On the Conjecture of Hardy & Littlewood concerning the Number of Primes of the Form $n^2 + a$

By Daniel Shanks

1. Introduction. In a famous paper, [1], Hardy and Littlewood developed a number of conjectures concerning the twin primes, the Goldbach problem, and other unsettled questions. One of these, Conjecture F, concerned the number of primes of the form $Am^2 + Bm + C$. We reword this conjecture, and at the same time reduce its generality somewhat, as follows:

CONJECTURE. *If a is an integer which is not a negative square, $a \neq -k^2$, and if $P_a(N)$ is the number of primes of the form $n^2 + a$ for $1 \leqq n \leqq N$, then*

$$(1) \qquad P_a(N) \sim \frac{1}{2} h_a \int_2^N \frac{dn}{\log n}$$

where the constant h_a is the infinite product

$$(2) \qquad h_a = \prod_{w \nmid a}^{\infty} \left(1 - \left(\frac{-a}{w}\right) \frac{1}{w-1}\right)$$

taken over all odd primes, w, which do not divide a, and for which $(-a/w)$ is the Legendre symbol.

In the trivial cases, $a = -k^2$, since $(k^2/w) = +1$ for every w, we have $h_a = 0$ on the one hand, and on the other there can be at most one prime of the form $n^2 - k^2 = (n - k)(n + k)$. For any other a, $h_a > 0$, and the conjecture indicates that there are infinitely many primes. But for no a has this been proven.

In particular, for $a = 1$, since $(-1/w)$ equals $+1$ or -1 according as $w = 4m + 1$ or $4m - 1$, we have

$$(3) \quad h_1 = (1 + \tfrac{1}{2})(1 - \tfrac{1}{4})(1 + \tfrac{1}{6})(1 + \tfrac{1}{10})(1 - \tfrac{1}{12}) \cdots = 1.37281346 \cdots$$

and therefore (1) implies that

$$(4) \qquad P_1(N) \sim 0.68640673 \int_2^N \frac{dn}{\log n}.$$

A. E. Western [2] verified that the number of primes of the form $n^2 + 1$ agreed well with the right side of (4) up to $N = 15{,}000$.

In a recent paper [3] a sieve method was developed for factoring numbers of the form $n^2 + 1$, and more generally of the form $n^2 + a$, and it was shown that the good agreement in (4) continues to hold out to $N = 180{,}000$; $(N^2 + 1 = 32{,}400{,}000{,}001)$. This verification, however, was not applied to (4) directly but to the related formula, (7), given below.

Let $\bar{\pi}_a(N)$ be the number of odd primes, q, which are $\leqq N$, which do not divide

Received April 26, 1960

a, and for which $(-a/q) = -1$. These are the primes which never divide $n^2 + a$. It is well known that

$$\bar{\pi}_a(N) \sim \frac{1}{2}\int_2^N \frac{dn}{\log n} \tag{5}$$

and therefore (1) can be rewritten as

$$\frac{P_a(N)}{\bar{\pi}_a(N)} \sim h_a. \tag{6}$$

Likewise (4) can be rewritten as

$$\frac{P_1(N)}{\bar{\pi}_1(N)} \sim 1.37281346 \cdots. \tag{7}$$

Since, in [3], we had $P_1(180{,}000) = 11223$, $\bar{\pi}_1(180{,}000) = 8178$, and $11223/8178 = 1.37234$, the agreement with the right side of (7) was even better than could be expected.

It is clear that the $\bar{\pi}_a(N)$ in (6) could be replaced by the asymptotically equal $\frac{1}{2}\pi(N)$ or by $\overset{+}{\pi}_a(N)$, (for the latter number we count the p's such that $(-a/p) = +1$). But (6) as it stands is to be preferred for two reasons. First, $\bar{\pi}_a(N)$ is generally much closer to $\frac{1}{2}\int_2^N dn/\log n$ than are either of the other two counts. See [4, sec. 10 and Table 7] for a discussion of the case $a = 1$. Second, the ratio in (6) has a simple geometric interpretation in the algebraic number field $R(\sqrt{-a})$. See [3, p. 82] for a discussion of the case $a = 1$, the Gauss plane.

