LIBERTAS MATHEMATICA, VOL. 2, 1982

ON THE NUMBER OF QUADRUPLES OF PRIMES IN ARITHMETIC PROGRESSION, BELOW A GIVEN BOUND

Emil Grosswald

I. INTRODUCTION

The old conjecture that for every positive integer k, there exist k primes that form an arithmetic progression is still open. It has been known for some time (see [1], [2]) that there are infinitely many triplets of primes in arithmetic progression, but even this modest result does not seem to have been shown to hold for $k = 4$.

Let $N_k(x)$ stand for the number of arithmetic progressions of k terms that consist only of primes and with largest term not in excess of x. In order to have a proof of the stated conjecture, it is sufficient to show that, given the positive integer k, it is possible to find x_k such that $N_k(x) > 0$ holds for all $x \geq x_k$.

In [3] it was shown that, on the basis of Hardy and Littlewoods's "Theorem X" of [5], one can prove that, for $x \to \infty$,

$$N_k(x) \sim C_k \frac{x^2}{\log^k x} \tag{1}$$

with an explicitly given value of C_k. Unfortunately, the proof of "Theorem X" uses a certain unproved hypothesis, as Hardy and Littlewood take special pains to make abundantly clear.

Formula (1) can also be obtained by using the independence of the distribution of residue classes modulo distinct primes, as was shown to the author by D. Zagier [7]. Unfortunately, one has to use these probabilistic considerations also over short intervals, so that also this "proof" of (1) has a gap.

In [3] the case $k = 3$ was handled by a third method (essentially the Hardy-Littlewood-Ramanujan circle method, in the Vinogradov-Estermann-Prachar version), which let to a complete proof of (1) and, in fact, of a stronger version of it.

Much of the work for $k = 3$ goes through at least for $k = 4$. The attempt to set up a complete proof along the same lines failed however, due to the difficulty to evaluate a certain double integral over the (two dimensional) "minor intervals". If one assumes, however, that its order is essentially what one would expect it to be (e.g., by analogy with the case $k = 3$), one obtaines by this third approach a strong form of exactly the same result as by the first two methods. Consequently, for $k = 4$, also this third method leads only to a heuristic "proof" of (1), just like the first two approaches. However, it appears rather interesting that three independent approaches, neither of which could be sharpened into a convincing proof, lead to exactly the same result. To this, one may add the following. A computer search was made up to $x = 50{,}000$ (see [4]) and the convergence to the results of that search, i.e. to the exact values of $N_4(x)$, of the values computed from formula (2) of the main Theorem (see Section 2) is quite remarkable. For these reasons and also in view of the fact that the case $k = 4$ presents some technical difficulties not encountered for $k \leq 3$, it may not be inappropriate to record here the results of these investigations in spite of the fact that they were not entirely successful.

2. MAIN RESULT

Let us denote by $E(x)$ the contribution of the minor intervals (see Section 3 for the precise definition of $E(x)$). With this notation, the following theorem holds.

Theorem. For every positive integer M,

$$N_4(x) = C_4 \frac{x^2}{\log^4 x}\left\{1 + \sum_{k=1}^{M} \frac{a_k}{\log^k x} + O\left(\frac{1}{\log^{M+1} x}\right)\right\} + O(E(x)), \qquad (2)$$

where $C_4 = \frac{3}{4} \prod_{p \geq 5} \left\{1 - \frac{3p-1}{(p-1)^3}\right\} \cong .4763 \ldots$,

$a_1 = 5 + \log 4 - \frac{5}{2} \log 3 \cong 3.639763639\ldots$ *and all further coefficients* a_j *are computable; the sum in parenthesis contains the first terms of an asymptotic series.*

Conjecture. For every K, $E(x) = O(x^2 (\log x)^{-K})$.

Corollary. The truth of the conjecture implies that, for $x \to \infty$,

$$N_4(x) = \frac{3}{4} \prod_{p \geq 5} \left\{1 - \frac{3p-1}{(p-1)^3}\right\} \frac{x^2}{\log^4 x}\left\{1 + \sum_{j=1}^{M} \frac{a_j}{\log^j x} + O\left(\frac{1}{\log^{M+1} x}\right)\right\}. \qquad (2')$$

3. PRELIMINARIES AND NOTATIONS

Let $S(\alpha)(= S_x(\alpha)) = \sum_{p \le x} e^{2\pi i p\alpha}$ and consider the double integral

$$I(= I(x)) = \int_0^1 \int_0^1 S(\alpha)S(\beta-2\alpha)S(\alpha-2\beta)S(\beta)\,d\beta\,d\alpha.$$

If we label the running primes in the different sums by p_i $(i = 1,2,3,4)$, we obtain

$$I = \sum_{p_i \le x} \int_0^1 \int_0^1 \exp[2\pi i(p_1\alpha+p_2(\beta-2\alpha)+p_3(\alpha-2\beta)+p_4\beta)]\,d\beta\,d\alpha$$

$$= \sum_{p_i \le x} \int_0^1 \int_0^1 \exp[2\pi i\alpha(p_1-2p_2+p_3)]\,\exp[2\pi i\beta(p_2-2p_3+p_4)]\,d\beta\,d\alpha$$

$$= \sum_{p_i \le x} \int_0^1 \exp[2\pi i\alpha(p_1-2p_2+p_3)]\,d\alpha \int_0^1 \exp[2\pi i\beta(p_2-2p_3+p_4)]\,d\beta.$$

Either one, or the other integral vanishes, unless both exponents vanish simultaneously, i.e., we have either $p_1 = p_2 = p_3 = p_4$, and this occurs $\pi(x)$ times; or else for $p_1 - p_2 = p_2 - p_3 = p_3 - p_4 \ne 0$. This common difference may be either positive, or negative and each of these cases occurs precisely $N_4(x)$ times.

