JOURNAL OF NUMBER THEORY **58**, 89–99 (1996)
ARTICLE NO. 0066

Symmetric and Asymmetric Primes

PETER FLETCHER*

2466 Meadow Creek Road, Christiansburg, Virginia 24073-7240

WILLIAM LINDGREN†

Department of Mathematics, Slippery Rock University of Pennsylvania, Slippery Rock, Pennsylvania 16057-9989

AND

CARL POMERANCE‡, §

Department of Mathematics, University of Georgia, Athens, Georgia 30602-7403

Communicated by Alan C. Woods

Received January 3, 1995; revised March 31, 1995

In a well-known proof of the quadratic reciprocity law, one counts the lattice points inside the rectangle with sides parallel to the axes and opposite vertices at the origin and $(p/2, q/2)$, where p and q are distinct odd primes. In particular, the Legendre symbols (p/q) and (q/p) depend, respectively, on the number of lattice points in the rectangle above and below the main diagonal. Say p, q form a *symmetric pair* if the number of lattice points above the main diagonal is equal to the number of lattice points below. Say a prime p is *symmetric* if it belongs to some symmetric pair, and otherwise call it *asymmetric*. We first characterize symmetric pairs p, q with the condition $(p-1, q-1) = |p-q|$. In particular, twin primes form a symmetric pair. Of the first 100,000 odd primes, about 5/6 of them are symmetric. However, we are able to prove that, asymptotically, almost all primes are asymmetric.

1. INTRODUCTION

Let P denote the set of all odd primes and define $S: P \times P \to N$ by

$$S(q, p) = \sum_{k=1}^{(p-1)/2} [qk/p].$$

* E-mail: fletcher@math.vt.edu.
† E-mail: wfl@sruvm.sru.edu.
‡ The third author is supported in part by an NSF grant.
§ E-mail: carl@math.uga.edu.

0022-314X/96 $18.00

Then S is the well-known function of Eisenstein's Lemma, which asserts, when p and q are different, that $S(p, q) + S(q, p) = (p-1)(q-1)/4$ (see [3, Th. 100]). We say that a pair of different odd primes (p, q) is *symmetric* provided that $S(p, q) = S(q, p)$. A prime that belongs to a symmetric pair is a *symmetric* prime; an odd prime that is not symmetric is *asymmetric*. We characterize the symmetric pairs (p, q) with $p \neq q$ as those pairs for which $q-p$ divides $p-1$. In particular, any pair of twin primes is a symmetric pair. The surprising connection between the twin-prime conjecture and Eisenstein's Lemma, which the characterization of symmetric pairs provides, suggests the conjecture that there are infinitely many symmetric primes. A count of the symmetric and asymmetric primes through the 100,000-th prime seems to support this conjecture, because in this count the symmetric primes outnumber the asymmetric primes by about five to one. We show, however, that this count is misleading: almost all primes are asymmetric. That is, the proportion of symmetric primes up to x, among all primes up to x, tends to 0 as $x \to \infty$. In fact, we show that the sum of the reciprocals of the symmetric primes is finite and the sum of the reciprocals of the asymmetric primes is infinite.

Throughout this paper, the letters p, q represent odd prime numbers and the letters d, j, k, m, n, r, t represent non-negative integers.

2. A Characterization of Symmetric Pairs

As Eisenstein observed, if S is the rectangle that has $(0, 0)$ and $(p/2, q/2)$ as opposite corners and l is the diagonal joining these two corners, then $S(q, p)$ is the number of interior lattice points of S that lie below l and $S(p, q)$ is the number of such points that lie above l. Thus the definition of symmetric pairs may be thought of geometrically; in this section we give an algebraic characterization of symmetric pairs. While it is true that one can use the geometric characterization and the division algorithm to show that 23 is the least asymmetric prime, or to show that all twin primes are symmetric, the algebraic characterization we obtain is a significant aid in determining which primes are symmetric.

