Honors 1: Undergraduate Math Lab[*]
The Circle Method and Germain Primes

Peter Sarnak[†], Steven Miller[‡], Alex Barnett[§]

November 18th, 2002

Courant Institute of Mathematical Sciences
New York University
New York, NY

Abstract

Using the Hardy-Littlewood Circle Method (and assuming no main term contribution from the Minor Arcs), we calculate the expected number of Germain primes. Calculations and notes by Steven Miller.

Contents

[*] Homepage: http://www.math.nyu.edu/~millerj/
[†] E-mail: sarnak@math.princeton.edu
[‡] E-mail: millerj@cims.nyu.edu or sjmiller@math.princeton.edu
[§] E-mail: barnett@nmr.mgh.harvard.edu

1 Preliminaries

1.1 Definitions

Let

$$e(x) \quad = \quad e^{2\pi i x} \tag{1}$$

and

$$\lambda(n) \;=\; \begin{cases} \log p & \text{if } n = p \text{ is prime} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Finally, define

$$c_q(a) \;=\; \sum_{\substack{r=1 \\ (r,q)=1}}^{q} e\Big(r\frac{a}{q}\Big). \tag{3}$$

1.2 Partial Summation

Lemma 1.1 (Partial Summation: Discrete Version). *Let* $A_N = \sum_{n=1}^{N} a_n$. *then*

$$\sum_{n=M}^{N} a_n b_n = A_N b_N - A_{M-1} b_M + \sum_{n=M}^{N-1} A_n (b_n - b_{n+1}) \tag{4}$$

Proof. Since $A_n - A_{n-1} = a_n$,

$$\begin{aligned} \sum_{n=M}^{N} a_n b_n &= \sum_{n=M}^{N} (A_n - A_{n-1}) b_n \\ &= (A_N - A_{N-1}) b_N + (A_{N-1} - A_{N-2}) b_{N-1} + \cdots + (A_M - A_{M-1}) b_M \\ &= A_N b_N + (-A_{N-1} b_N + A_{N-1} b_{N-1}) + \cdots + (-A_M b_{M+1} + A_M b_M) - a_{M-1} b_M \\ &= A_N b_N - a_{M_1} b_M + \sum_{n=M}^{N-1} A_n (b_n - b_{n+1}). \end{aligned} \tag{5}$$

□

Lemma 1.2 (Abel's Summation Formula - Integral Version). *Let $h(x)$ be a continuously differentiable function. Let $A(x) = \sum_{n \le x} a_n$. Then*

$$\sum_{n \le x} a_n h(n) = A(x)h(x) - \int_1^x A(u)h'(u)du \tag{6}$$

See, for example, W. Rudin, *Principles of Mathematical Analysis*, page 70.

1.3 Siegel-Walfisz

Theorem 1.3. *[Siegel-Walfisz] Let $C, B > 0$, and let a and q be relatively prime. Then*

$$\sum_{\substack{p \le x \\ p \equiv a(q)}} \log p \ = \ \frac{x}{\phi(q)} + O\Big(\frac{x}{\log^C x}\Big) \tag{7}$$

for $q \le \log^B x$, and the constant above does not depend on x, q or a (ie, it only depends on C and B).

1.4 Germain Integral

Define

$$\begin{aligned} f_{1N}(x) &= \sum_{p_1 \le N} \log p_1 \cdot e(p_1 x) \\ f_{2N}(x) &= \sum_{p_2 \le N} \log p_2 \cdot e(-2p_2 x) \\ f_N(x) &= \sum_{p_1 \le N} \sum_{p_2 \le N} \log p_1 \log p_2 \cdot e\Big((p_1 - 2p_2)x\Big). \end{aligned} \tag{8}$$

Consider

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} f_N(x)e(-x)dx \ = \ \sum_{p_1 \le N} \sum_{p_2 \le N} \log p_1 \log p_2 \int_{-\frac{1}{2}}^{\frac{1}{2}} e\Big((p_1 - 2p_2 - 1)x\Big)dx. \tag{9}$$

Note

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} e\Big((p_1 - 2p_2 - 1)x\Big)dx \ = \ \begin{cases} 1 & \text{if } p_1 - 2p_2 - 1 = 0 \\ 0 & \text{if } p_1 - 2p_2 - 1 \neq 0 \end{cases} \tag{10}$$

Thus, we get a contribution of $\log p_1 \log p_2$ if p_1 and $p_2 = \frac{p_1 - 1}{2}$ are both primes. Thus,

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} f_N(x)e(-x)dx \ = \ \sum_{\substack{p_1 \leq N \\ p_2 = \frac{p_1-1}{2} \text{ prime}}} \log p_1 \log p_2. \tag{11}$$

The above is a weighted counting of Germain primes.

1.5 Major and Minor Arcs

Let B be a positive integer, $Q = \log^B N$, and define the Major Arc $\mathcal{M}_{a,q}$

$$\mathcal{M}_{a,q} \ = \ \Big\{x \in [0,1) : \ \Big|x - \frac{a}{q}\Big| \ < \ \frac{Q}{N}\Big\}. \tag{12}$$

We also add in one interval centered at either 0 or 1, ie, the "interval" (or wrapped-around interval)

$$\Bigg[0, \frac{Q}{N}\Bigg] \ \cup \ \Bigg[1 - \frac{Q}{N}, 1\Bigg]. \tag{13}$$

For convenience, we often use the interval $[-\frac{1}{2}, \frac{1}{2}]$ instead of $[0,1]$, in which case we would have

$$\Bigg[-\frac{1}{2}, -\frac{1}{2} + \frac{Q}{N}\Bigg] \bigcup \Bigg[\frac{1}{2} - \frac{Q}{N}, \frac{1}{2}\Bigg]. \tag{14}$$

For functions that are periodic of period one, we could instead consider

$$\Bigg[\frac{1}{2} - \frac{Q}{N}, \frac{1}{2} + \frac{Q}{N}\Bigg]. \tag{15}$$

The Major Arcs are defined by

$$\mathcal{M} \ = \ \bigcup_{q \leq Q} \ \bigcup_{\substack{a=1 \\ (a,q)=1}}^{q} \mathcal{M}_{a,q}. \tag{16}$$

The Minor Arcs, m, are whatever is *not* in the Major Arcs.