In the present paper [5] we first develop an interesting and rapidly converging formula for computing the h_a and we tabulate these constants for $a = -4(1)4$. We then present short tables of $P_a(N)$ and $\bar{\pi}_a(N)$ for $a = \pm 2, \pm 3, +4$, and for $N = 10{,}000(10{,}000)180{,}000$ which show that (6) also gives good agreement in these five cases. Finally we present an elementary (sieve) argument which makes it plausible that the Hardy-Littlewood conjecture is true for every a. Further, an analysis of this computation enables us to isolate the essential difficulty in obtaining a proof.

2. The Right Side of (6). To compute the h_a we will want the following

LEMMA. *For* $|x| < \frac{1}{2}$,

$$\frac{1}{1 - 2x} = \prod_{s=1}^{\infty} \left(\frac{1 + x^s}{1 - x^s}\right)^{b(s)} \tag{8}$$

where the exponents $b(s)$ *are given by* $b(1) = b(2) = b(3) = 1$, $b(4) = 2$, $b(5) = 3$, $b(6) = 5$, *and, in general, if* d *is an odd divisor of* s *and* $\mu(d)$ *is its Möbius function, then*

$$b(s) = \frac{1}{2s} \sum_d \mu(d) 2^{s/d}. \tag{9}$$

Examples of (9): A.) If $s = p$, an odd prime, $d = 1$ or $d = p$ and [6]

$$b(p) = (2^p - 2)/2p = (2^{p-1} - 1)/p. \tag{9a}$$

B.) If $s = 2^k$, then d can only equal 1 and

$$b(s) = 2^{s-1}/s. \tag{9b}$$

Therefore $b(7) = 9$ and $b(8) = 16$.

PROOF OF THE LEMMA. After taking the logarithm of both sides of (8),

$$-\ln(1 - 2x) = \sum_{s=1}^{\infty} b(s) \ln[(1 + x^s)/(1 - x^s)], \tag{10}$$

we expand both sides in Maclaurin series and identify the corresponding coefficients. This yields the condition, for $s = 2^k m$, with m odd,

$$2^{s-1} = \sum_{d/m} \frac{s}{d} b\left(\frac{s}{d}\right). \tag{11}$$

Now applying the Möbius inversion formula we obtain (9). Since from (11) we also have $b(s) \leqq 2^s/2s$ it follows that (10) converges if $|x| < \frac{1}{2}$ and the steps may be reversed to yield (8).

Now for any $a \neq -k^2$ let p_i be the odd primes such that $(-a/p) = +1$, let q_i be the odd primes such that $(-a/q) = -1$, and let $r_1 = 2, r_2, r_3, \cdots, r_c$ be the (finite number of) primes which divide $2a$. Further, for $s = 1, 2, 3, \cdots$, let

$$L_a(s) = \left[\prod_{p,q}\left(1 - \frac{1}{p^s}\right)\left(1 + \frac{1}{q^s}\right)\right]^{-1}, \tag{12}$$

the product being taken over the p's and q's in numerical order. Finally for $s = 2, 3, 4, \cdots$, let

$$\zeta_a(s) = \zeta(s) \prod_{i=1}^{c} (1 - r_i^{-s}) \tag{13}$$

where $\zeta(s)$ is the Riemann zeta function.

THEOREM. *If*

$$f_a^{(0)} = \zeta_a(2)/L_a(1) \text{ and } K_a^{(0)}(s) = \zeta_a(2s)/L_a(s)\zeta_a(s) \tag{14}$$

for $s = 2, 3, 4, \cdots$, *then*

$$h_a = f_a^{(0)} \cdot \prod_{s=2}^{\infty} [K_a^{(0)}(s)]^{b(s)}, \tag{15}$$

where $b(s)$ *is given by* (9). *More generally, for more rapid convergence, we may select a positive integer* u *and define*

$$f_a^{(u)} = f_a^{(0)} \prod_{i=1}^{u}\left(1 - \frac{2}{p_i(p_i - 1)}\right) = f_a^{(0)} \prod_{i=1}^{u}\left(1 - \frac{2}{p_i}\right)\left(\frac{p_i + 1}{p_i - 1}\right), \tag{16}$$

and

$$K_a^{(u)}(s) = K_a^{(0)}(s) \prod_{i=1}^{u}\left(1 + \frac{2}{p_i^s - 1}\right) = K_a^{(0)}(s) \prod_{i=1}^{u}\left(\frac{p_i^s + 1}{p_i^s - 1}\right). \tag{17}$$

