{n\neq m}\sum\Lambda(n)\Lambda(m)a_n(x)a_m(x)\left(\frac{\sin\frac{T}{2}\log\frac{m}{n}}{\frac{T}{2}\log\frac{m}{n}}\right)^2$$

$$= S_1 + S_2 .$$

By the prime number theorem

$$S_1 = \frac{T}{x^2}\sum_{n\leqslant x}\Lambda^2(n)n + Tx^2\sum_{n>x}\frac{\Lambda^2(n)}{n^3}$$

$$= T\log x + O(T) .$$

Hence we have

$$(4.8)\quad B(x,T) = T\log x + O(T) + S_2 .$$

We now prove $S_2 \ll x$. To do this, we need to use a sieve estimate for prime twins. The estimate is (see [10], Th. 3.11 and Th. 5.8.1)

$$(4.9)\quad \sum_{k<n\leqslant x}\Lambda(n)\Lambda(n-k) \ll c(k)x ,$$

where

$$(4.10)\quad c(k) = \prod_{\substack{p \mid k \\ p > 2}} \frac{p-1}{p-2} \quad .$$

We shall also use a special case of Lemma 17.4 in [22]:

$$(4.11)\quad \sum_{k \leqslant y} c(k) \ll y \quad .$$

Now,

$$S_2 = \frac{2T}{x} \sum_{n=1}^{\infty} \sum_{m \leqslant n/2} + \frac{2T}{x} \sum_{n=1}^{\infty} \sum_{n/2 < m < n}$$

$$= \sum\nolimits_1 + \sum\nolimits_2 \quad .$$

First,

$$\sum\nolimits_1 \ll \frac{T}{x} \sum_{n=1}^{\infty} \sum_{m \leqslant n/2} \Lambda(n)\Lambda(m) a_n(x) a_m(x) \cdot \frac{1}{T^2}$$

$$\ll \frac{1}{Tx} \left(\sum_{n=1}^{\infty} \Lambda(n) a_n(x) \right)^2 \ll \frac{x}{T} \ll x$$

by the prime number theorem (or Chebyshev's estimate).

In the range $n/2 \leqslant m < n$, $\log \frac{n}{m} \cong \frac{n-m}{n}$. (Here $f \cong g$ means $f \ll g$ and $g \ll f$.) Hence

$$\sum_2 \ll \frac{T}{x} \sum_{n=1}^{\infty} \sum_{n/2<m<n} \Lambda(n)\Lambda(m) a_n(x) a_m(x) \min\left(1, \frac{n^2}{T^2(n-m)^2}\right) .$$

Letting $m = n - k$, we obtain

$$\ll \frac{T}{x} \sum_{n=1}^{\infty} \sum_{k=1}^{[n/2]} \Lambda(n)\Lambda(n-k) a_n(x) a_{n-k}(x) \min\left(1, \frac{n^2}{T^2k^2}\right).$$

We break this sum into pieces

$$\sum_{n=1}^{\infty} = \sum_{n\leqslant x} + \sum_{x<n\leqslant 2x} + \sum_{j=1}^{\infty} \left\{ \sum_{2^j x<n\leqslant 2^{j+1}x} \right\} .$$

For the first sum we have $a_n(x) = (n/x)^{\frac{1}{2}}$, so that

$$\frac{T}{x} \sum_{n\leqslant x} \ll \frac{T}{x^2} \sum_{n\leqslant x} \sum_{k\leqslant n/2} \Lambda(n)\Lambda(n-k) n^{\frac{1}{2}} (n-k)^{\frac{1}{2}} \min\left(1, \frac{n^2}{T^2k^2}\right)$$

$$\ll \frac{T}{x} \sum_{n\leqslant x} \sum_{k\leqslant n/2} \Lambda(n)\Lambda(n-k) \min\left(1, \frac{n^2}{T^2k^2}\right) .$$