We begin with the following lemma, which is a consequence of the proof of Eisenstein's Lemma.

Lemma 2.1. *Let p and q be two odd prime numbers. Then* $S(q, p) = \sum_{k=1}^{(q-1)/2} [(p/q)(k - \frac{1}{2}) + \frac{1}{2}]$.

Proof. The lattice point definition of $S(q, p)$ given above is seen to be equivalent to the definition in Section 1 by organizing the lattice points into columns parametrized by the first coordinate k and letting k run from

1 to $(p-1)/2$: there are $[qk/p]$ points in the kth column. Equally, these lattice points may be organized into rows parametrized by the second coordinate j, where j runs from 1 to $(q-1)/2$. The number of lattice points on this row is $[p/2]-[pj/q]$. We change j into $(q+1)/2-k$, so that the new variable k also runs from 1 to $(q-1)/2$, and note that

$$\left[\frac{p}{2}\right]-\left[\frac{p}{q}\left(\frac{q+1}{2}-k\right)\right]=-\left[-\frac{p-1}{2}+\frac{p}{q}\left(\frac{q+1}{2}-k\right)\right]$$
$$=\left[\frac{p}{q}\left(k-\frac{1}{2}\right)+\frac{1}{2}\right],$$

the second equality following from the identity $-[-x]=[x+1]$, for x not an integer. This proves the lemma.

LEMMA 2.2. *Let m and n be natural numbers. Then the following are equivalent*:

(1) $[(m/n)\,k]=[(m/n)(k-(1/2))]$ *for all positive integers* $k\leqslant n/2$,

(2) m *divides* $n-1$.

Proof. The lemma is trivial if $n=1$, so assume that $n\geqslant 2$. Note that $k=1$ satisfies $k\leqslant n/2$. If $m/n\geqslant 1$, then (2) fails and (1) fails for $k=1$. Thus we assume that $m/n<1$. Note that $[(m/n)\,k]=[(m/n)(k-(1/2))]$ if, and only if, $\{mk/n\}\geqslant m/(2n)$, where $\{x\}=x-[x]$ denotes the fractional part of x. Let $j=[(n-1)/m]$ and set $n-1=jm+r$ where $0\leqslant r\leqslant m-1$. For any integer t we have

$$\frac{m}{n}(tj+1)=\frac{mtj+m}{jm+r+1}=t+\frac{m-t(r+1)}{n}. \qquad (*)$$

In order to show that (1) implies (2), we prove the contrapositive. Suppose that $r>0$ and set $t=[m/(r+1)]$. Then $m-t(r+1)$, which is the remainder when m is divided by $r+1$, is less than $m/2$, and (*) yields $\{(m/n)(tj+1)\}<m/(2n)$. It remains to show that $tj+1$ is a valid choice for k in (1). But

$$tj+1=\left[\frac{m}{r+1}\right]\frac{n-r-1}{m}+1\leqslant\frac{n-r-1}{r+1}+1=\frac{n}{r+1}\leqslant\frac{n}{2}.$$

Now suppose that (2) holds. Then $r=0$, $jm=n-1$ and, for all integers t, (*) reduces to $m(tj+1)/n=t+(m-t)/n$. For $0\leqslant t\leqslant m-1$ and $1\leqslant d\leqslant j$ we thus have $\{jm(tj+d)/n\}=(dm-tn)$. Let k be a positive integer with $k\leqslant n/2$. Since $k\leqslant n-1$, there exists t with $0\leqslant t\leqslant m-1$ and d with $1\leqslant d\leqslant j$

such that $k = tj + d$. Thus $\{mk/n\} = (dm - t)/n$. Now $tj + 1 \leqslant n/2$, so that $t \leqslant (n-2)/(2j) < m/2$. Thus $\{mk/n\} > (dm - m/2)/n \geqslant m/(2n)$. That is, (1) holds.