Then

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} f_N(x)e(-x)dx \;=\; \int_{\mathcal{M}} f_N(x)e(-x)dx \;+\; \int_{\mathrm{m}} f_N(x)e(-x)dx. \tag{17}$$

We will assume that there is no net contribution over the minor arcs. Thus, in the sequel we investigate

$$\int_{\mathcal{M}} f_N(x)e(-x)dx. \tag{18}$$

1.6 Reformulation of Germain Integral

$$\begin{aligned} f_{1N}(x) &= \sum_{m_1 \le N} \lambda(m_1) \cdot e(m_1 x) \\ f_{2N}(x) &= \sum_{m_2 \le N} \lambda(m_2) \cdot e(-2m_2 x) \\ f_N(x) &= \sum_{m_1 \le N} \sum_{m_2 \le N} \lambda(m_1)\lambda(m_2) \cdot e\Big((m_1 - 2m_2)x\Big). \end{aligned} \tag{19}$$

We investigate

$$\int_{\mathcal{M}} f_N(x)e(-x)dx. \tag{20}$$

We will show the Major Arcs contribute, up to lower order terms, $T_2 N$, where T_2 is a constant independent of N. The length of the Major Arc $\mathcal{M}_{a,q}$ is $\frac{Q}{N}$. We sum over $(a,q) = 1$ and $q \le Q$. Thus, the total length is bounded by

$$\sum_{q \le Q} q \cdot \frac{Q}{N} \;\ll\; \frac{Q^3}{N} \;\ll\; \frac{\log^B}{N}. \tag{21}$$

By choosing B sufficiently large, we will be able to make all the errors from the Major Arc calculations less than the main term from the Major Arcs. Of course, we have absolutely no control over what happens on the minor arcs, and we will simply assume there is no contribution from the minor arcs.

Thus, on the Major Arc $\mathcal{M}_{a,q}$, success will be in finding a function of size N^2 such that the error from this function to $f_N(x)$ on $\mathcal{M}_{a,q}$ is much smaller than N^2, say N^2 divided by a large power of $\log N$.

Similarly, when we integrate over the Major Arcs, we will find the main terms will be of size N; again, success will be in showing the errors in the approximations are much smaller than N, say N divided by a large power of $\log N$.

We are able to do this because of the Siegel-Walfisz Theorem (Theorem 1.3). Given *any* $B > 0$, we can find a $C > 0$ such that, if $q \leq \log^B N$, then

$$\sum_{\substack{p \leq N \\ p \equiv r(q)}} \log p \ = \ \frac{N}{\phi(q)} + O\Big(\frac{N}{\log^C N}\Big), \tag{22}$$

$(r, q) = 1$. Thus, we can take C enormous, large enough so that even when we multiply by the length of the Major Arcs (of size $\frac{\log^{3B} N}{N}$, we still have something small.

2 $f_N(x)$ and $u(x)$

2.1 $f\left(\frac{a}{q}\right)$

We now calculate $f_N\left(\frac{a}{q}\right)$ for $q \leq \log^B N$.

Up to lower order terms,

$$\begin{aligned}
f_N\Big(\frac{a}{q}\Big) &= \sum_{p_1 \le N} \log p_1 \cdot e\Big(p_1 \frac{a}{q}\Big) \sum_{p_2 \le N} \log p_2 \cdot e\Big(-2p_2 \frac{a}{q}\Big) \\
&= \sum_{r_1=1}^{q} \sum_{\substack{p_1 \le N \\ p_1 \equiv r_1 (q)}} \log p_1 \cdot e\Big(p_1 \frac{a}{q}\Big) \sum_{r_2=1}^{q} \sum_{\substack{p_2 \le N \\ p_2 \equiv r_1 (q)}} \log p_2 \cdot e\Big(-2p_2 \frac{a}{q}\Big) \\
&= \sum_{r_1=1}^{q} e\Big(r_1 \frac{a}{q}\Big) \sum_{r_2=1}^{q} e\Big(r_2 \frac{-2a}{q}\Big) \sum_{\substack{p_1 \le N \\ p_1 \equiv r_1 (q)}} \log p_1 \sum_{\substack{p_2 \le N \\ p_2 \equiv r_2 (q)}} \log p_2 \\
&= \frac{N^2}{\phi^2(q)} \sum_{\substack{r_1=1 \\ (r_1,q)=1}}^{q} e\Big(r_1 \frac{a}{q}\Big) \sum_{\substack{r_2=1 \\ (r_2,q)=1}}^{q} e\Big(r_2 \frac{-2a}{q}\Big) \\
&= \frac{N^2}{\phi^2(q)} c_q(a) c_q(-2a), \qquad (23)
\end{aligned}$$

where the second to last line follows from the Siegel-Walfisz Theorem (Theorem 1.3). We restrict to $(r_i, q) = 1$ because if $(r_i, q) > 1$, there is at most one prime $p_i \equiv r_i \bmod q$.

2.2 $u(x)$

Let

$$u(x) \;=\; \sum_{m_1 \le N} \sum_{m_2 \le N} e\Big((m_1 - 2m_2)x\Big). \qquad (24)$$

We will often look at

$$\frac{c_q(a)c_q(-2a)}{\phi^2(q)} u(x). \qquad (25)$$

Note

$$u(0) \;=\; N^2. \qquad (26)$$

3 $f_N(\alpha) - \frac{c_q(a)c_q(-2a)}{\phi^2(q)} u(\alpha - \frac{a}{q})$, $\alpha \in \mathcal{M}_{a,q}$

Let

$$C_q(a) \;=\; \frac{c_q(a)c_q(-2a)}{\phi^2(q)}. \tag{27}$$

We write α as $\beta + \frac{a}{q}$, $\beta \in \Big[-\frac{Q}{N}, \frac{Q}{N}\Big]$, $Q = \log^B N$. As always, we ignore lower order terms.

Note $f_N(x)$ is approximately $C_q(a)N^2$ for x near $\frac{a}{q}$. We now expand and show $f_N(\alpha)$ is $C_q(a)u\Big(\alpha - \frac{a}{q}\Big)$ plus errors of size $\frac{N^2}{\log^{C-2B} N}$ for $\alpha \in \mathcal{M}_{a,q}$.

3.1 Setup

$$\begin{aligned}
S_{a,q}(\alpha) &= f_N(\alpha) - C_q(a)u\Big(\alpha - \frac{a}{q}\Big) \\
&= \sum_{m_1,m_2\le N} \lambda(m_1)\lambda(m_2)e\Big((m_1-2m_2)\alpha\Big) \;-\; C_q(a)\sum_{m_1,m_2\le N} e\Big((m_1-2m_2)\beta\Big) \\
&= \sum_{m_1,m_2\le N}\Big[\lambda(m_1)\lambda(m_2)e\Big((m_1-2m_2)\frac{a}{q}\Big) - C_q(a)\Big]e\Big((m_1-2m_2)\beta\Big) \\
&= \sum_{m_1\le N}\Bigg[\sum_{m_2\le N}\Big[\lambda(m_1)\lambda(m_2)e\Big((m_1-2m_2)\frac{a}{q}\Big) - C_q(a)\Big]e(-2m_2\beta)\Bigg]e(m_1\beta)
\end{aligned} \tag{28}$$