Then for every $u = 0, 1, 2, \cdots$,

$$h_a = f_a^{(u)} \prod_{s=2}^{\infty} [K_a^{(u)}(s)]^{b(s)}. \tag{18}$$

PROOF. For every $s = 2, 3, 4, \cdots$,

$$\zeta(s) = \left[\prod_{p,q,r}\left(1-\frac{1}{p^s}\right)\left(1-\frac{1}{q^s}\right)\left(1-\frac{1}{r^s}\right)\right]^{-1}$$

and we easily verify that

$$1 = K_a^{(0)}(s)\prod_p\left(\frac{p^s+1}{p^s-1}\right). \tag{19}$$

We likewise find that

$$h_a = f_a^{(0)}\prod_p\left(1-\frac{2}{p}\right)\left(\frac{p+1}{p-1}\right) \tag{20}$$

so for any positive integer m, we have from (19) and (20)

$$h_a = f_a^{(u)}\prod_{s=2}^{m}[K_a^{(u)}(s)]^{b(s)}\cdot\prod_{i=u+1}^{\infty}\left(1-\frac{2}{p_i}\right)\left(\frac{p_i+1}{p_i-1}\right)\cdot\prod_{s=2}^{m}\prod_{i=u+1}^{\infty}\left(\frac{p_i^s+1}{p_i^s-1}\right)^{b(s)}.$$

Since m is finite the order of the products may be changed to give

$$h_a = f_a^{(u)}\prod_{s=2}^{m}[K_a^{(u)}(s)]^{b(s)}\cdot\prod_{i=u+1}^{\infty}\left(1-\frac{2}{p_i}\right)\cdot\prod_{s=1}^{m}\left(\frac{p_i^s+1}{p_i^s-1}\right)^{b(s)}.$$

Now every $p > 2$, and we may therefore use (8) with $x = 1/p_i$ to obtain

$$h_a = f_a^{(u)}\prod_{s=2}^{m}[K_a^{(u)}(s)]^{b(s)}\cdot\prod_{i=u+1}^{\infty}\prod_{s=m+1}^{\infty}\left(\frac{p_i^s-1}{p_i^s+1}\right)^{b(s)}.$$

But it may be readily seen that the double infinite product on the right converges (monotonically increasing) to 1 as $m \to \infty$, and it thus follows that the right side of (18) converges (monotonically decreasing) to h_a as $m \to \infty$.

The computation of the h_a from (18) requires knowledge of the $L_a(s)$. Now every $L_a(s)$ has a Dirichlet series

$$L_a(s) = \sum_{n=1}^{\infty} d_n(a)n^{-s}$$

with real periodic coefficients. Specifically we have

$$\begin{aligned} L_1(s) &= 1 - 3^{-s} + 5^{-s} - 7^{-s} + - + -,\\ L_2(s) &= 1 + 3^{-s} - 5^{-s} - 7^{-s} + + - -,\\ L_{-2}(s) &= 1 - 3^{-s} - 5^{-s} + 7^{-s} + - - +,\\ L_3(s) &= 1 - 5^{-s} + 7^{-s} - 11^{-s} + - + -,\\ L_{-3}(s) &= 1 - 5^{-s} - 7^{-s} + 11^{-s} + - - +,\\ L_4(s) &= 1 - 3^{-s} + 5^{-s} - 7^{-s} + - + -. \end{aligned} \tag{21}$$

The $L_a(1)$, which enter into $f_a^{(0)}$ as defined by eq. (14), may be obtained in closed form by use of Gauss sums and Fourier series, [7]. Specifically, for $a > 0$ we have the simple

$$L_a(1) = \frac{\pi}{2\sqrt{a}}\, q_a \tag{22}$$

where the q_a for $1 \leqq a \leqq 100$ are listed in Table 1.