PROPOSITION 2.1. *Let p and q be two odd prime numbers. The following statements are equivalent*:

(1) *The pair* (p, q) *is symmetric*,

(2) *For each positive integer* $k \leqslant (q-1)/2$, $[pk/q] = [(p/q)(k - (1/2)) + 1/2]$,

(3) *For each positive integer* $k \leqslant (q-1)/2$, $[((p-q)/q)\,k] = [((p-q)/q)(k-(1/2))]$,

(4) *For each positive integer* $k \leqslant (q-1)/2$, $[(|p-q|/q)\,k] = [(|p-q|/q)(k-(1/2))]$,

(5) $p - q$ *divides* $q - 1$,

(6) $p - q$ *divides* $p - 1$,

(7) $|p - q| = (p-1, q-1)$.

Proof. It is evident from Lemma 2.1 that (2) implies (1). It also follows from Lemma 2.1 that (1) implies (2). For if $p < q$ it follows that for each positive integer k, $[pk/q] \leqslant [(p/q)(k-(1/2)) + 1/2]$, and if $q < p$, then $[pk/q] \geqslant [(p/q)(k-(1/2)) + 1/2]$. Thus

$$\sum_{k=1}^{(q-1)/2} [pk/q] = \sum_{k=1}^{(q-1)/2} \left[\frac{p}{q}\left(k - \frac{1}{2}\right) + \frac{1}{2}\right]$$

exactly when (2) holds.

The equivalences of (2) and (3) and of (5), (6) and (7) are straightforward; the equivalence of (4) and (5) is a special case of Lemma 2.2; and the equivalence of (3) and (4) follows from the identity $[|x|] = -[1 + x]$ for $x < 0$ and x not an integer.

COROLLARY 2.1. *Let* (p, q) *be a symmetric pair with* $p > q$. *Then* $p \leqslant 2q - 1$.

COROLLARY 2.2. *If both* q *and* $p = 2q - 1$ *are odd primes*, (p, q) *is symmetric.*

COROLLARY 2.3. *Any pair of twin primes is a symmetric pair.*

COROLLARY 2.4. *Let* $n > 1$ *and suppose that both* $2n + 1$ *and* $4n + 3$ *are primes. If* $4n + 5$ *is not prime*, $4n + 3$ *is asymmetric.*

By Corollary 2.3, 23 is the least candidate for an asymmetric prime. Corollary 2.4 finds that this prime is asymmetric and in fact supplies the first sixteen asymmetric primes of the form $4n+3$. The asymmetric prime 907 is not given by Corollary 2.4, and of course Corollary 2.4 does not find asymmetric primes such as 173 that are of the form $4n+1$.

We list the asymmetric primes less than the 1001st odd prime, (7933):

23, 47, 83, 167, 173, 263, 359, 383, 389, 467, 479, 503, 509, 557, 563, 587, 653, 719, 797, 839, 863, 887, 907, 971, 983, 1103, 1187, 1259, 1283, 1307, 1367, 1439, 1499, 1511, 1523, 1571, 1579, 1637, 1733, 1823, 1907, 1913, 2039, 2063, 2099, 2203, 2207, 2411, 2447, 2459, 2543, 2579, 2663, 2767, 2819, 2879, 2903, 2927, 2963, 3023, 3089, 3203, 3491, 3593, 3623, 3779, 3803, 3863, 3923, 3947, 3989, 4007, 4013, 4073, 4079, 4139, 4283, 4349, 4373, 4391, 4493, 4583, 4679, 4703, 4919, 5003, 5039, 5087, 5189, 5323, 5381, 5387, 5399, 5471, 5483, 5507, 5623, 5717, 5807, 5843, 5927, 5939, 6047, 6173, 6263, 6317, 6323, 6389, 6599, 6607, 6619, 6653, 6719, 6857, 6863, 6899, 6983, 7013, 7027, 7079, 7159, 7187, 7247, 7283, 7499, 7517, 7523, 7607, 7643, 7699, 7703, 7727, 7817, 7823, 7883.