We now apply Partial Summation multiple times. First, we apply Partial Summation to the m_2-sum:

$$\begin{aligned}
S_{2;a,q} &= \sum_{m_2\le N}\Big[\lambda(m_1)\lambda(m_2)e\Big((m_1-2m_2)\frac{a}{q}\Big) - C_q(a)\Big]e(-2m_2\beta) \\
&= \sum_{m_2\le N} a_{m_2}b_{m_2} \\
&= A_2(N)e(-2N\beta) + 4\pi i\beta\int_0^N \sum_{m_2\le u} a_{m_2}e(-u\beta)du.
\end{aligned} \tag{29}$$

We hit the above with $e(m_1\beta)$, and sum from $m_1 = 1$ to N. We get two pieces:

$$\begin{aligned}
S_{1\sum;a,q} &= \sum_{m_1\le N} A_2(N)e(-2N\beta)\cdot e(m_1\beta) \\
S_{1\int;a,q} &= \sum_{m_1\le N} 4\pi i\beta \int_0^N \sum_{m_2\le u} a_{m_2} e(-u\beta)du\cdot e(m_1\beta) \\
S_{a,q} &= S_{1\sum;a,q} + S_{1\int;a,q}. \qquad (30)
\end{aligned}$$

3.2 $S_{1\sum;a,q}$

$$\begin{aligned}
S_{1\sum;a,q} &= \sum_{m_1\le N} A_2(N)e(-2N\beta)\cdot e(m_1\beta) \\
&= e(-2N\beta)\sum_{m_1\le N} A_2(N)e(m_1\beta) \\
&= e(-2N\beta)\sum_{m_1\le N}\sum_{m_2\le N}\Big[\lambda(m_1)\lambda(m_2)e\Big((m_1-2m_2)\frac{a}{q}\Big)-C_q(a)\Big]e(m_1\beta) \\
&= e(-2N\beta)\Bigg[A_1(N)e(N\beta) \\
&\quad -2\pi i\beta\int_0^N\sum_{m_1\le t}\sum_{m_2\le N}\Big[\lambda(m_1)\lambda(m_2)e\Big((m_1-2m_2)\frac{a}{q}\Big)-C_q(a)\Big]e(t\beta)dt. \qquad (31)
\end{aligned}$$

3.2.1 First Piece

The first piece, the $A_1(N)e(N\beta)$ term, is small for $q \le Q$. Why? We have (up to lower order terms)

$$\begin{aligned}
A_1(N)e(N\beta) &= \sum_{m_1,m_2\le N}\lambda(m_1)\lambda(m_2)e\Big((m_1-2m_2)\frac{a}{q}\Big)-\sum_{m_1,m_2\le N}C_q(a) \\
&= C_q(a)N^2 - N^2C_q(a) \ = \ 0. \qquad (32)
\end{aligned}$$

Thus, because of our choice of functions, the leading terms vanish, and the remaining term is small.

3.2.2 Second Piece

We now study the second piece. Note $|\beta| \leq \frac{Q}{N} = \frac{\log^2 B}{N}$, and $C_q(a) = \frac{c_q(a)}{\phi^2(q)} \frac{c_q(-2a)}{\phi^2(q)}$.

Up to lower order terms, the m_2-sum will leave us with

$$\beta \frac{c_q(-2a)N}{\phi(q)} \int_0^N \sum_{m_1 \leq t} \left[\lambda(m_1) e\Big(m_1 \frac{a}{q}\Big) - \frac{c_q(a)}{\phi(q)} \right] e(t\beta) dt. \tag{33}$$

Note $f_N(x)$ is a multiple of N^2 for x near $\frac{a}{q}$. Thus, we want to make sure the above is well dominated by N^2.

For $t \leq \sqrt{N}$, this is immediate. For $t \geq \sqrt{N}$, using Siegel-Walfisz (Theorem 1.3), we can make the bracketed quantity in the integrant dominated by $\frac{N}{\log^C N}$ for any C when $q \leq \log^B N$. Thus, we integrate a quantity that is at most $\frac{N}{\log^C N}$ over an interval of length N, we multiply by $N\beta \ll Q = \log^B N$.

Thus, choosing C appropriately, the integral contributes $\frac{N^2}{\log^{C-B} N}$, and hence is negligible.

Remark 3.1. *Note, of course, that the contribution is only negligible while* $|\beta| \leq \frac{Q}{N}$.

Lemma 3.2. $S_{1 \sum;a,q}$ *is a lower order correction.*

3.3 $S_{1 \int;a,q}$

We must evaluate

$$S_{1 \int;a,q} \quad = \quad \sum_{m_1 \leq N} 4\pi i \beta \int_0^N \sum_{m_2 \leq u} a_{m_2} e(-u\beta) du \cdot e(m_1 \beta), \tag{34}$$

where

$$a_{m_2} \quad = \quad \Big[\lambda(m_1)\lambda(m_2) e\Big((m_1 - 2m_2)\frac{a}{q}\Big) - C_q(a) \Big]. \tag{35}$$

We bring the sum over m_1 inside the integral and again use Partial Summation.

We will ignore the integration and β for now, as these will contribute $\beta N \ll Q = \log^B N$ times the maximum value of the integrand. We will leave the $e(-u\beta)du$ with this integration.

When $u \leq \sqrt{N}$, we can immediately show the above is a lower order correction. Thus, below we always assume $u \geq \sqrt{N}$.

3.3.1 First Piece

We have

$$\begin{aligned}
S_{1\int\sum;a,q} &= \sum_{\substack{m_1 \leq N \\ m_2 \leq u}} \Big[\lambda(m_1)\lambda(m_2) e\Big((m_1 - 2m_2)\frac{a}{q}\Big) - C_q(a)\Big] e(N\beta) \\
&= e(N\beta)\left[\sum_{\substack{m_1 \leq N \\ m_2 \leq u}} \lambda(m_1)\lambda(m_2) e\Big((m_1 - 2m_2)\frac{a}{q}\Big) - C_q(a) \sum_{\substack{m_1 \leq N \\ m_2 \leq u}} 1\right]. \\
&= e(N\beta)\Bigg[C_q(a)uN - C_q(a)uN + \text{Lower Order Terms}\Bigg], \qquad (36)
\end{aligned}$$

where by the Siegel-Walfisz Theorem (Theorem 1.3), the error in the bracketed quantity is of size $\frac{uN}{\log^C N}$.

We then integrate from $u = \sqrt{N}$ to N and multiply by β, giving a contribution bounded by

$$\beta N \cdot \frac{N^2}{\log^C N} \ll \frac{\log^B}{N} \frac{N^3}{\log^C N} \ll \frac{N^2}{\log^{C-B} N}, \qquad (37)$$

again getting a lower order correction to $f_N(x)$ for x near $\frac{a}{q}$ (remember $f_N(x)$ is of size N^2).