When both exponents vanish, the integrals reduce to unity, so that the corresponding summand of I is one. It follows that $I(x) = \pi(x) + 2N_4(x)$, or, equivalently, that

$$N_4(x) = \frac{1}{2}\,(I(x)-\pi(x)). \tag{3}$$

In order to compute the integral I, the unit intervals are split into "major intervals" centered around rationals with "small" denominators, and their complements, the "minor intervals". Specifically, let $N = [\log^u x]$ ($[z]$ = greatest integer function, u = integer, to be selected later) and consider the fractions a/q, $(a,q) = 1$, with $1 \le q \le N$. If $\tau = N/x$, then the major intervals along the α-axis are defined by $\alpha = a/q + \delta$, with a and q as before and $|\delta| \le \tau$. The two dimensional intervals $\alpha = \frac{a_1}{q_1} + \delta_1$, $\beta = \frac{a_2}{q_2} + \delta_2$, centered at $(\frac{a_1}{q_1}, \frac{a_2}{q_2})$, arranged in some arbitrary, but fixed order, will be indexed by an index $j \in J$. The union of the complementary ("minor") intervals will be denoted by U. With these notations it is well known (see e.g. [6], pp. 179-185) that if we set $\tilde{S}(q,\delta) = \frac{\mu(q)}{\phi(q)} \sum_{2 \le n \le x} e^{2\pi i n\delta}/\log n$,

$$S(\alpha) = S(a/q+\delta) = \tilde{S}(q,\delta) + T(\delta), \tag{4}$$

with

$$T(\delta) = O(x(x|\delta|+1)e^{-c\sqrt{\log x}}), \tag{5}$$

for any real δ with $0 \le |\delta| \le \frac{1}{2}$, and

$$T(\delta) = O(xe^{-c\sqrt{\log x}}) \tag{6}$$

for $|\delta| \le \tau$.

Here and in what follows c stands for some positive constant, not necessarily the same in all expressions. Also,

$$\tilde{S}(q,\delta) = O(\frac{x}{\phi(q)\log x}) \tag{7}$$

for arbitrary real δ;

$$\tilde{S}(q,\delta) = O(\frac{1}{\phi(q)|\delta|}) \tag{8}$$

for $0 < |\delta| \le \frac{1}{2}$.

The proofs of (7) and (8) are essentially found in [6], p. 185, except that we find it preferable not to replace $\phi(q)$ in the denominator by its lower bound $q^{1-\varepsilon}$, as is done in [6].

4. SKETCH OF THE PROOF

Let $\alpha = a_1/q_1 + \delta_1$, $\beta = a_2/q_2 + \delta_2$, $\alpha - 2\beta = (a_1q_2-2a_2q_1)/q_1q_2 + (\delta_1-2\delta_2) = a_3/q_3 + \delta_3$, $(a_3,q_3) = 1$; and $\beta - 2\alpha = (a_2q_1-a_1q_2)/q_1q_2 + (\delta_2-2\delta_1) = a_4/q_4 + \delta_4$, $(a_4,q_4) = 1$. In accordance with an earlier indication, we shall abbreviate temporarily the summations $\sum\limits_{1\le q_1\le N}\ \sum\limits_{1\le q_2\le N}\ \sum\limits_{\substack{1\le a_1\le q_1\\(a_1,q_1)=1}}\ \sum\limits_{\substack{1\le a_2\le q_2\\(a_2,q_2)=1}}$ by $\sum\limits_J$. Whenever this is not sufficiently clear, the symbol $\sum\limits_{q_1,q_2}\ \sum\limits_{a_1,a_2}$ will occasionally by used for the same summation. With this notation,

$$I = \sum_J \int_{-\tau}^{\tau}\int_{-\tau}^{\tau} S(\frac{a_1}{q_1}+\delta_1)S(\frac{a_2}{q_2}+\delta_2)S(\frac{a_3}{q_3}+\delta_3)S(\frac{a_4}{q_4}+\delta_4)d\delta_2 d\delta_1 + E(x), \tag{9}$$

where $E(x) = \iint_U S(\alpha)S(\beta-2\alpha)S(\alpha-2\beta)S(\beta)d\beta d\alpha$.

From here on we proceed successively as follows:

(a) We replace each $S(a/q+\delta)$ by (4), so as to obtain an integrand of the form $\prod\limits_{i=1}^{4} \tilde{S}(q_i,\delta_i) + R$ in (9).

(b) We estimate R by (5) and (6) and compute its contribution to the integral over a major interval and then we sum that contribution over all major intervals $j \in J$; the result, say $E_1(x)$, turns out to be $O(x^2 e^{-c\sqrt{\log x}})$.

(c) We extend the limits of the integrals in (9) from $\pm\tau$ to $\pm 1/2$, thereby introducing a new error, say $E_2(x)$, with $E_2(x) = O(x^2/(\log x)^{1+u})$.