This list shows that of the first hundred odd primes 87 are symmetric and of the first thousand odd primes 865 are symmetric.

We are indebted to Gary Roberts for a *Mathematica* program, available on request, which gives the following information about the density of symmetric primes.

Number of primes	$10K$	$20K$	$40K$	$60K$	$80K$	$100K$
Approximate density	.8474	.8420	.8370	.8345	.8331	.8326

The table above suggests perhaps that the density of symmetric primes converges to a positive number. Looking at a random point farther out, say the hundred primes in the interval [776531401, 776533741] (the 40,000,000th through the 40,000,099th odd primes), we find that 84 are symmetric. This might be thought of as further experimental evidence that the density converges to a value perhaps near 0.8. However, the asymptotic density is 0, as we now shall see.

3. The Distribution of Symmetric Primes

In what follows we write $f(x) \ll g(x)$ if there are positive numbers c, x_0 such that $|f(x)| \leqslant cg(x)$ for all $x \geqslant x_0$. We could in principle work out actual numerical values for the numbers c, x_0 that are implicit in the use of the notation, but we usually have not done so.

THEOREM 3.1. *For all sufficiently large numbers x, the number of symmetric primes up to x is* $\leqslant x/(\log x)^{1.027}$.

Proof. Suppose p, q are a symmetric pair of primes with $p<q$. Let $d=q-p$, so that $d=(p-1, q-1)$. There is an integer m with $p=dm+1$ and $q=dm+d+1$. For a fixed positive integer d, let $M(x, d)$ denote the number of integers $m \leqslant x/d$ with $dm+1$, $dm+d+1$ both prime, and for a fixed m, let $D(x, m)$ denote the number of integers $d \leqslant x/m$ with $dm+1$, $dm+d+1$ both prime. That is, if we consider the ordered pairs (d, m) with $dm \leqslant x$ and with $dm+1$ and $dm+d+1$ both prime, then $M(x, d)$ is the number of such pairs with first coordinate d and $D(x, m)$ is the number of such pairs with second coordinate m.

Let $y=e^{(\log x)^{\alpha}}$ with α to be determined, $0<\alpha<1$. Let $z>0$ and let $L=\sum_{z<p\leqslant x} 1/(p-1)$. Let $\omega_z(n)$ denote the number of distinct prime factors of n exceeding z, and let $\Omega_z(n)$ denote the number of such prime factors of n counted with multiplicity. Let $S(x)$ denote the number of primes $p \leqslant x$ such that p is the smaller member of a symmetric pair. Then for any number β with $0<\beta<1$,

$$
\begin{aligned}
S(x) \leqslant & \sum_{d \leqslant y} M(x, d)+\sum_{m \leqslant y} D(x, m)+\sum_{\substack{d \leqslant x/y \\ \omega_z(d) \leqslant \beta L}} M(x, d) \\
& +\sum_{\substack{m \leqslant x/y \\ \omega_z(m) \leqslant \beta L}} D(x, m)+\sum_{\substack{p \leqslant x \\ \Omega_z(p-1)>2\beta L}} 1 \\
= & S_1+S_2+S_3+S_4+S_5, \qquad \text{say.}
\end{aligned}
$$

Indeed, if $p<q$ is a symmetric pair with $p \leqslant x$, then there are integers d, m with $p-1=dm$ and $q=dm+d+1$. If p is not counted in S_1 and S_2, then $d<x/y$, $m<x/y$. And if p is not counted by S_3 or S_4, then $\omega_z(d), \omega_z(m)>\beta L$. Thus $\Omega_z(p-1)=\Omega_z(d)+\Omega_z(m) \geqslant \omega_z(d)+\omega_z(m)>2\beta L$, so that p is counted in S_5.