3.3.2 Second Piece

Again, $u \geq \sqrt{N}$, and we have

$$2\pi i\beta \int_0^N \sum_{m_1 \leq t} \left[\sum_{m_2 \leq u} \Big[\lambda(m_1)\lambda(m_2) e\Big((m_1 - 2m_2)\frac{a}{q}\Big) - C_q(a)\Big]\right] e(t\beta) dt. \qquad (38)$$

Again, for $t \leq \sqrt{N}$, the contribution will be a lower order correction. For $t, u \geq \sqrt{N}$,

Again, executing the sum over m_1 and m_2 will give us

$$C_q(a)ut - C_q(a)ut + \text{Lower Order Terms}, \tag{39}$$

with the lower order terms of size $\frac{ut}{\log^C N}$.

Integrating over t (from $\sqrt{N}$ to N), then integrating over u (from $\sqrt{N}$ to N) and then multiplying by β^2 gives an error bounded by

$$\beta^2 N^2 \cdot \frac{N^2}{\log^C N} \ \ll\ \frac{\log^{2B} N}{N^2} \frac{N^4}{\log^C N} \ \ll\ \frac{N^2}{\log^{C-2B} N}, \tag{40}$$

again a lower order correction.

4 Integrals of $u(x)$

4.1 Formulations

Remember

$$u(x) \ = \sum_{m_1, m_2 \leq N} e\Big((m_1 - 2m_2)x\Big). \tag{41}$$

We need to study $\int_{-\frac{1}{2}}^{\frac{1}{2}} f_N(x) e(-x) dx$. We have shown that

$$f_N(\alpha) \ = \ C_q(a) u\Big(\alpha - \frac{a}{q}\Big) + O\Big(\frac{N^2}{\log^{C-2B} N}\Big), \quad \alpha \in \mathcal{M}_{a,q}. \tag{42}$$

Thus, we must evaluate

$$\begin{aligned} \int_{\mathcal{M}_{a,q}} u\Big(\alpha - \frac{a}{q}\Big) \cdot e(-\alpha) d\alpha \ &= \ \int_{\frac{a}{q} - \frac{Q}{N}}^{\frac{a}{q} + \frac{Q}{N}} u\Big(\alpha - \frac{a}{q}\Big) \cdot e(-\alpha) d\alpha \\ &= \ \int_{-\frac{Q}{N}}^{\frac{Q}{N}} u(\beta) \cdot e\Big(-\frac{q}{q} - \beta\Big) d\beta \\ &= \ e\Big(-\frac{a}{q}\Big) \int_{-\frac{Q}{N}}^{\frac{Q}{N}} u(\beta) e(-\beta) d\beta. \end{aligned} \tag{43}$$

4.2 $\int_{-\frac{1}{2}}^{\frac{1}{2}} u(x)e(-x)dx$

$$\begin{aligned}\int_{-\frac{1}{2}}^{\frac{1}{2}} u(x)e(-x)dx &= \int_{-\frac{1}{2}}^{\frac{1}{2}} \sum_{m_1,m_2\leq N} e\Big((m_1-2m_2)x\Big)\cdot e(-x)dx \\ &= \sum_{m_1,m_2\leq N} \int_{-\frac{1}{2}}^{\frac{1}{2}} e\Big((m_1-2m_2-1)x\Big)dx. \qquad (44)\end{aligned}$$

If $m_1-2m_2-1=0$, the integral gives 1. There are approximately $\frac{N}{2}$ ways to choose $m_1,m_2\leq N$ such that $m_1-2m_2-1=0$.

Assume now $m_1-2m_2-1\neq 0$. Then the integral vanishes.

Hence,

Lemma 4.1.

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} u(x)e(-x)dx = \frac{N}{2}+O(1). \qquad (45)$$

4.3 $\int_{-\frac{1}{2}}^{-\frac{Q}{N}} + \int_{\frac{Q}{N}}^{\frac{1}{2}} u(x)e(-x)dx$

Define

$$\begin{aligned}I_1 &= \Big[-\frac{1}{2},-\frac{1}{2}+\frac{Q}{N}\Big] \\ I_2 &= \Big[-\frac{1}{2}+\frac{Q}{N},-\frac{Q}{N}\Big] \\ I_3 &= \Big[\frac{Q}{N},\frac{1}{2}-\frac{Q}{N}\Big] \\ I_4 &= \Big[\frac{1}{2}-\frac{Q}{N},\frac{1}{2}\Big] \\ I &= I_1\cup I_2\cup I_3\cup I_4. \qquad (46)\end{aligned}$$

4.4 Integral over I_2, I_3

We have

$$\begin{aligned}\int_{I_i} u(x)e(-x)dx &= \int_{I_i} \sum_{m_1,m_2 \le N} e\Big((m_1 - 2m_2 - 1)x\Big)dx \\ &= \int_{I_i} \sum_{m_1 \le N} e(m_1 x) \sum_{m_2 \le N} e(-2m_2 x) \cdot e(-x)dx \\ &= \int_{I_i} \frac{e(x) - e((N+1)x)}{1 - e(x)} \frac{e(-2x) - e(-2(N+1)x)}{1 - e(-2x)} e(-x)dx. \end{aligned} \tag{47}$$

On I_2 and I_3, the integral is

$$\ll \int_{I_i} \frac{2}{x}\frac{2}{x}dx \ \ll \ \frac{N}{Q} \ = \ \frac{N}{\log^B N}, \tag{48}$$

see, for example, Nathanson (Additive Number Theory: The Classical Bases, Chapter 8).

4.5 Integral over I_1, I_4

Each of these intervals has length $\frac{Q}{N} = \frac{\log^B N}{N}$. There are $\frac{N}{2} + O(1)$ pairs such that $m_1 - 2m_2 - 1 = 0$. Each of these pairs will contribute (bound the integrand by 1) $\frac{Q}{N}$. As there are at most $\frac{N}{2}$ pairs, these contribute at most $\frac{N}{2}\frac{Q}{N} \ll \log^B N$.

Henceforth we assume $m_1 - 2m_2 - 1 \neq 0$. We write

$$I_1 \cup I_4 \ = \ \left[\frac{1}{2} - \frac{Q}{N}, \frac{1}{2} + \frac{Q}{N}\right] \ = \ I'. \tag{49}$$

We have

$$\begin{aligned} & \sum_{\substack{m_1,m_2\le N\\ m_1-2m_2-1\neq 0}} \int_{I'} e\Big((m_1-2m_2-1)x\Big)dx \\ = \;& e\Big(-\frac{1}{2}\Big) \sum_{\substack{m_1,m_2\le N\\ m_1-2m_2-1\neq 0}} (-1)^{m_1} \int_{-\frac{Q}{N}}^{\frac{Q}{N}} e\Big((m_1-2m_2-1)x\Big)dx \\ = \;& e\Big(-\frac{1}{2}\Big)\frac{1}{2\pi i} \sum_{\substack{m_1,m_2\le N\\ m_1-2m_2-1\neq 0}} (-1)^{m_1} \frac{2\sin\Big((m_1-2m_2-1)\frac{Q}{N}\Big)}{m_1-2m_2-1}, \qquad (50) \end{aligned}$$

because, changing variables by sending x to $(x-\frac{1}{2})+\frac{1}{2}$ gives factors of $e\Big((m_1-2m_2-1)\frac{1}{2}\Big) = e(-\frac{1}{2})e(\frac{m_1}{2})e(-m_2)$, and $e(\frac{m_1}{2}) = (-1)^{m_1}$.