(d) We compute the principal term, given by

$$\sum_J \int_{-1/2}^{1/2}\int_{-1/2}^{1/2}\Big(\prod_{k=1}^{4}\frac{\mu(q_k)}{\phi(q_k)}\Big)\sum_{2\le n_i\le x}\frac{e^{2\pi i(n_1-2n_2+n_3)\delta_1}e^{2\pi i(n_2-2n_3+n_4)\delta_2}}{\log n_1\ \log n_2\ \log n_3\ \log n_4}\,d\delta_2 d\delta_1 .$$

This equals

$$\sum_J\Big(\prod_{k=1}^{4}\frac{\mu(q_k)}{\phi(q_k)}\Big)\sum_{2\le n_i\le x}\frac{1}{\log n_1 \log n_2\ \log n_3\ \log n_4}\int_{-1/2}^{1/2}e^{2\pi i(n_1-2n_2+n_3)\delta_1}d\delta_1 \times$$

$$\int_{-1/2}^{1/2}e^{2\pi i(n_2-2n_3+m_4)\delta_2}d\delta_2 = \sum_J\Big(\prod_{k=1}^{4}\frac{\mu(q_k)}{\phi(q_k)}\Big)\sideset{}{'}\sum_{2\le n_i\le x}\frac{1}{\log n_1\ \log n_2\ \log n_3\ \log n_4},$$

where the dash signifies that one sums only over values n_i that satisfy $n_1 - n_2 = n_2 - n_3 = n_3 - n_4$. As the last sum is independent of the a_j's and q_k's, it may be factored out and the result reads

$$\sideset{}{'}\sum_{2\le n_i\le x}\frac{1}{\log n_1\ \log n_2\ \log n_3\ \log n_4}\ \sum_J\Big(\prod_{k=1}^{4}\frac{\mu(q_k)}{\phi(q_k)}\Big). \tag{10}$$

The first factor of (10) equals

$$2\sideset{}{''}\sum_{2\le n_i\le x}\frac{1}{\log n_1\ \text{lof}\ n_2\ \log n_3\ \log n_4} + O\Big(\frac{x}{\log^4 x}\Big), \tag{11}$$

where $\sum''$ means that the summation condition has been strengthened to $n_1 - n_2 = n_2 - n_3 = n_3 - n_4 > 0$. The sum of (11) (in fact, a more general one) has been studied in [3], where it is shown that it equals

$$\frac{x^2}{6\log^4 x}\Big\{1 + \sum_{k=1}^{M}\frac{a_k}{\log^k x} + O\Big(\frac{1}{\log^{M+1} x}\Big)\Big\};$$

here all coefficients a_k are computable, M is an arbitrary positive integer and the curly brackets enclose an asymptotic series.

(e) The second factor of (10) leads to the "singular series". One observes that it is multiplicative. Hence, one may start from the expression

$$S = \lim_{p\to\infty} \prod_{p \leq P} \sum_{\nu=0}^{\infty} \sum_{\mu=0}^{\infty} \sum_{\substack{1 \leq a_1 \leq p^{\nu} \\ (a_1,p)=1}} \sum_{\substack{1 \leq a_2 \leq P^{\mu} \\ (a_2,P)=1}} \frac{\mu(p^{\nu})}{\phi(p^{\nu})} \frac{\mu(p^{\mu})}{\phi(p^{\mu})} \frac{\mu(q_3)}{\phi(q_3)} \frac{\mu(q_4)}{\phi(q_2)} ,$$

show that it converges, and that it is the limit for $N \to \infty$ of the second factor of (10). The computation yields

$$S = \frac{9}{2} \prod_{p \geq 5} \{1 - \frac{3p-1}{(p-1)^3}\} .$$

By putting together these results, one obtains

$$I = \frac{1}{3} \frac{x^2}{\log^4 x} \cdot \frac{9}{2} \prod_{p \geq 5} \{1 - \frac{3p-1}{(p-1)^3}\} \cdot \{1 + \sum_{k=1}^{M} \frac{a_k}{\log^k x} + O(\frac{1}{\log^{M+1} x})\} + O(x^2 e^{-c\sqrt{\log x}}) +$$
$$O(x^2 (\log x)^{-1-u}) + E(x).$$

On account of (3,, and if we take $u \geq M + 4$, this completes the proofs of (2) and (2').

5. SOME DETAILS OF THE PROOF

The error term R consists of the sum of 15 among the 16 terms of the product $\prod_{i=1}^{4} \{\tilde{S}(q_i,\delta_i)+T(\delta_i)\}$. By using only the trivial estimate $|\tilde{S}| \leq x$ and (5), we obtain $|R| = O(x^4 e^{-c\sqrt{\log x}} (1+x\,\tilde{\delta}))$, $\tilde{\delta} = \max(|\delta_1|,|\delta_2|)$. Integration in both variables over intervals of lengths τ leads to $\int_0^\tau \int_0^\tau |R|\, d\delta_1 d\delta_2 = O((x^4\tau^2+x^5\tau^3)e^{-c\sqrt{\log x}}) = O(x^2 N^3 e^{-c\sqrt{\log x}})$ by the definition of τ. Summing this first over all invervals with given q_1,q_2, for $1 \leq a_1 \leq q_1$, $1 \leq a_2 \leq q_2$, $(a_1,q_1) = (a_2,q_2) = 1$ and then for all q_1,q_2, $1 \leq q_1$, $q_2 \leq N$, we obtain $E_1 = O(x^2 N^7 e^{-c\sqrt{\log x}}) = O(x^2 e^{-c\sqrt{\log x}})$, because $N = O(\log^u x)$.