We must specify the numbers α, β, z, but first we estimate $S_1, \ldots, S_5$ as functions of these parameters. From the sieve [2, Cor. 2.4.1] (with $l=1$, $k=d$, $a=1$, $b=d$) we have

$$
M(x, d) \ll\left(\frac{d}{\varphi(d)}\right)^2 \frac{x/d}{\log^2(x/d)} \ll \frac{x(\log \log x)^2}{d \log^2(x/d)} \tag{1}
$$

uniformly for $d<x$. We have used $d/\varphi(d) \ll \log \log x$, see [3, Th. 328]. In the same way that Cor. 2.4.1 ultimately follows from Th. 2.3 in [2], we have uniformly for $m<x$,

$$
D(x, m) \ll \frac{m}{\varphi(m)} \cdot \frac{m+1}{\varphi(m+1)} \frac{x/m}{\log^2(x/m)} \ll \frac{x(\log \log x)^2}{m \log^2(x/m)}. \tag{2}
$$

We now estimate S_1, S_2. We have by (1) and our choice of y that

$$S_1 = \sum_{d \leqslant y} M(x, d) \ll \frac{x(\log\log x)^2}{\log^2 x} \sum_{d \leqslant y} \frac{1}{d} \ll \frac{x(\log\log x)^2 \log y}{\log^2 x}$$
$$= \frac{x(\log\log x)^2}{(\log x)^{2-\alpha}},$$

and similarly by (2),

$$S_2 = \sum_{m \leqslant y} D(x, m) \ll \frac{x(\log\log x)^2}{(\log x)^{2-\alpha}}.$$

To estimate S_3, S_4 we use the following lemma.

LEMMA 3.1. *Uniformly for* $0 < \beta \leqslant 1$, $x > z \geqslant 1.5$ *and* $L = \sum_{z < p \leqslant x} 1/(p-1)$, *we have*

$$\sum_{\substack{n \leqslant x \\ \omega_z(n) \leqslant \beta L}} \frac{1}{n} \ll (\log z)^{1-\beta+\beta\log\beta} (\log x)^{\beta-\beta\log\beta}.$$

Proof. The numbers $n \leqslant x$ may be uniquely factored as $n = kl$ where every prime factor of k is $\leqslant z$ and every prime factor of l is in $(z, x]$. For any integer $j \geqslant 0$ and any finite set of primes $\mathscr{P}$, if we expand

$$\left(\sum_{p \in \mathscr{P}} \left(\frac{1}{p} + \frac{1}{p^2} + \frac{1}{p^3} + \cdots\right)\right)^j$$

by the multinomial theorem, then among the terms with j distinct factors are those of the form $1/p_1^{a_1} \cdot 1/p_2^{a_2} \cdots 1/p_j^{a_j}$ where $p_1, ..., p_j$ are distinct primes in $\mathscr{P}$ and $a_1, ..., a_j > 0$. Since the multiplicity of such terms is $j!$ it follows that

$$\frac{1}{j!}\left(\sum_{p \in \mathscr{P}} \left(\frac{1}{p} + \frac{1}{p^2} + \frac{1}{p^3} + \cdots\right)\right)^j = \frac{1}{j!}\left(\sum_{p \in \mathscr{P}} \frac{1}{p-1}\right)^j$$

is an upper estimate for the sum of the reciprocals of those integers with exactly j distinct prime factors that are all in $\mathscr{P}$. Applying this first to $\mathscr{P}$ being the primes up to z and then to $\mathscr{P}$ being the primes in $(z, x]$, we have

$$\sum_{\substack{n \leqslant x \\ \omega_z(n) \leqslant \beta L}} \frac{1}{n} \leqslant \sum_{i=0}^{\infty} \frac{1}{i!}\left(\sum_{p \leqslant z} \frac{1}{p-1}\right)^i \sum_{j=0}^{[\beta L]} \frac{1}{j!}\left(\sum_{z < p \leqslant x} \frac{1}{p-1}\right)^j$$
$$= \exp\left(\sum_{p \leqslant z} \frac{1}{p-1}\right) \sum_{j=0}^{[\beta L]} \frac{1}{j!} L^j.$$