4.5.1 $0 < |m_1-2m_2-1| \le N^{1-\epsilon}$

Let $w = m_1-2m_2-1$. We will do the case $0 < w \le N^{1-\epsilon}$, the case with $-N^{1-\epsilon} > w > 0$ being handled similarly.

For each w, there are at most N pairs of m_1, m_2 giving rise to such a w. For such w, $\frac{\sin(w\frac{Q}{N})}{w} \ll \frac{Q}{N}$ (because we are taking the sin of a quantity very close to zero).

Thus, these pairs contribute at most

$$\ll \; N\cdot\frac{Q}{N} \; \ll \; Q \; = \; \log^B N. \qquad (51)$$

Inserting absolute values in Equation 50 gives a contribution of at most $\log^B N$ for such w, $0 < w < N^{1-\epsilon}$.

4.5.2 $N^{1-\epsilon} < |m_1-2m_2-1| \le N$

Again, let $w = m_1-2m_2-1$ and assume $N^{1-\epsilon} < |w| \le N$. We will only consider $w > 0$; $w < 0$ is handled similarly.

The cancellation is due to the presence of the factor $(-1)^{m_1}$; note that for the pair (m_1, m_2) we only care about the parity of m_1.

Consider w and $w-1$.

For $m_1-2m_2-1 = w$, the solutions are

$$\begin{aligned} m_1 &= w+3, \qquad m_2 = 1 \\ m_1 &= w+5, \qquad m_2 = 2 \\ m_1 &= w+7, \qquad m_2 = 3 \end{aligned} \tag{52}$$

and so on; thus there are about $\frac{N-w}{2}$ pairs, all with parity $-(-1)^w$.

For $m_1 - 2m_2 - 1 = w - 1$, we again have about $\frac{N-w}{2}$ pairs, but now the parity is $(-1)^w$. Thus, each of the $\frac{N-w}{2}$ pairs with $m_1 - 2m_2 - 1 = w$ is matched with one of the $\frac{N-w}{2}$ pairs with $m_1 - 2m_2 - 1 = w - 1$, and we are off by at most $O(1)$ pairs, which will contribute

$$\ll \sum_{w=N^{1-\epsilon}}^{N} \frac{1}{w} \ll \log N. \tag{53}$$

For the remaining terms, we subtract in pairs, using the first order Taylor Expansion of $\sin(x)$. We have

$$\sum_{w=N^{1-\epsilon}}^{N} \left[\frac{\sin\left(w\frac{Q}{N}\right)}{w} - \frac{\sin\left(w\frac{Q}{N} - \frac{Q}{N}\right)}{w-1} \right]. \tag{54}$$

The Main Term of the Taylor Expansion gives $\ll \frac{1}{w^2}$, which when summed over w gives $\frac{1}{N^{1-\epsilon}}$. As we have about $\frac{N-w}{2} \ll N$ pairs for each w, this contributes at most $N \cdot \frac{1}{N^{1-\epsilon}} \ll N^\epsilon$.

We also have the first order term from the Taylor Expansion:

$$\sin\left(w\frac{Q}{N} - \frac{Q}{N}\right) \;=\; \sin\left(w\frac{Q}{N}\right) + O\left(\frac{Q}{N}\right). \tag{55}$$

This error leads to (remembering there are $\frac{N-w}{2} \ll N$ pairs for each w)

$$\ll N \sum_{w=N^{1-\epsilon}}^{N} \frac{\frac{Q}{N}}{w-1} \ll Q \log N^\epsilon \ll \log^{B+1} N. \tag{56}$$

4.6 Collecting the Pieces

We have shown

$$\begin{aligned}\int_{[-\frac{1}{2},\frac{1}{2}]} u(x)e(-x)dx &= \frac{N}{2}+O(1)\\ \int_{[-\frac{1}{2},\frac{1}{2}]-[-\frac{Q}{N},\frac{Q}{N}]} u(x)e(-x)dx &= O\Big(\frac{N}{\log^B N}\Big). \qquad (57)\end{aligned}$$

Therefore

Lemma 4.2.

$$\int_{-\frac{Q}{N}}^{\frac{Q}{N}} u(x)e(-x)dx = \frac{N}{2}+O\Big(\frac{N}{\log^B N}\Big). \qquad (58)$$

Remembering that we had

$$\begin{aligned}\int_{\mathcal{M}_{a,q}} u\Big(\alpha-\frac{a}{q}\Big)\cdot e(-\alpha)d\alpha &= \int_{\frac{a}{q}-\frac{Q}{N}}^{\frac{a}{q}+\frac{Q}{N}} u\Big(\alpha-\frac{a}{q}\Big)\cdot e(-\alpha)d\alpha\\ &= \int_{-\frac{Q}{N}}^{\frac{Q}{N}} u(\beta)\cdot e\Big(-\frac{q}{q}-\beta\Big)d\beta\\ &= e\Big(-\frac{a}{q}\Big)\int_{-\frac{Q}{N}}^{\frac{Q}{N}} u(\beta)e(-\beta)d\beta, \qquad (59)\end{aligned}$$

we see that

Lemma 4.3.

$$\int_{\mathcal{M}_{a,q}} u\Big(\alpha-\frac{a}{q}\Big)\cdot e(-\alpha)d\alpha = e\Big(-\frac{a}{q}\Big)\frac{N}{2}. \qquad (60)$$