6. FURTHER DETAILS OF THE PROOF

Denote the integrand of (9) by $V_j(\delta_1,\delta_2)$; then the change of limits of the integrals leads to a certain new error, say e_j for the j-th major interval. By the symmetry between α and β it is sufficient (see Fig. 1) to compute the order of magnitude of the integrals

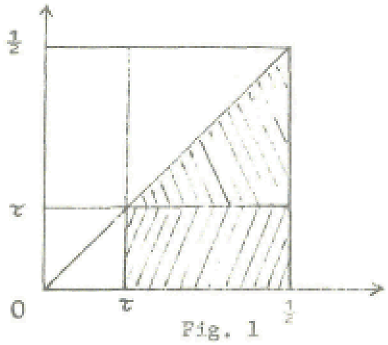

Fig. 1

$\int_{\tau}^{1/2}\int_0^{\delta_1} |V_j(\delta_1,\delta_2)|\, d\delta_2 d\delta_1$. The integrand, however, consists of 4 factors. Each of these has to be estimated either by (7), or by (8), according to the value of the respective δ. In addition to the difficulty with $|\delta| > 1/2$, one also observes that for certain ranges of the values of δ_1 and δ_2, either δ_3, or δ_4 may vanish, or else be so small that (8) is not usable. We therefore define $\eta = \tau/3$ and shall estimate any $\tilde{S}$ by (7), also whenever the respective δ satisfies $|\delta| < \eta$. It is clear that $\tau \le \delta_1 \le 1/2$, so that $\tilde{S}(q_1,\delta_1)$ can be estimated by (8) for all δ_1. Also $|\delta_2| \le 1/2$, so that $S(q_2,\delta_2)$ will be estimated by (8), except for $0 \le \delta_2 \le \eta$, when (7) has to be used. For other two factors one has to divide the rectangle of integration into subrectangles; in some of them, (7), in others (8) will have to be used for the individual factors. Specifically, for δ_3 one observes that $-\frac{1}{2} \le -\delta_1 \le \delta_3 = \delta_1 - 2\delta_2 \le \frac{1}{2} - 2\delta_2 \le \frac{1}{2}$; hence, $|\delta_3| \le \frac{1}{2}$ and we use (8) for $\tilde{S}(q_3,\delta_3)$, except when $|\delta_3| < \eta$, i.e. for $(\delta_1-\eta)/2 < \delta_2 < (\delta_1+\eta)/2$, when we have to use (7). In the same way, by requiring $\delta_4 = \delta_2 - 2\delta_1$ to satisfy $\eta \le |\delta_4| \le \frac{1}{2}$ for use of (8), it follows that if $\tau \le \delta_1 \le \frac{1}{4}$ we may use (8) for $\tilde{S}(q_4,\delta_4)$, while for $\frac{1}{4} < \delta_1 \le \frac{1}{2}$, we have to use (7) if $0 \le \delta_2 \le 2\delta_1 - \frac{1}{2}$, and (8) for $2\delta_1 - \frac{1}{2} < \delta_2 < \delta_1$. Now one has to interleaf the points of subdivision of δ_2 and this requires some further refinement of the subdivisions. So, e.g., $2\delta_1 - \frac{1}{2} < (\delta_1-\eta)/2$ if $\delta_1 < (1-\eta)/3$, but $2\delta_1 - \frac{1}{2} > (\delta_1-\eta)/2$ if $\delta_1 > (1-\eta)/3$. If we set $e_j = \int_{\tau}^{1/2}\int_0^{\delta_1} |V_j(\delta_1,\delta_2)|\, d\delta_2 d\delta_1$, then the total error introduced into I by the change of limits is at most $2 \sum_{j \in J} e_j = E_2(x)$, say. By using previous points of subdivision, we now obtain the following. For the j-th two-dimentional interval, define $\phi_j = \prod_{k=1}^{4} \phi(q_k)$; then

$$e_j\phi_j \le \int_{\tau}^{1/4}(I_1/\delta_1)d\delta_1 + \int_{\frac{1}{4}}^{\frac{1}{4}+\frac{\eta}{2}}(I_2/\delta_1)d\delta_2 + \int_{\frac{1}{4}+\frac{\eta}{2}}^{\frac{1}{3}-\frac{\eta}{3}}(I_3/\delta_1)d\ {}_2 + \int_{\frac{1}{3}-\frac{\eta}{3}}^{\frac{1}{3}+\frac{\eta}{3}}(I_4/\delta_1)d\delta_2 +$$

$$\int_{\frac{1}{3}+\frac{\eta}{3}}^{1/2}(I_5/\delta_1)d\delta_1.$$

Here

$$I_1 = \frac{x}{\log x} \int_0^{\eta} \frac{d\delta_2}{|\delta_3||\delta_4|} + \int_{\eta}^{(\delta_1-\eta)/2} \frac{d\delta_2}{\delta_2|\delta_3||\delta_4|} + \frac{x}{\log x} \int_{(\delta_1-\eta)/2}^{(\delta_1+\eta)/2} \frac{d\delta_2}{\delta_2|\delta_4|} + \int_{(\delta_1+\eta)}^{\delta_1} \frac{d\delta_2}{\delta_2|\delta_3||\delta_4|}$$