We shall use (see [3, Th. 427])

$$\sum_{p \leqslant x} \frac{1}{p-1} = \log\log x + O(1). \tag{3}$$

Thus $\exp(\sum_{p \leqslant z} 1/(p-1)) \ll \log z$, so

$$\sum_{\substack{n \leqslant x \\ \omega_z(n) \leqslant \beta L}} \frac{1}{n} \ll \log z \sum_{j=0}^{[\beta L]} \frac{1}{j!} L^j. \tag{4}$$

Now

$$\sum_{j=0}^{[\beta L]} \frac{1}{j!} L^j = \sum_{j=0}^{[\beta L]} \left(\frac{1}{\beta}\right)^j \frac{(\beta L)^j}{j!} \leqslant \left(\frac{1}{\beta}\right)^{\beta L} \sum_{j=0}^{\infty} \frac{(\beta L)^j}{j!} = \left(\frac{e}{\beta}\right)^{\beta L}.$$

Using $L = \log\log x - \log\log z + O(1)$ from (3), we have $(e/\beta)^{\beta L} \ll (\log x/\log z)^{\beta - \beta\log\beta}$, so that the lemma follows from (4).

From Lemma 3.1 and (1) we have, with $z = \log^2 x$,

$$\begin{aligned} S_3 = \sum_{\substack{d \leqslant x/y \\ \omega_z(d) \leqslant \beta L}} M(x, d) &\ll \frac{x(\log\log x)^2}{\log^2 y} \sum_{\substack{d \leqslant x/y \\ \omega_z(d) \leqslant \beta L}} \frac{1}{d} \\ &\ll \frac{x(\log x)^{\beta - \beta\log\beta} (\log\log x)^3}{\log^2 y} \\ &= \frac{x(\log\log x)^3}{(\log x)^{2\alpha - \beta + \beta\log\beta}}. \end{aligned}$$

Similarly from Lemma 3.1 and (2) we have, with $z = \log^2 x$,

$$S_4 \ll \frac{x(\log\log x)^3}{(\log x)^{2\alpha - \beta + \beta\log\beta}}.$$

To estimate S_5 we first establish the following lemma.

LEMMA 3.2. *For any $z > 0$, $x > 0$, the number of integers $n \leqslant x$ with $\Omega_z(n) \neq \omega_z(n)$ is $\leqslant x/z$.*

Proof. If $\Omega_z(n) \neq \omega_z(n)$, then there is some prime $p > z$ with $p^2 \mid n$. The number of integers $n \leqslant x$ divisible by p^2 is $[x/p^2]$. Thus it suffices to show that for all $z > 0$,

$$\sum_{p > z} \frac{1}{p^2} < \frac{1}{z}. \tag{5}$$

Note that if p_0 is an odd prime, then

$$\sum_{p\geqslant p_0}\frac{1}{p^2}<\sum_{k=0}^{\infty}\frac{1}{(p_0+2k)^2}$$
$$<\frac{1}{p_0^2}+\frac{1}{(p_0+2)^2}+\sum_{k=2}^{\infty}\frac{1}{(p_0+2k-1)(p_0+2k)}$$
$$=\frac{1}{p_0^2}+\frac{1}{(p_0+2)^2}+\sum_{k=2}^{\infty}\left(\frac{1}{p_0+2k-1}-\frac{1}{p_0+2k}\right)$$
$$=\frac{1}{p_0^2}+\frac{1}{(p_0+2)^2}+\frac{1}{p_0+3}<\frac{1}{p_0},$$

so that (5) holds for $z\geqslant 2$. Also, using the above display with $p_0=17$,

$$\sum_{p\geqslant 2}\frac{1}{p^2}=\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{5^2}+\frac{1}{7^2}+\frac{1}{11^2}+\frac{1}{13^2}+\sum_{p\geqslant 17}\frac{1}{p^2}$$
$$<\frac{1}{2^2}+\cdots+\frac{1}{13^2}+\frac{1}{17}<\frac{1}{2},$$

so that (5) also holds for $0<z<2$.