5 Determination of the Main Term

We now calculate the contribution from the Major Arcs. Up to lower order terms,

$$\begin{aligned}
\int_{\mathcal{M}} f_N(x)e(-x)dx &= \sum_{q\le Q}\sum_{\substack{a=1\\(a,q)=1}}^{q}\int_{\frac{a}{q}-\frac{Q}{N}}^{\frac{a}{q}+\frac{Q}{N}} f_N(\alpha)e(-\alpha)d\alpha \\
&= \sum_{q\le Q}\sum_{\substack{a=1\\(a,q)=1}}^{q}\int_{\frac{a}{q}-\frac{Q}{N}}^{\frac{a}{q}+\frac{Q}{N}} C_q(a)u\Big(\alpha-\frac{a}{q}\Big)e(-\alpha)d\alpha \\
&= \sum_{q\le Q}\sum_{\substack{a=1\\(a,q)=1}}^{q} e\Big(-\frac{a}{q}\Big)\int_{-\frac{Q}{N}}^{\frac{Q}{N}} C_q(a)u(\beta)e(-\beta) \\
&= \sum_{q\le Q}\sum_{\substack{a=1\\(a,q)=1}}^{q} C_q(a)e\Big(-\frac{a}{q}\Big)\frac{N}{2} \\
&= \frac{N}{2}\sum_{q\le Q}\sum_{\substack{a=1\\(a,q)=1}}^{q}\frac{c_q(a)c_q(-2a)}{\phi^2(q)}\cdot e\Big(-\frac{a}{q}\Big) \\
&= \frac{N}{2}\sum_{q=1}^{Q}\Bigg[\sum_{\substack{a=1\\(a,q)=1}}^{q} C_q(a)e\Big(-\frac{a}{q}\Big)\Bigg] \\
&= \frac{N}{2}\sum_{q=1}^{Q}\rho_q \\
&= \mathfrak{S}_N\frac{N}{2}, \qquad (61)
\end{aligned}$$

where we have defined

$$\begin{aligned}
c_q(a) &= \sum_{\substack{r=1\\(r,q)=1}}^{q} e\Big(r\frac{a}{q}\Big)\\
C_q(a) &= \frac{c_q(a)c_q(-2a)}{\phi^2(q)}\\
\rho_q &= \sum_{\substack{a=1\\(a,q)=1}}^{q} C_q(a)e\Big(-\frac{a}{q}\Big)\\
\mathfrak{S}_N &= \sum_{\substack{a=1\\(a,q)=1}}^{q} \rho_q. \qquad (62)
\end{aligned}$$

5.1 Properties of $C_q(a)$ and ρ_q

We will follow the presentation of Nathanson (Additive Number Theory: The Classical Bases, Chapter 8 and Appendix A).

5.1.1 $c_q(a)$ is Multiplicative

We follow Nathanson, Pages $320-321$, Theorem $A.23$. Note that we are labeling by r what he labels a, and we are labeling by a what he labels n.

Lemma 5.1. *$c_q(a)$ is multiplicative; ie, if $(q,q')=1$, then $c_{qq'}(a)=c_q(a)c_{q'}(a)$.*

Proof: We have

$$\sum_{\substack{\widetilde{r}=1\\(\widetilde{r},qq')=1}}^{qq'} e\Big(\widetilde{r}\frac{a}{qq'}\Big). \qquad (63)$$

Exercise 5.2. *Show that we can write the $\widetilde{r}$s above as $\widetilde{r}\equiv rq'+r'q$ mod qq', where $1\leq r\leq q$, $1\leq r'\leq q'$, and $(r,q)=(r',q')=1$.*

Thus

$$\begin{aligned}
c_q(a)c_{q'}(a) &= \sum_{\substack{r=1\\(r,q)=1}}^{q} e\Big(r\frac{a}{q}\Big) \sum_{\substack{r'=1\\(r',q')=1}}^{q'} e\Big(r'\frac{a}{q'}\Big)\\
&= \sum_{\substack{r=1\\(r,q)=1}}^{q} \sum_{\substack{r'=1\\(r',q')=1}}^{q'} e\Big(\frac{(rq'+r'q)a}{qq'}\Big)\\
&= \sum_{\substack{\tilde{r}=1\\(\tilde{r},qq')=1}}^{qq'} e\Big(\tilde{r}\frac{a}{q}\Big) \;=\; c_{qq'}(a). \qquad (64)
\end{aligned}$$

5.1.2 $c_q(a)$ for $(a,q)=1$

Exercise 5.3. *Show that*

$$h_d(a) \;=\; \sum_{r=1}^{d} e\Big(r\frac{a}{d}\Big) \;=\begin{cases} d & \textit{if } d|a\\ 0 & \textit{otherwise}\end{cases} \qquad (65)$$

Recall the moebius function:

$$\mu(d) \;=\; \begin{cases} (-1)^r & \text{if } d \text{ is the product of } r \text{ distinct primes}\\ 0 & \text{otherwise}\end{cases} \qquad (66)$$

Exercise 5.4. *Prove*

$$\sum_{d|(r,q)} \mu(d) \;=\; \begin{cases} 1 & \textit{if } (r,q)=1\\ 0 & \textit{otherwise}\end{cases} \qquad (67)$$

Then

$$\begin{aligned}
c_q(a) &= \sum_{\substack{r=1\\(r,q)=1}}^{q} e\Big(r\frac{a}{q}\Big)\\
&= \sum_{r=1}^{q} e\Big(r\frac{a}{q}\Big) \sum_{d|(r,q)} \mu(d)\\
&= \sum_{d|q} \mu(d) \sum_{\substack{r=1\\d|r}}^{q} e\Big(r\frac{a}{q}\Big)\\
&= \sum_{d|q} \mu(d) \sum_{l=1}^{\frac{q}{d}} e\Big(l\frac{a}{\frac{q}{d}}\Big)\\
&= \sum_{d|q} \mu(d) h_{\frac{q}{d}}(a)\\
&= \sum_{d|q} \mu\Big(\frac{q}{d}\Big) h_d(a)\\
&= \sum_{\substack{d|q\\d|a}} \mu\Big(\frac{q}{d}\Big)\cdot d\\
&= \sum_{d|(a,q)} \mu\Big(\frac{q}{d}\Big) d. \qquad (68)
\end{aligned}$$

Note that if $(a,q)=1$, then there is only one term above, namely $d=1$, which yields

$$c_q(a) \;=\; \mu(q) \;\text{ if } (a,q)=1. \qquad (69)$$

Corollary 5.5. *If $q=p^k$, $k\geq 2$ and $(a,q)=1$, then $c_q(a)=0$.*

5.1.3 $C_q(a)$ is Multiplicative

We have shown $c_{qq'}(a) = c_q(a)c_{q'}(a)$ if $(q,q')=1$. Recall the Euler phi-function, $\phi(q)$, is the number of numbers less than q which are relatively prime to q.

Exercise 5.6. *Prove that $\phi(q)$ is multiplicative; ie, if $(q,q')=1$, then $\phi(qq')=\phi(q)\phi(q')$.*

We now have

Lemma 5.7. *$C_q(a)$ is multiplicative.*

Proof: Assume $(q,q')=1$. We have

$$\begin{aligned} C_{qq'}(a) &= \frac{c_{qq'}(a)c_{qq'}(-2a)}{\phi^2(qq')} \\ &= \frac{c_q(a)c_{q'}(a)c_q(-2a)c_{q'}(-2a)}{\phi^2(q)\phi^2(q')} \\ &= \frac{c_q(a)c_q(-2a)}{\phi^2(q)} \cdot \frac{c_{q'}(a)c_{q'}(-2a)}{\phi^2(q')} \\ &= C_q(a)C_{q'}(a). \qquad (70) \end{aligned}$$

5.1.4 ρ_q is Multiplicative

We first prove a needed lemma.