We have, for one part,

$$\sum_{x/2\leqslant n\leqslant x} \sum_{k\leqslant n/2} \Lambda(n)\Lambda(n-k) \min\left(1, \frac{n^2}{T^2k^2}\right) \ll$$

$$\ll \sum_{x/2\leqslant n\leqslant x} \left\{ \sum_{k\leqslant x/T} \Lambda(n)\Lambda(n-k) + \frac{x^2}{T^2} \sum_{x/T<k\leqslant n/2} \frac{\Lambda(n)\Lambda(n-k)}{k^2} \right\} \ll$$

$$\ll \sum_{k<\frac{x}{T}} \sum_{k<n\leqslant x} \Lambda(n)\Lambda(n-k) \quad +$$

$$+ \frac{x^2}{T^2} \sum_{j\leqslant O(\log T)} \left\{ \sum_{\frac{2^j x}{T}\leqslant k<\frac{2^{j+1}x}{T}} \frac{1}{k^2} \sum_{\frac{x}{2}\leqslant n<x} \Lambda(n)\Lambda(n-k) \right\}$$

$$\ll \sum_{k\leqslant\frac{x}{T}} c(k)x + \sum_{j=0}^{\infty} \frac{1}{2^{2j}} \left\{ \sum_{k<\frac{2^{j+1}x}{T}} c(k)x \right\}$$

$$\ll \frac{x^2}{T} \quad .$$

We replace x by x/2 , x/4 , ... It follows that

$$\frac{T}{x} \sum_{n\leqslant x} \ll x \quad .$$

For $x\leqslant n<2x$, we have $x/2\leqslant n-k<2x$ and so $a_n(x)\ll 1$ and $a_{n-k}(x)\ll 1$. Thus, a similar argument as the one above shows

$$\frac{T}{x} \sum_{x<n\leqslant 2x} \ll x \quad .$$

Finally, consider

$$\frac{T}{x} \sum_{2^j x\leqslant n\leqslant 2^{j+1}x} \quad .$$

In this range $a_n(x) \ll 2^{-3j/2}$ and $a_{n-k}(x) \ll 2^{-3j/2}$. Hence

$$\frac{T}{x} \sum_{2^j x \leq n \leq 2^{j+1} x} \ll \frac{T}{x} \sum_{2^j x \leq n \leq 2^{j+1} x} \sum_{k \leq \frac{n}{2}} \Lambda(n)\Lambda(n-k) 2^{-3j} \min\left(1, \frac{2^{2j}x^2}{k^2T^2}\right)$$

$$\ll \frac{T}{2^{3j}x} \sum_{k \leq \frac{2^j x}{T}} \sum_{k < n < 2^{j+1}x} \Lambda(n)\Lambda(n-k) +$$

$$+ \frac{x}{2^j T} \sum_{\frac{2^j x}{T} \leq k < 2^j x} \frac{1}{k^2} \sum_{k < n < 2^{j+1}x} \Lambda(n)\Lambda(n-k)$$

$$\ll \frac{T}{2^{2j}} \sum_{k < \frac{2^j x}{T}} c(k) +$$

$$+ \frac{x^2}{T} \sum_{0 \leq r \leq O(\log T)} \left\{ \sum_{\frac{2^{j+r}x}{T} \leq k < \frac{2^{j+r+1}x}{T}} \frac{c(k)}{k^2} \right\}$$

$$\ll \frac{x}{2^j} + \frac{T}{2^{2j}} \sum_{r=0}^{\infty} 2^{-2r} \left\{ \sum_{k < \frac{2^{j+r+1}x}{T}} c(k) \right\}$$

$$\ll \frac{x}{2^j} + \frac{T}{2^{2j}} \sum_{r=0}^{\infty} \frac{2^{-r}2^j x}{T} \ll \frac{x}{2^j} \quad .$$

Summing over j, we conclude

$$\sum\nolimits_2 \ll x \quad ,$$

and hence $S_2 \ll x$. Thus we have proved, by (4.8),

(4.9) $$B(x,T) = T \log x + O(T) + O(x) \quad .$$

We note, for any $\delta > 0$,

(4.10) $$TB(x,T) - (T-\delta)B(x,T-\delta) \leqslant 2\delta A(x,T) \leqslant (T+\delta)B(x,T+\delta) - TB(x,T) \quad .$$

To see this, we note

$$(T+\delta)B(x,T+\delta) - TB(x,T) = \frac{2}{x}\int_T^{T+\delta} (T+\delta-t)\left|\sum_{n=1}^{\infty} \frac{\Lambda(n)a_n(x)}{n^{it}}\right|^2 dt + 2\delta A(x,T)$$

$$\geqslant 2\delta A(x,T) \quad ,$$

and similarly for the lower estimate.