We now estimate S_5 with $z=\log^2 x$. Divide the primes $p\leqslant x$ with $\Omega_z(p-1)>2\beta L$ into two classes: if $\omega_z(p-1)=\Omega_z(p-1)$ and some prime q divides $p-1$ with $q>x^{1/\log\log x}$, then p is in the first class; the remaining primes comprise the second class. From [1], the number of integers up to x with all prime factors $\leqslant x^{1/\log\log x}$ is $<x/(\log x)^{\log\log\log x}$ for all large x. From this result and Lemma 3.2 with $z=\log^2 x$, it follows that the number of primes in the second class is $O(x/\log^2 x)$.

If p is in the first class and q is the largest prime factor of $p-1$, then $p-1=nq$, where $n<x^{1-1/\log\log x}$, and $\omega(n)\geqslant\omega_z(n)=\Omega_z(n)>2\beta L-1$, where $\omega(n)$ is the number of different prime factors of n. Thus,

$$S_5=\sum_{\substack{p\leqslant x\\ \Omega_z(p-1)>2\beta L}}1\leqslant\sum_{\substack{n<x^{1-1/\log\log x}\\ \omega(n)>2\beta L-1}}\sum_{\substack{q\text{ prime}\\ q<x/n\\ qn+1\text{ prime}}}1+O\left(\frac{x}{\log^2 x}\right).$$

By the sieve (see [2, Cor. 2.4.1]), the sum on q is

$$\ll\frac{n}{\varphi(n)}\frac{x/n}{\log^2(x/n)}\ll\frac{x(\log\log x)^3}{n\log^2 x},$$

using $n < x^{1-1/\log\log x}$. Thus, as in the proof of Lemma 3.1,

$$S_5 \ll \frac{x(\log\log x)^3}{\log^2 x} \sum_{\substack{n < x^{1-1/\log\log x} \\ \omega(n) > 2\beta L - 1}} \frac{1}{n} + \frac{x}{\log^2 x}$$

$$\leqslant \frac{x(\log\log x)^3}{\log^2 x} \sum_{j > 2\beta L - 1} \frac{1}{j!} \left(\sum_{p < x^{1-1/\log\log x}} \frac{1}{p-1} \right)^j + \frac{x}{\log^2 x}. \tag{6}$$

Let $\bar{L}$ denote the sum on p, so that from (3), $\bar{L} = \log\log x + O(1)$. Assume now that β is chosen so that $1/2 < \beta < 1$. Since $L = \log\log x - \log\log\log x + O(1)$, we have $2\beta L - 1 > \bar{L}$ for all large x, depending on the choice of β. Again as in the proof of Lemma 3.1, we have

$$\sum_{j > 2\beta L - 1} \frac{1}{j!} \bar{L}^j = \sum_{j > 2\beta L - 1} \left(\frac{\bar{L}}{2\beta L - 1} \right)^j \frac{(2\beta L - 1)^j}{j!} < \left(\frac{e\bar{L}}{2\beta L - 1} \right)^{2\beta L - 1}$$

$$= \exp((2\beta L - 1)(1 + \log \bar{L} - \log L - \log(2\beta) + O(1/L)))$$

$$= \exp\left(2\beta L \left(1 - \log(2\beta) + \frac{\log\log\log x}{L} \right) + O(1) \right)$$

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

$$< (\log x)^{2\beta - 2\beta\log(2\beta)} (\log\log x)^2.$$

Thus, from (6) we have that

$$S_5 \ll \frac{x(\log\log x)^5}{(\log x)^{2 - 2\beta + 2\beta\log(2\beta)}} + \frac{x}{\log^2 x}.$$

We now choose α, β so that our estimates for S_1, S_2, S_3, S_4, S_5 are all about equal. That is, we choose α, β with $0 < \alpha < 1$, $1/2 < \beta < 1$ and

$$2 - \alpha = 2\alpha - \beta + \beta\log\beta = 2 - 2\beta + 2\beta\log(2\beta).$$

An approximate numerical solution to this system is $\alpha = 0.972$, $\beta = 0.621$, which gives $S(x) = o(x/(\log x)^{1.027})$ for $x \to \infty$.