and $I_2,\ldots,I_5$ have similar expressions. The integrals that enter the expressions of the I_k are easily obtained. There are all together 24 integrals to be evaluated. All are elementary. It follows from these calculations that $\int_{\tau}^{1/4}(I_1/\delta_1)d\delta_1 = O(x^2(\log x)^{-u-1})$, while all others are $O(x(\log x)^{-u-2})$. Hence, $e_j\phi_j = O(x^2(\log x)^{-u-1})$ and $e_j = \phi_j^{-1}O(x^2(\log x)^{-u-1})$. It should suffice, if we indicate here, as an example, only the computation of $\int_{\tau}^{1/4}(I_1/\delta_1)d\delta_1$.

With $\delta_3 = \delta_1 - 2\delta_2$ and $\delta_4 = \delta_2 - 2\delta_1$, we obtain the following.

$$I_1 = \frac{x}{\log x} \int_0^{\eta} \frac{d\delta_2}{|\delta_3||\delta_3|} + \int_{\eta}^{\frac{\delta_1-\eta}{2}} \frac{d\delta_2}{\delta_2|\delta_3||\delta_4|} + \frac{x}{\log x} \int_{\frac{\delta_1-\eta}{2}}^{\frac{\delta_1+\eta}{2}} \frac{d\delta_2}{\delta_2|\delta_4|} + \int_{\frac{\delta_1+\eta}{2}}^{\delta_1} \frac{d\delta_2}{\delta_2|\delta_3||\delta_4|} =$$

$$\frac{x}{\log x} \cdot \frac{1}{3\delta_1} \log \frac{2\delta_1-\eta}{2\delta_1-4\eta} + \frac{1}{6\delta_1^2} \log \frac{16\eta^7(2\delta_1-\eta)}{(\delta_1-2\eta)^4(\delta_1-\eta)^3(3\delta_1+\eta)} + \frac{x}{\log x}\,\frac{1}{2\delta_1}\,\frac{(3\delta_1-\eta)(\delta_1-\eta)}{(3\delta_1+\eta)(\delta_1+\eta)} +$$

$$\frac{1}{6\delta_1^2} \log \frac{(\delta_1+\eta)^3(3\delta_1-\eta)}{16\eta^4} O\{\frac{x}{\log x}\,\frac{1}{\delta_1} + \frac{1}{\delta_1^2}\log\frac{1}{\eta}\}$$ in $3\eta = \tau \leq \delta_1 \leq 1/4$. Consequently,

$$\int_{\tau}^{1/4}(I_1/\delta_1)d\delta_1 = O(\frac{x}{\log x}\,\frac{1}{\eta}) + O(\frac{1}{\eta 2}\log\frac{1}{\eta}) + O(\frac{x^2}{N\log x}) + O(\frac{x^2}{N^2}\log\frac{x}{N}) = O(x^2(\log x)^{-u-1}) +$$

$$O(x^2(\log x)^{-2u+1}) = O(x^2(\log x)^{-u-1}),$$ for $u \geq 2$.

The value of $E_2(x)$ is found, by summing $2\ e_j$ over all $j \in J$. We obtain $E_2(x) = O(x^2(\log x)^{-u-1}) \sum_{j\in J} \phi_j^{-1}$. We shall verify that the sum over ϕ_j^{-1} converges as $N \to \infty$, so that, finally, $E_2(x) = O(x^2(\log x)^{-u-1})$, as we wanted to show.

In order to show that $\lim_{N\to\infty} \sum_{j\in J} \phi_j^{-1} = \lim_{N\to\infty} \sum_{q_1,q_2\leq N} \sum_{a_1,a_2} \frac{1}{\phi(q_1)\phi(q_2)\phi(q_3)\phi(q_4)}$

converges to a finite value we observe that the functions are multiplicative as functions of q_1,q_2, so that the limit for $N \to \infty$ exists, if

$$\lim_{P\to\infty} \prod_{p\le P} \sum_{\nu=0}^{\infty} \sum_{\mu=0}^{\infty} \sum_{\substack{1\le a_1\le p^\nu \\ (a_1,p)=1}} \sum_{\substack{1\le a_2\le p^\mu \\ (a_2,p)=1}} \frac{1}{\phi(p^\nu)\phi(p^\mu)\phi(q_3)\phi(q_4)}$$

exists and then the two limits are equal. Here q_3 and q_4 are defined for $q_1 = p^\nu$, $q_2 = p^\mu$ as in Section 4.

The exact numerical value of this expression is irrelevant; for this reason, we shall only sketch here the proof that the factor corresponding to $p \ge 5$ equals

$$F(p) = 1 + \frac{5}{p^2}\,\frac{1-(2/5p)+(3/5p^2)}{(1-p^{-1})^2(1-p^{-2})^2}\,. \tag{12}$$

This suffices to show that $\prod_p F(p)$ converges, hence that $\lim_{N\to\infty} \sum_{j\in J} \phi_j^{-1}$ is finite, as we had to prove. The primes $p = 2$ and $p = 3$ require separate considerations and lead to factors 71/9 and 85/32, respectively. The lengthy details of those computations may be suppressed.