Lemma 5.8. *Consider $C_{q_1}(a_1q_2)$. Then*

$$C_{q_1}(a_1q_2) \ = \ C_{q_1}(a_1) \qquad (71)$$

if $(q_1,q_2)=1$.

Proof:

$$\begin{aligned} C_{q_1}(a_1q_2) &= \sum_{\substack{r_1=1 \\ (r_1,q_1)=1}}^{q_1} e\Big(r_1\frac{a_1q_2}{q_1}\Big) \\ &= \sum_{\substack{r_1=1 \\ (r_1,q_1)=1}}^{q_1} e\Big(r_1q_2\frac{a_1}{q_1}\Big) \\ &= \sum_{\substack{r=1 \\ (r,q_1)=1}}^{q_1} e\Big(r\frac{a_1}{q_1}\Big) \ = \ C_{q_1}(a), \qquad (72) \end{aligned}$$

because $(q_1,q_2)=1$ implies that as r_1 goes through all residue classes that are relatively prime to q_1, so too does $r=r_1q_2$. □

Lemma 5.9. *ρ_q is multiplicative.*

Recall

$$\rho_q \;=\; \sum_{\substack{a=1\\(a,q)=1}}^{q} C_q(a)e\Big(-\frac{a}{q}\Big). \tag{73}$$

Assume $(q_1,q_2)=1$. Then we can write the congruence classes mod q_1q_2 as $a_1q_2+a_2q_1$, with $1\le a_1\le q_1$, $1\le a_2\le q_2$ and $(a_1,q_1)=(a_2,q_2)=1$.

$$\begin{aligned}
\rho_{q_1q_2} \;&=\; \sum_{\substack{a=1\\(a,q_1q_2)=1}}^{q_1q_2} C_{q_1q_2}(a)e\Big(-\frac{a}{q_1q_2}\Big)\\
&=\; \sum_{\substack{a=1\\(a,q_1q_2)=1}}^{q_1q_2} C_{q_1}(a)C_{q_2}(a)e\Big(-\frac{a}{q_1q_2}\Big)\\
&=\; \sum_{\substack{a_1=1\\(a_1,q_1)=1}}^{q_1}\;\sum_{\substack{a_2=1\\(a_2,q_2)=1}}^{q_2} C_{q_1}(a_1q_2+a_2q_1)C_{q_2}(a_1q_2+a_2q_1)e\Big(-\frac{a_1q_2+a_2q_1}{q_1q_2}\Big).
\end{aligned} \tag{74}$$

Exercise 5.10. *With a_1,a_2,q_1,q_2 as above,*

$$C_{q_1}(a_1q_2+a_2q_1) \;=\; C_{q_1}(a_1q_2) \quad \textit{and} \quad C_{q_2}(a_1q_2+a_2q_1) \;=\; C_{q_2}(a_2q_1). \tag{75}$$

Thus, we have

$$\begin{aligned}
\rho_{q_1q_2} \;&=\; \sum_{\substack{a_1=1\\(a_1,q_1)=1}}^{q_1}\;\sum_{\substack{a_2=1\\(a_2,q_2)=1}}^{q_2} C_{q_1}(a_1q_2)C_{q_2}(a_2q_1)e\Big(-\frac{a_1q_2+a_2q_1}{q_1q_2}\Big)\\
&=\; \sum_{\substack{a_1=1\\(a_1,q_1)=1}}^{q_1} C_{q_1}(a_1q_2)e\Big(-\frac{a_1}{q_1}\Big)\sum_{\substack{a_2=1\\(a_2,q_2)=1}}^{q_2} C_{q_2}(a_2q_1)e\Big(-\frac{a_2}{q_2}\Big)\\
&=\; \sum_{\substack{a_1=1\\(a_1,q_1)=1}}^{q_1} C_{q_1}(a_1)e\Big(-\frac{a_1}{q_1}\Big)\sum_{\substack{a_2=1\\(a_2,q_2)=1}}^{q_2} C_{q_2}(a_2)e\Big(-\frac{a_2}{q_2}\Big)\\
&=\; \rho_{q_1}\cdot\rho_{q_2}.
\end{aligned} \tag{76}$$

Thus, ρ_q is multiplicative. □

5.1.5 Calculation of ρ_q

Lemma 5.11. *$\rho_{p^k} = 0$ if $k \geq 2$ and p is a prime.*

Proof: This follows immediately from $C_{p^k}(a) = 0$. □

Lemma 5.12. *If $p > 2$ is prime, $\rho_p = -\frac{1}{(p-1)^2}$.*

Proof:

$$\begin{aligned} \rho_p &= \sum_{\substack{a=1 \\ (a,p)=1}}^{p} C_p(a) e\Big(-\frac{a}{p}\Big) \\ &= \sum_{a=1}^{p-1} \frac{c_p(a)c_p(-2a)}{\phi^2(p)} e\Big(-\frac{a}{p}\Big). \end{aligned} \tag{77}$$

But as $p > 2$, $c_p(a) = c_p(-2a) = \mu(p)$ as $(a,p) = 1$. As $\mu^2(p) = 1$ and $\phi(p) = p - 1$ we have

$$\begin{aligned} \rho_p &= \sum_{a=1}^{p-1} \frac{1}{(p-1)^2} e\Big(-\frac{a}{p}\Big) \\ &= \frac{1}{(p-1)^2}\Big[-e\Big(-\frac{0}{p}\Big) + \sum_{a=0}^{p-1} e\Big(-\frac{a}{p}\Big)\Big] \\ &= -\frac{1}{(p-1)^2}. \end{aligned} \tag{78}$$

Lemma 5.13. *If $p = 2$, then $\rho_2 = 1$.*

Proof:

$$\begin{aligned} \rho_2 &= \sum_{\substack{a=1 \\ (a,2)=1}}^{2} C_2(a) e\Big(-\frac{a}{2}\Big) \\ &= C_2(1) e\Big(-\frac{1}{2}\Big) \\ &= \frac{c_2(1)c_2(-2)}{\phi^2(2)} \cdot e^{-\pi i} \\ &= \frac{e^{\pi i} e^{-2\pi i}}{1^2} \cdot e^{-\pi i} \ = \ 1, \end{aligned} \tag{79}$$

where we have used $c_2(1) = e^{\pi i}$ and $c_2(-2) = e^{-2\pi i}$.

Exercise 5.14. *Prove* $c_2(1) = e^{\pi i}$ *and* $c_2(-2) = e^{-2\pi i}$.