An almost identical proof gives that $S'(x) = o(x/(\log x)^{1.027})$ for $x \to \infty$, where $S'(x)$ is the number of primes $q \leqslant x$ for which q is the larger member of a symmetric pair. The only difference is that we redefine the functions $D(x, m)$, $M(x, d)$ by replacing $dm + d + 1$ with $dm - d + 1$.

Since the number of symmetric primes up to x is $\leqslant S(x) + S'(x)$, we have the theorem.

Remarks. An immediate corollary of Theorem 3.1 is that the sum of the reciprocals of the symmetric primes is finite. Since the sum of the reciprocals of all of the primes is infinite, it follows that the sum of the reciprocals of the asymmetric primes is infinite. We conjecture that there are infinitely many symmetric primes. From Corollary 2.3, this follows immediately from the twin prime conjecture. However, there are probably many more symmetric primes than twin primes. For example, any two primes that differ by 4 and are 1 mod 4 form a symmetric pair. In light of Theorem 3.1 it is natural to conjecture that there is some number $\sigma > 1$ such that the number of symmetric primes up to x is $x/(\log x)^{\sigma + o(1)}$ as $x \to \infty$. If such a number σ exists, then from Theorem 3.1, $\sigma \geqslant 1.027$.

We conjecture that $\sigma = 2 - (1 + \log\log 2)/\log 2 = 1.08607\ldots$. Here is a brief heuristic that supports this conjecture. For each divisor d of $p-1$, the prime p has two "chances" to be a symmetric prime, namely that either $p-d$ or $p+d$ is prime. So if we say a number near p is prime with "probability" $1/\log p$, then the probability that p is symmetric is $1-(1-1/\log p)^{2\tau(p-1)}$, where τ counts the number of positive divisors of its argument. The "normal" value of $\tau(p-1)$ is about $(\log p)^{\log 2} < (\log p)^{0.7}$, which "shows" why symmetric primes are rare. If p is exceptional in that $\tau(p-1) > \log p$, then it is not unlikely that p is symmetric, and the number of such exceptional primes up to x can (rigorously) be shown to be $x/(\log x)^{\sigma + o(1)}$ as $x \to \infty$, with the above value of σ. In particular, if $\tau(p-1)$ is larger than a large constant times $\log p$, then p is "very likely" to be symmetric, while if $\tau(p-1)$ is less than a small constant times $\log p$, then p is "very unlikely" to be symmetric. In particular, for $\log 2 < \lambda \leqslant 1$, the number of symmetric primes up to x with $\tau(p-1) < (\log p)^{\lambda}$ should be $x/(\log x)^{3-\lambda+(\lambda\log\lambda - 1 - \log\log 2)/\log 2 + o(1)}$ as $x \to \infty$. This exponent on $\log x$ is decreasing with value σ at $\lambda = 1$.

Acknowledgment

We thank the referee for a helpful suggestion which simplified the proof of Lemma 3.1.

References

1. N. G. de Bruijn, On the number of positive integers $\leqslant x$ and free of prime factors $> y$, *Nederl. Akad. Wetensch. Proc. Ser. A* **54** (1951), 50–60.
2. H. Halberstam and H.-E. Richert, "Sieve Methods," Academic Press, London, 1974.
3. G. H. Hardy and E. M. Wright, "An Introduction to the Theory of Numbers," 4th ed., Oxford Univ. Press, London, 1968.