For general $p \ge 5$, let us consider first the case $\nu > \mu$. Then $\frac{a_1}{p^\nu} - \frac{2a_2}{p^\mu} = \frac{a_1-2a_2p^{\nu-\mu}}{p^\nu}$, so that $q_3 = p^\nu$. Similarly, $\frac{a_2}{p^\mu} - \frac{2a_1}{p^\nu} = \frac{a_2p^{\nu-\mu}-2a_1}{p^\nu}$ and $q_4 = p^\nu$. Hence, for all $\phi(p^\nu)\phi(p^\mu) = p^{\nu+\mu-2}(p-1)^2$ couples (a_1,a_2), $q_3 = q_4 = p^\nu$ and the sum over $1 \le a_1 \le p^\nu$, $1 \le a_2 \le p$, $p \nmid a_1a_2$ amounts to

$\frac{p^{\mu+\nu-2}(p-1)^2}{\{p^{\nu-1}(p-1)\}^3p^{\mu-1}(p-1)} = \frac{1}{p^{2\nu-2}(p-1)^2}$. For fixed μ and $\nu = \mu+r$ $(r = 1,2,\ldots)$, the total contribution to the sum equals $\sum_{r=1}^{\infty} \frac{1}{p^{2(\mu+r)-2}(p-1)^2} = \frac{p^2}{p^{2\mu}(p-1)^2(p^2-1)}$.

The contribution for all $\nu > \mu \ge 0$ is, therefore $\frac{p^2}{(p-1)^2(p^2-1)} \sum_{\mu=0}^{\infty} \frac{1}{p^{2\mu}} =$

$\frac{p^4}{(p-1)^2(p^2-1)^2}$. By symmetry, we obtain the same contribution from the terms with $\mu > \nu$, so that all terms with $\mu \ne \nu$ contribute a total of $\frac{2p^4}{(p-1)^4(p+1)^2}$.

In a similar way, but with more computation, one shows that the contribution of the terms with $\mu = \nu > 0$ equals $\frac{p^2(p-3)(p+1)+2p^4}{(p-1)^4(p+1)^2}$. The contribution of the term with $\mu = \nu = 0$ is one and, if we add all these contributions, we obtain (12).

7. COMPUTATION OF THE PRINCIPAL TERM

If we combine (9) with the facts that $E_1(x) = O(x^2 e^{-c\sqrt{\log x}})$ and $E_2(x) = O(x^2(\log x)^{-1-u})$, we obtain that

$$I = \sum_{q_1,q_2} \sum_{a_1,a_2} \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} \Big(\prod_{k=1}^{4} \frac{\mu(q_k)}{\phi(q_k)} \Big) \sum_{2 \le n_i \le x} \Big(\prod_{i=1}^{4} \frac{e^{2\pi i n_i \delta_i}}{\log n_i} \Big) d\delta_2 d\delta_1 + O(x^2(\log x)^{-1-u}) + E(x). \qquad (13)$$

If we replace here δ_3 and δ_4 by their values, the multiple sum becomes

$$\sum_{q_1,q_2} \sum_{a_1,a_2} \Big(\prod_{k=1}^{4} \frac{\mu(q_k)}{\phi(q_k)} \Big) \sum_{2 \le n_i \le x} \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} \frac{e^{2\pi i\delta_1(n_1-2n_2+n_3)}\, e^{2\pi i\delta_2(n_2-2n_3+n_4)}}{\log n_1 \log n_2 \log n_3 \log n_4}\, d\delta_2 d\delta_1 .$$

This equals

$$\sum_{q_1,q_2} \sum_{a_1,a_2} \Big(\prod_{k=1}^{4} \frac{\mu(q_k)}{\phi(q_k)} \Big) \sum_{2 \le n_i \le x} \frac{1}{\log n_1 \log n_2 \log n_3 \log n_4} \times \int_{-1/2}^{1/2} e^{2\pi i\delta_1(n_1-2n_2+n_3)} d\delta_1 \int_{-1/2}^{1/2} e^{2\pi i\delta_2(n_2-2n_3+n_4)} d\delta_2 .$$

One or the other of the integrals vanishes, unless we have $n_1 - 2n_2 + n_3 = n_2 - 2n_3 + n_4 = 0$, when both reduce to unity. The principal term reduces to

$$\sideset{}{'}\sum_{2 \le n_i \le x} \frac{1}{\log n_1 \log n_2 \log n_3 \log n_4} \sum_{q_1,q_2} \sum_{q_1,a_2} \Big(\prod_{k=1}^{4} \frac{\mu(q_k)}{\phi(q_k)} \Big) . \qquad (14)$$

By recalling the definition of $\sum'$, we see that the first sum contains (i) terms with $n_1 = n_2 = n_3 = n_4$, i.e., $\sum_{2 \le n \le x} \frac{1}{\log^4 n} \doteq \frac{x}{\log^4 x}$; (ii) terms with positive common difference; and (iii) terms with negative common difference. The contributions of (ii) and (iii) are obviously equal; hence, recalling the definition of $\sum''$, the first sum equals

$$2 \sideset{}{''}\sum_{2 \le n_i \le x} \frac{1}{\log n_1 \log n_2 \log n_3 \log n_4} + O\Big(\frac{x}{\log^4 x}\Big) .$$