By (4.9) and (4.10),

$$2T\delta \log x - \delta^2 \log x + O((T+\delta)^2) + O(x(T+\delta)) \leqslant 2\delta A(x,T)$$

and

$$2\delta A(x,T) \leqslant 2T\delta \log x + \delta^2 \log x + O((T+\delta)^2) + O(x(T+\delta)) \quad .$$

We thus have

(4.11) $$|A(x,T) - T \log x| < \frac{\delta \log x}{2} + O(\delta) + O(T) + O(x) + O\left(\frac{T}{\delta}(x+T)\right) \quad .$$

Picking

$$\delta = \begin{cases} T & \text{if } 1 \leqslant x \leqslant 2 \\ \dfrac{T}{(\log x)^{\frac{1}{2}}} & \text{if } 2 \leqslant x \leqslant T \\ \left(\dfrac{Tx}{\log x}\right)^{\frac{1}{2}} & \text{if } \quad x \geqslant T \end{cases} ,$$

we obtain, for $1 \leqslant x$,

$$\text{(4.12)} \quad A(x,T) = T \log x + O(T(\log x)^{\frac{1}{2}}) + O((xT \log x)^{\frac{1}{2}}) + O(x) + O(T) .$$

By (4.5), we have, for $T^{\varepsilon} \leqslant x \leqslant T^2$,

$$F(x,T) = \frac{1}{2\pi} (A(x,T) - A(x,T^{\frac{1}{2}})) + O(T)$$

$$= \frac{T \log x}{2\pi} + O(x) + O((xT \log x)^{\frac{1}{2}}) + O(T(\log x)^{\frac{1}{2}}) \quad .$$

This proves Lemma B, since for $x > T^2$ the result is trivial.

5. Another Method

An alternative method is based on the following lemma pointed out to me by Prof. H. L. Montgomery.

Lemma 4. Assume the Riemann Hypothesis. Let $x \geqslant 1$, $e^{\delta} = 1 + \frac{1}{T}$, and $T \geqslant 4$. Then

$$(5.1)\qquad \int_0^{\infty}\Bigg(\sum_{y<n<y+\frac{y}{T}} \Lambda(n)a_n(x) - x\int_{xe^{-\delta}}^{x} a_v(y)\,\frac{dv}{v}\Bigg)^2 \frac{dy}{y}$$

$$= \frac{4x\delta^2}{\pi}\int_0^{\infty}\left(\frac{\sin\frac{\delta}{2}t}{\frac{\delta}{2}t}\right)^2 \left|\sum_{\gamma}\frac{x^{i\gamma}}{1+(t-\gamma)^2}\right|^2 dt + O(\delta\log^2(\tfrac{1}{\delta}))\ ,$$

where $a_u(w) = \min\left[\left(\frac{u}{w}\right)^{1/2}, \left(\frac{w}{u}\right)^{3/2}\right]$.

To prove this, we use (4.2). Let

$$G_{\delta}(t) = \left(\frac{\sin\frac{\delta}{2}t}{\frac{\delta}{2}t}\right)\left(\sum_{n=1}^{\infty}\Lambda(n)a_n(x)n^{-it} - \frac{2x^{1-it}}{\left(\frac{1}{2}+it\right)\left(\frac{3}{2}-it\right)}\right).$$

We compute the Fourier transform of $G_{\delta}(t)$. Here

$$\hat{f}(y) = \int_{-\infty}^{\infty} f(u)e(-yu)\,du\ , \quad e(u) = e^{2\pi iu}\ .$$

Since

$$(5.2)\quad \int_{-\infty}^{\infty} e(-\beta t)\,\frac{\sin \frac{\delta}{2} t}{\frac{\delta}{2} t}\,dt = \begin{cases} \frac{2\pi}{\delta} & \text{if } |\beta| < \frac{\delta}{4\pi} \\ \frac{\pi}{\delta} & \text{if } |\beta| = \frac{\delta}{4\pi} \\ 0 & \text{if } |\beta| > \frac{\delta}{4\pi} \end{cases},$$

by the absolute convergence of the series,

$$\hat{G}_\delta(y) = \frac{2\pi}{\delta} \sum_{|y + \frac{\log n}{2\pi}| \leqslant \frac{\delta}{4\pi}}{}' \Lambda(n) a_n(x) -$$

$$- 2x \int_{-\infty}^{\infty} \frac{e\left(-\left(y + \frac{\log x}{2\pi}\right) t\right)}{\left(\frac{1}{2} + it\right)\left(\frac{3}{2} - it\right)} \left(\frac{\sin \frac{\delta}{2} t}{\frac{\delta}{2} t}\right) dt ,$$

where the sum Σ' indicates we take ½ the terms $\left|x + \frac{\log n}{2}\right| = \frac{\delta}{4\pi}$ (if they exist). A residue calculation shows

$$\int_{-\infty}^{\infty} e(-\beta t)\left(\frac{1}{2} + it\right)^{-1}\left(\frac{3}{2} - it\right)^{-1} dt = \pi \min(e^{-3\pi\beta}, e^{\pi\beta}) .$$