5.2 Determination of $\mathfrak{S}_N$ and $\mathfrak{S}$

Recall

$$\mathfrak{S}_N = \sum_{q \le Q} \rho_q. \tag{80}$$

We define

$$\mathfrak{S} = \sum_{q} \rho_q. \tag{81}$$

Exercise 5.15. *Let* h_q *be any multiplicative sequence (with whatever growth conditions are necessary to ensure the convergence of all sums below). Then*

$$\sum_q h_q = \prod_{p \ prime} \Big(1 + \sum_{k=1}^{\infty} h_{p^k}\Big). \tag{82}$$

5.2.1 $\mathfrak{S}$

We have

$$\begin{aligned} \mathfrak{S} &= \sum_q \rho_q \\ &= \prod_{p \text{ prime}} \Big(1 + \sum_{k=1}^{\infty} \rho_{p^k}\Big) \\ &= \prod_p \Big(1 + \rho_p\Big) \end{aligned} \tag{83}$$

because $\rho_{p^k} = 0$ for $k \geq 2$ and p prime by Lemma 5.11. We have previously shown (see Lemmas 5.12 and 5.13) that $\rho_2 = 1$ and $\rho_p = -\frac{1}{(p-1)}$ for $p > 2$ prime. Therefore

$$\begin{aligned}\mathfrak{S} &= \prod_p \Big(1+\rho_p\Big) \\ &= (1+\rho_2)\prod_{p>2}(1+\rho_p) \\ &= 2\prod_{p>2}\left[1-\frac{1}{(p-1)^2}\right] \\ &= 2T_2, \qquad (84)\end{aligned}$$

where

Definition 5.16 (Twin Prime Constant).

$$T_2 \ = \ \prod_{p>2}\left[1-\frac{1}{(p-1)^2}\right] \ \approx \ .6601618158 \qquad (85)$$

is the twin prime constant.

5.2.2 $\mathfrak{S}_N$

We need to estimate $|\mathfrak{S}-\mathfrak{S}_N|$. As ρ_q is multiplicative and zero if $q = p^k$ $(k \geq 2)$, we see we need only look at sums of ρ_p. As $\rho_p = -\frac{1}{(p-1)^2}$, one can show that the difference between $\mathfrak{S}$ and $\mathfrak{S}_N$ tends to zero as $N \to \infty$.

Thus,

Lemma 5.17.

$$\mathfrak{S} \ = \ 2T_2. \qquad (86)$$

5.3 Number of Germain Primes and Weighted Sums

Combining the above arguments, we have shown that, up to lower order terms,

$$\begin{aligned}\sum_{\substack{p\leq N \\ p,\frac{p-1}{2}\text{ prime}}} \log(p)\cdot\log\left(\frac{p-1}{2}\right) &= \mathfrak{S}\frac{N}{2} \\ &= 2T_2\frac{N}{2} \\ &= T_2N. \qquad (87)\end{aligned}$$

Note that we are counting Germain prime pairs by $\left(\frac{p-1}{2}, p\right)$ and not $(p, 2p+1)$. Such a difference in counting will introduce a factor of 2.

We can pass from this weighted sum to a count of the number of Germain prime pairs $\left(\frac{p-1}{2}, p\right)$ with $p \leq N$.

Again we follow Nathanson, Chapter 8. Define

$$\begin{aligned} \pi_G(N) &= \sum_{\substack{p \leq N \\ p, \frac{p-1}{2} \text{ prime}}} 1 \\ G(N) &= \sum_{\substack{p \leq N \\ p, \frac{p-1}{2} \text{ prime}}} \log(p) \cdot \log\left(\frac{p-1}{2}\right). \end{aligned} \tag{88}$$

Clearly

$$G(N) \leq \log^2 N \cdot \pi_G(N). \tag{89}$$

Therefore,

Lemma 5.18. *Up to lower order terms,*

$$\pi_G(N) \geq \frac{G(N)}{\log^2 N} = \frac{T_2 N}{\log^2 N}. \tag{90}$$

We now provide a bound in the opposite direction.

$$\pi_G(N^{1-\delta}) = \sum_{\substack{p \leq N^{1-\delta} \\ p, \frac{p-1}{2} \text{ prime}}} 1 \ll \frac{N^{1-\delta}}{\log N}. \tag{91}$$

Then

$$\begin{aligned} G(N) &\geq \sum_{\substack{p \geq N^{1-\delta} \\ p,\frac{p-1}{2} \text{ prime}}} \log p \cdot \log\left(\frac{p-1}{2}\right) \\ &= (1-\delta)^2 \log^2 N \sum_{\substack{p \geq N^{1-\delta} \\ p,\frac{p-1}{2} \text{ prime}}} 1 \\ &= (1-\delta)^2 \log^2 N \Big(\pi_G(N) - \pi_G(N^{1-\delta})\Big) \\ &\geq (1-\delta)^2 \log^2 N \pi_G(N) + O\left((1-\delta)^2 \log^2 N \cdot \frac{N^{1-\delta}}{\log N}\right). \end{aligned} \tag{92}$$

Therefore

$$\begin{aligned} \log^2 N \cdot \pi_G(N) &\leq (1-\delta)^{-2} \cdot G(N) + O\left(\log^2 N \cdot \frac{N^{1-\delta}}{\log N}\right) \\ 0 \ \leq \ \log^2 N \cdot \pi_G(N) - G(N) &\leq \Big[(1-\delta)^{-2} - 1\Big] G(N) + O\Big(\log N \cdot N^{1-\delta}\Big). \end{aligned} \tag{93}$$

If $0 < \delta < \frac{1}{2}$, then $(1-\delta)^{-2} - 1 \ll \delta$. We thus have

$$0 \ \leq \ \log^2 N \cdot \pi_G(N) - G(N) \ \ll \ N\left[\delta + O\Big(\frac{\log N}{N^\delta}\Big)\right]. \tag{94}$$

Choose $\delta = \frac{2 \log\log N}{\log N}$. Then we get

$$0 \ \leq \ \log^2 N \cdot \pi_G(N) - G(N) \ \leq \ O\Big(N \frac{\log\log N}{\log N}\Big). \tag{95}$$

Recalling $G(N) \approx T_2 N$ gives

Lemma 5.19.

$$\pi_G(N) \ \leq \ \frac{T_2 N}{\log^2 N}. \tag{96}$$

Combining with the other bound we have finally shown

Theorem 5.20. *Assuming there is no contribution to the main term from the Minor Arcs, up to lower order terms we have*

$$\pi_G(N) \;=\; \frac{T_2 N}{\log^2 N}, \tag{97}$$

where T_2 is the twin prime constant

$$T_2 \;=\; \prod_{p>2}\left[1 - \frac{1}{(p-1)^2}\right] \;\approx\; .6601618158. \tag{98}$$