The second factor of (14), say S_N, is a sum over the multiplicative product $\prod_{k=1}^{4} \frac{\mu(q_k)}{\phi(q_k)}$. It also turns out to be a convergent sum, as $N \to \infty$; hence, $\lim_{N \to \infty} S_N = S$, with

$$S = \prod_p \sum_{\nu=0}^{\infty} \sum_{\mu=0}^{\infty} \sum_{\substack{1 \le a_1 \le p^\nu \\ p\nmid a_1}} \sum_{\substack{1 \le a_2 \le p^\mu \\ p\nmid a_2}} \frac{\mu(p^\nu)}{\phi(p^\nu)} \frac{\mu(p^\mu)}{\phi(p^\mu)} \frac{\mu(q_3)}{\phi(q_3)} \frac{\mu(q_4)}{\phi(q_4)}, \tag{15}$$

where q_3, q_4 are defined with respect to $q_1 = p^\nu$, $q_2 = p^\mu$. As in the computation of $\sum_{j \in J} \phi_j^{-1}$, the primes $p = 2$ and $p = 3$ require separate considerations. The computations are, however, much easier here, because $\mu(p^\nu) = 0$ for $\nu > 1$.

Claim 1. The prime $p = 2$ contributes the factor 4.

Proof. There are four terms to be considered. The term with $\nu = \mu = 0$, contributes $\{\frac{\mu(1)}{\phi(1)}\}^4 = 1$. The second term with $\nu = 1$, $\mu = 0$ has $q_1 = 2$, $q_2 = 1$, $a_1 = a_2 = 1$, $\frac{1}{2} - \frac{2\cdot 1}{1} = \frac{-3}{2} = \frac{a_3}{q_3}$, $\frac{1}{1} - \frac{2.1}{2} = \frac{0}{1} = \frac{a_4}{q_4}$, so that $q_3 = 2$, $q_4 = 1$ and $\prod_{k=1}^{4} \frac{\mu(q_k)}{\phi(q_k)} = (\frac{\mu(2)}{\phi(2)})^2 (\frac{\mu(1)}{\phi(1)})^2 = +1$; by symmetry, this is also the contribution of the third term, with $\nu = 0$, $\mu = 1$.

In the last term $\nu = \mu = 1$, $q_1 = q_2 = 2$, $a_1 = a_2 = 1$ and proceeding as before we find $q_3 = q_4 = 2$, so that $\prod_{k=1}^{4} \frac{\mu(q_k)}{\phi(q_k)} = (\frac{\mu(2)}{\phi(2)})^4 = 1$. The sum of the contributions of the 4 terms equals $+4$.

Claim 2. The prime $p = 3$ contributes the factor 9/8.

Proof. The term with $\nu = \mu = 0$ contributes $\{\frac{\mu(1)}{\phi(1)}\}^4 = 1$.

If $\nu = 1$, $\mu = 0$, $q_1 = 3$, $q_2 = 1$, $a_1 = 1$ or 2, $a_2 = 1$. If $a_1 = 1$ then $\frac{1}{3} - \frac{2.1}{1} = \frac{a_3}{q_3}$, so that $q_3 = 3$; $\frac{1}{1} - \frac{2.1}{3} = \frac{a_4}{q_4}$, so that $q_4 = 3$. Similarly we verify that also for $a_2 = 2$, $q_3 = q_4 = 3$. Hence, in each case $\prod_{k=1}^{4} \frac{\mu(q_k)}{\phi(q_k)} = (\frac{\mu(3)}{\phi(3)})^3 \frac{\mu(1)}{\phi(1)} = \frac{-1}{8}$, or together $\frac{-1}{4}$. The terms with $\nu = 0$, $\mu = 1$ make the same contribution. If $\nu = \mu = 1$, $q_1 = q_2 = 3$, $a_1 = 1$, or 2, $a_2 = 1$, or 2. We verify that if $a_1 = a_2$, then $q_3 = q_4 = 3$ (this happens twice, for $a_1 = a_2 = 1$ and for $a_1 = a_2 = 2$), while if $a_1 \ne a_2$, then $q_3 = q_4 = 1$

(this happens also twice, for $a_1 = 1, a_2 = 2$ and for $a_1 = 2, a_1 = 1$). Hence, the contribution of these terms is $2(\frac{\mu(3)}{\phi(3)})^4 + 2(\frac{\mu(3)}{\phi(3)})^2(\frac{\mu(1)}{\phi(1)})^2 = \frac{2}{16} + \frac{2}{4} = \frac{5}{8}$.

Summing up, the factor corresponding to $p = 3$ equals $1 + 2.(\frac{-1}{4}) + \frac{5}{8} = \frac{9}{8}$, as claimed.

Claim 3. Any prime $p > 3$, contributes a factor $\{1 - \frac{3p-1}{(p-1)^3}\}$.

Proof. The term with $\nu = \mu = 0$, contributes $\{\frac{\mu(1)}{\phi(1)}\}^4 = + 1$.