Thus the integral above is

$$= - 2x \int_{\infty}^{\infty} \frac{e\left(-\left(y + \frac{\log x}{2}\right) t\right)}{\left(\frac{1}{2} + it\right)\left(\frac{3}{2} - it\right)} \left\{ \frac{2\pi}{\delta} \int_{-\frac{\delta}{4\pi}}^{\frac{\delta}{4\pi}} e(-tu)\,du \right\} dt$$

$$= - \frac{4\pi x}{\delta} \int_{-\frac{\delta}{4\pi}}^{\frac{\delta}{4\pi}} \left\{ \int_{-\infty}^{\infty} e\left(-\left(y + u + \frac{\log x}{2\pi}\right) t\right)\left(\frac{1}{2} + it\right)^{-1}\left(\frac{3}{2} - it\right)^{-1} dt \right\} du$$

$$= - \frac{4\pi^2 x}{\delta} \int_{-\frac{\delta}{4\pi}}^{\frac{\delta}{4\pi}} \min\,(x^{-3/2}\, e^{-3\pi y - 3\pi u}, x^{1/2} e^{\pi y + \pi u})\, du \quad ,$$

letting $v = xe^{2\pi u}$,

$$= - \frac{2\pi x}{\delta} \int_{xe^{-\delta/2}}^{xe^{\delta/2}} \min\left(\left(\frac{v}{e^{-2\pi y}}\right)^{1/2}, \left(\frac{e^{-2\pi y}}{v}\right)^{3/2}\right) \frac{dv}{v}$$

$$= - \frac{2\pi x}{\delta} \int_{xe^{-\delta/2}}^{xe^{\delta/2}} a_v(e^{-2\pi y})\, \frac{dv}{v} \quad .$$

We thus have

$$\hat{G}_\delta(y) = \frac{2\pi}{\delta} \sideset{}{'}\sum_{|y + \frac{\log n}{2\pi}| < \frac{\delta}{4\pi}} \Lambda(n) a_n(x) - \frac{2\pi x}{\delta} \int_{xe^{-\delta/2}}^{xe^{\delta/2}} a_v(e^{-2\pi y})\, \frac{dv}{v} \quad .$$

By Parseval's theorem, we conclude, by (4.2),

$$(5.3) \quad \int_{-\infty}^{\infty} \left(\frac{2\pi}{\delta} \sum_{|y + \frac{\log n}{2\pi}| < \frac{\delta}{4\pi}} \Lambda(n) a_n(x) - \frac{2\pi x}{\delta} \int_{xe^{-\delta/2}}^{xe^{\delta/2}} a_v(e^{-2\pi y})\, \frac{dv}{v}\right)^2 dy$$

$$= \int_{-\infty}^{\infty} \left(\frac{\sin \frac{\delta}{2} t}{\frac{\delta}{2} t}\right)^2 \left| - 2x^{1/2} \sum_{\gamma} \frac{x^{i\gamma - it}}{1 + (t-\gamma)^2} + x^{-1/2}(\log(|t|+2) + O(1)) + O\left(\frac{x^{-3/2}}{|t|+2}\right)\right|^2 dt \quad .$$

We have dropped the dash since $\Sigma = \Sigma'$ a.e.. On making the change of variables $y \to \frac{-1}{2\pi}(\log y + \frac{\delta}{2})$ in the left-hand side of (5.3), we obtain the left side of (5.1) multiplied by $\frac{2\pi}{\delta^2}$. Next

$$\int_{-\infty}^{\infty} \left(\frac{\sin\frac{\delta}{2}t}{\frac{\delta}{2}t}\right)^2 \left| x^{-1/2}(\log(|t|+2)+O(1)) + O(x^{-3/2}(|t|+2)^{-1})\right|^2 dt$$

$$\ll x^{-1}\int_{-\infty}^{\infty}\left(\frac{\sin v}{v}\right)^2 \log^2\left(\left|\frac{v}{\delta}\right| + 2\right)\frac{dv}{\delta}$$

$$\ll \frac{x^{-1}}{\delta}\log^2\left(\frac{1}{\delta}\right) .$$

Next, by (3.6),

$$\int_{-\infty}^{\infty} \left(\frac{\sin\frac{\delta}{2}t}{\frac{\delta}{2}t}\right)^2 \left| - 2x^{\frac{1}{2}}\sum_{\gamma}\frac{x^{i\gamma-it}}{1+(t-\gamma)^2}\right|^2 dt$$

$$\ll x\int_{-\infty}^{\infty}\left(\frac{\sin v}{v}\right)^2 \log^2\left(\left|\frac{v}{\delta}\right| + 2\right)\frac{dv}{\delta}$$

$$\ll \frac{x\log^2\left(\frac{1}{\delta}\right)}{\delta} .$$

Thus, by the Cauchy Schwarz inequality, the right side of (5.3) is

$$= 4x\int_{\infty}^{\infty} \left(\frac{\sin\frac{\delta}{2}t}{\frac{\delta}{2}t}\right)^2 \left| \sum_{\gamma}\frac{x^{i\gamma}}{1+(t-\gamma)^2}\right|^2 dt + O\left(\frac{1}{\delta}\log^2\left(\frac{1}{\delta}\right)\right) .$$

The lemma now follows on noting the symmetry of the zeros with the real axis implies the integrand above is an even function.

We now apply Conjecture C in a slightly more general form than previously stated. Let $Q(T) \uparrow \infty$ as $T \to \infty$, where $Q(T)$ will be specified later in different ways. We then conjecture

$$(5.4) \qquad F(x,T) \ll T \log T \qquad \text{uniformly for } T \leqslant x \leqslant TQ(T) \quad .$$

Here $F(x,T)$ is defined by (4.3). Conjecture C is equivalent to (5.4) when $Q(T) = T^{1+\delta}$. We note for future use that (5.4) implies, for any $\alpha \geqslant 1$, and the same T,

$$(5.5) \qquad F(x,\alpha T) \ll \alpha T \log T \qquad \text{uniformly for } T \leqslant x \leqslant TQ(T) \quad . ^{*}$$

This is trivial by (1.36) when $\alpha \geqslant \log T$. For $\alpha < \log T$ it follows by replacing T by αT in (5.4) and noting (5.4) holds for $T \leqslant x \leqslant \alpha T$ by Lemma B.

Lemma 5. Assuming (5.4), we have the right hand side of (5.1) is

$$\ll \quad \frac{x \log T}{T} \quad ,$$

uniformly for $T \leqslant x \leqslant TQ(T)$.

Since $\delta \sim \frac{1}{T}$ as $T \to \infty$, the right side of (5.1) is

$$\ll \quad \frac{x}{T^2} \int_0^\infty \left(\frac{\sin \frac{\delta}{2} t}{\frac{\delta}{2} t} \right)^2 \left| \sum_\gamma \frac{x^{i\gamma}}{1 + (t-\gamma)^2} \right|^2 dt + O\left(\frac{\log^2 T}{T} \right) \quad .$$

By (4.4),

* See corrigendum after page 72.

$$\int_0^\infty \left(\frac{\sin\frac{\delta}{2}t}{\frac{\delta}{2}t}\right)^2 \left|\sum_\gamma \frac{x^{i\gamma}}{1+(t-\gamma)^2}\right|^2 dt$$

$$\ll \int_0^T \left|\sum_\gamma \frac{x^{i\gamma}}{1+(t-\gamma)^2}\right|^2 dt + \sum_{k=1}^\infty \int_{2^{k-1}T}^{2^kT} \left(\frac{\sin\frac{\delta}{2}t}{\frac{\delta}{2}t}\right)^2 \left|\sum_\gamma \frac{x^{i\gamma}}{1+(t-\gamma)^2}\right|^2 dt$$

$$\ll F(x,T) + O(\log^3 T) + \sum_{k=1}^\infty \frac{1}{2^{2k}} \int_0^{2^kT} \left|\sum_\gamma \frac{x^{i\gamma}}{1+(t-\gamma)^2}\right|^2 dt$$

$$\ll T\log T + \sum_{k=1}^\infty \frac{1}{2^{2k}} \{F(x,2^kT) + O(k^3\log^3 T)\}$$

$$\ll T\log T + \sum_{k=1}^\infty \frac{1}{2^{2k}} (2^kT\log T)$$

$$\ll T\log T \quad ,$$

for $T \leqslant x \leqslant TQ(T)$, by (5.4) and (5.5). Lemma 5 now follows.