If $\nu = 1, \mu = 0, q_1 = p, q_2 = 1$, then for any $a_1 \not\equiv 0 \pmod p$, $\frac{a_1}{p} - \frac{2.1}{1} = \frac{a_3}{q_3}$ and $\frac{1}{1} - \frac{2a_1}{p} = \frac{a_4}{q_4}$, whence (using $p \neq 2$) it follows that $q_3 = q_4 = p$. Each such term contributes $(\frac{\mu(p)}{\phi(p)})^3 \frac{\mu(1)}{\phi(1)} = (\frac{-1}{p-1})^3 = \frac{-1}{(p-1)^3}$. The sum over all a_1, $1 \leq a_1 \leq p - 1$ equals $\frac{-1}{(p-1)^2}$ and the same contribution is made by the terms with $\nu = 0, \mu = 1$. If $\nu = \mu = 1$, $q_1 = q_2 = p$, a_1 and a_2 each take all values from 1 to $p-1$. For fixed a_1, $\frac{a_1-2a_2}{p} = \frac{a_3}{q_3}$ shows that for $p-2$ values of a_2, $a_3 \not\equiv 0 \pmod p$; hence $q_3 = p$. For the one value say a_2' with $2a_2' \equiv a_1 \pmod p$, $q_3 = 1$. Similarly, $q_4 = p$, except for one value of a_2, say a_2'', when $q_4 = 1$. We claim that $a_2' \neq a_2''$. Indeed, otherwise, from $a_1 - 2a_2 \equiv a_2 - 2a_1 \equiv 0 \pmod p$ follows $a_1 + a_2 \equiv 0 \pmod p$, $a_1 - 2a_2 \equiv - 3a_2 \equiv 0 \pmod p$, which is impossible for $p \nmid a_2$, $p \geq 5$. It follows that (for fixed a_1) $p - 3$ values of a_2 lead to $q_3 = q_4 = p$, one value leads to $q_3 = p, q_4 = 1$, and one value leads to $q_3 = 1, q_4 = 4$, for a contribution of the summands with fixed $a_1 \not\equiv 0 \pmod p$ of $(p-3)(\frac{\mu(p)}{\phi(p)})^4 + 2(\frac{\mu(p)}{\phi(p)})^3 = \frac{p-3}{(p-1)^4} - \frac{2}{(p-1)^3} = - \frac{p+1}{(p-1)^4}$. The sum for all $p-1$ of a_1 equals $- \frac{p+1}{(p-1)^3}$. The total contribution of all terms adds up to

$1 + 2(\frac{-1}{(p-1)^2}) + (- \frac{p+1}{(p-1)^3})\ 1 - \frac{3p-1}{(p-1)^3}$, as claimed.

Combining the three Claims, (15) becomes

$$S = 4 \cdot \frac{9}{8} \prod_{p \geq 5} \{1 - \frac{3p-1}{(p-1)^3}\}$$

and the principal term (14) now reads

$$2\cdot\frac{9}{2}\prod_{p\geq 5}\{1-\frac{3p-1}{(p-1)^3}\}\{\sum_{2\leq n_i\leq x}{}''\frac{1}{\log n_1\log n_2\log n_3\log n_4}+O(\frac{x}{\log^4 x})\}\ .$$

The sum is known (see [3]) to be equal to

$$\frac{x^2}{6\log^4 x}\{1+\sum_{k=1}^{M}\frac{a_k}{\log^k x}+O(\frac{1}{\log^{M+1}x})\}\ ,$$

where M may be any positive integer and where the a_k are certain computable constants.

With this the principal term becomes

$$\frac{3}{2}\prod_{p\geq 5}\{1-\frac{3p-1}{(p-1)^3}\}\frac{x^2}{\log^4 x}\{1+\sum_{k=1}^{M}\frac{a_k}{\log^k x}+O(\frac{1}{\log^{M+1}x})\}.$$

We now substitute this in (13) for the multiple sum of integrals and obtain

$$I=2C_4\frac{x^2}{\log^4 x}\{1+\sum_{k=1}^{M}\frac{a_k}{\log^k x}+O(\frac{1}{\log^{M+1}x})\}+O(x^2(\log x)^{-1-u})+E(x),\quad\text{with}$$

$$C_4=\frac{3}{4}\prod_{p\geq 5}\{1-\frac{3p-1}{(p-1)^3}\}\doteq .4763\ \ldots\quad .$$

If we select $u\geq M+4$ and substitute the value of I in (3), we obtain the Theorem and, on the basis of the conjectural $E(x)=O(x^2(\log x)^{-k})$, this finishes also the proof of the Corollary.

BIBLIOGRAPHY

[1] Chowla, S., "There exists an infinity of 3-combinations of primes in A. P. - Proc. Lahore Philos. Soc. 6, No. 2 (1944), pp. 15-16.

[2] Van der Corput, J. G., "Über Summen von Primzahlen und Primzahlquadraten" - Mathem. Annalen Vol. 116 (1939), pp. 1-50.

[3] Grosswald, E., "Arithmetic progressions that consist only of primes" - J. of Number Theory, Vol. 14 (1982), pp. 9-31.

[4] Grosswald, E. and Hagis, P., "Arithmetic progressions consisting only of primes" (Mathematics of Computation, Vol. 33 (1979)), pp. 1343-1352.

[5] Hardy, G. H. and Littlewood, J. E., "Some problems of 'Partitio Numerorum' Part III," Acta Mathematica, Vol. 44 (1922), pp. 1-70.

[6] Prachar, K., *Primzahlverteilung*, - Springer-Verlag, Berlin-Göttingen-Heidelberg, 1957.

[7] Zagier, D., private communication.