Alternative Proof of Theorem 1. Assuming Conjecture C, we can take $Q(T) = T^{1+\varepsilon}$, for some $\varepsilon > 0$, in (5.4) and Lemma 5. Combining this with Lemma 4, we have, for $T \leqslant x \leqslant T^{2+\varepsilon}$,

$$\int_{\frac{x}{2}}^{2x} \left| \sum_{y<n<y+\frac{y}{T}} \Lambda(n)a_n(x) - x\int_{xe^{-\delta}}^{x} a_v(y)\frac{dv}{v}\right|^2 \frac{dy}{y} \ll \frac{x\log T}{T} \quad .$$

Next, for v, y, and n in $[\frac{x}{2}, 2x]$, we have $a_n(x)$ and $a_v(y)$ are $\cong 1$.

Hence

$$\sum_{y<n<y+\frac{y}{T}} \Lambda(n) a_n(x) = \sum_{y<p<y+\frac{y}{T}} \log p \, a_p(x) + O\left(\frac{y^{\frac{1}{2}}\log y}{T}\right) .$$

Take $T = \frac{3x}{H}$, and suppose $\frac{x}{2} \leqslant p_n \leqslant x$ and $p_{n+1} - p_n \geqslant H$. Then

$$\int_{p_n}^{p_{n+1}-\frac{H}{2}} \left| \sum_{y<p<y+\frac{y}{T}} \log p \, a_p(x) + O\left(\frac{y^{\frac{1}{2}}\log y}{T}\right) - x\int_{xe^{-\delta}}^{x} a_v(y) \frac{dv}{v} \right|^2 \frac{dy}{y}$$

$$= \int_{p_n}^{p_{n+1}-\frac{H}{2}} \left| O\left(\frac{y^{\frac{1}{2}}\log y}{T}\right) - x\int_{xe^{-\delta}}^{x} a_v(y) \frac{dv}{v} \right|^2 \frac{dy}{y}$$

$$\gg \frac{1}{x} \int_{p_n}^{p_{n+1}-\frac{H}{2}} \left| \frac{x}{T} + O\left(\frac{x^{\frac{1}{2}}\log x}{T}\right) \right|^2 dy$$

$$\gg \frac{x}{T^2} (p_{n+1} - p_n) \quad , \quad \text{as } T \to \infty .$$

Combining these estimates and summing over the p_n, we obtain, since the intervals $[p_n, p_{n+1} - \frac{H}{2}]$ are disjoint,

$$\sum_{\substack{\frac{x}{2} \leqslant p_n \leqslant x \\ p_{n+1} - p_n \geqslant H}} (p_{n+1} - p_n) \ll T \log T \ll \frac{x \log x}{H} .$$

This holds for $T \leqslant x \leqslant T^{2+\varepsilon}$, which implies $3 \leqslant H \leqslant x^{\frac{1}{2}+\varepsilon}$. For $0 < H \leqslant 3$ this is trivial and for $H > x^{\frac{1}{2}+\varepsilon}$ this holds by (1.7). Hence, the result holds uniformly for $H > 0$. Replacing x by $\frac{x}{2}, \frac{x}{4}, \ldots$ successively and adding proves Theorem 1.

We now prove a result of Heath-Brown [15].

Theorem 2. Assume the Riemann Hypothesis and the conjecture (5.4) holds in the range $T \leqslant x \leqslant Tf(T)$. If $f(x)/\log x \to \infty$, then almost all intervals $[x, x+f(x)]$ contain a prime.

Proof. Let $S = \{y \mid x \leqslant y \leqslant 2x$, there does not exist a prime in $[y, y+\frac{y}{T}]\}$. $S^* = \{y \mid x \leqslant y \leqslant 2x$, there does not exist any primes or prime powers in $[y, y+\frac{y}{T}]\}$. Clearly,

$$|S| = |S^*| + O\left(\frac{x^{1/2}y}{T}\right) = |S^*| + O\left(\frac{x^{3/2}}{T}\right) .$$

By Lemmas 4 and 5 we have, for $T \leqslant x \leqslant TQ(T)$

$$\int_0^\infty \left| \sum_{y<n<y+\frac{y}{T}} \Lambda(n) a_n(x) - x \int_{xe^{-\delta}}^{x} a_v(y) \frac{dv}{v} \right|^2 \frac{dy}{y} \ll \frac{x \log T}{T} .$$

We note $a_v(y) \cong 1$ for $y \in S^*$ and $xe^{-\delta} \leqslant v \leqslant x$, thus the left hand side is

$$\geqslant \int_{S^*} \left| x \int_{xe^{-\delta}}^{x} a_v(y) \frac{dv}{v} \right|^2 \frac{dy}{y}$$