TABLE 1

a	q_a	a	q_a	a	q_a	a	q_a
1	$\frac{1}{2}$	26	6	51	6	76	6
2	1	27	3	52	4	77	8
3	1	28	2	53	6	78	4
4	1	29	6	54	6	79	5
5	2	30	4	55	4	80	8
6	2	31	3	56	8	81	6
7	1	32	4	57	4	82	4
8	2	33	4	58	2	83	9
9	2	34	4	59	9	84	8
10	2	35	6	60	4	85	4
11	3	36	4	61	6	86	10
12	2	37	2	62	8	87	6
13	2	38	6	63	4	88	4
14	4	39	4	64	4	89	12
15	2	40	4	65	8	90	8
16	2	41	8	66	8	91	6
17	4	42	4	67	3	92	6
18	2	43	3	68	8	93	4
19	3	44	6	69	8	94	8
20	4	45	4	70	4	95	8
21	4	46	4	71	7	96	8
22	2	47	5	72	4	97	4
23	3	48	4	73	4	98	8
24	4	49	4	74	10	99	6
25	2	50	6	75	6	100	4

TABLE 2

a	h_a
−4	0
−3	1.38342429
−2	1.85005441
−1	0
0	0
1	1.37281346
2	0.71306310
3	1.12073275
4	1.37281346

The $L_a(1)$ for negative a are a little more complicated and will not be listed here. As regards $L_a(s)$ for other values of s, $L_1(s)$ is a well known function, but except for a few scattered results, [8], values of the other L's do not seem to have been published. J. W. Wrench, Jr. has computed unpublished tables of $L_a(s)$ for $a = \pm 2$ and ± 3. With his permission the author used these tables, together with (18), to compute the four corresponding values of h_a in Table 2. The remaining entries, $h_{-4} = h_{-1} = h_0 = 0$ and $h_4 = h_1$, are trivial.

The variation of the h_a in Table 2 is notable. For example, there should be more than two and one-half times as many primes of the form $n^2 - 2$ as of the form $n^2 + 2$. As a side remark, we note from (15) that $f_a^{(0)} = 2\zeta_a(2)\sqrt{a}/\pi q_a$ is the leading factor of h_a. Thus for $a > 0$, $n^2 + a$ will therefore have few or many primes according as q_a is large or small (relative to $2\sqrt{a}/\pi$). From Table 1 we see that there will be few primes for $a = 2, 5, 11, 14, 26, 41, 89$, and 194, ($q_{194} = 20$) and there will be many primes for $a = 7, 37, 58$, and 163, ($q_{163} = 3$). The famous function of Euler, $n^2 + n + 41$, equals $\frac{1}{4}[(2n + 1)^2 + 163]$ and its well-known richness in primes is thus closely related to the small value of q_{163}. This, in turn, is related in class number theory to the unique factorization of the integers in the algebraic number field $R(\sqrt{-163})$.

3. The Left Side of (6). Tables of $P_a(N)$ and $\bar{\pi}_a(N)$ for $a = \pm 2, \pm 3, +4$, and $N = 100k$ ($k = 1, 2, \cdots, 1800$) were computed with an IBM 704 program based on the sieve method and the p-adic square roots of $-a$, [3, sec. 9]. At the same time the prime divisors of $n^2 + a$ which do not exceed N were counted, and from these counts the values of $\bar{\pi}_a(N)$ are easily obtained. Summaries of these results are given in Tables 3, 4, and 5. In the last of these, the results for $a = 4$ are compared with the previous results [3] for $a = 1$.

4. Both Sides of (6). In Figure 1 we plot $P_a(N)/\bar{\pi}_a(N)$ versus N together with the conjectured limits, h_a, for $a = \pm 2$ and ± 3. The cases $a = 1$ and $a = 4$, (which should be asymptotically equal since $h_1 = h_4$), are not included in this figure for clarity. If included, these two graphs would intertwine that for the case $a = -3$.

5. An Elementary Interpretation. The over-all impression of the foregoing results is that (6) and its equivalent (1) are almost surely true for $a = 1, \pm 2, \pm 3, 4$.

TABLE 3

N	$P_2(N)$	$\bar{\pi}_2(N)$	$P_2(N)/\bar{\pi}_2(N)$	$P_{-2}(N)$	$\bar{\pi}_{-2}(N)$	$P_{-2}(N)/\bar{\pi}_{-2}(N)$
10000	446	622	0.6737	1153	625	1.8448
20000	817	1134	0.7205	2140	1140	1.8772
30000	1180	1632	0.7230	3087	1631	1.8927
40000	1494	2117	0.7057	3977	2112	1.8830
50000	1821	2580	0.7058	4824	2587	1.8647
60000	2160	3051	0.7080	5643	3041	1.8556
70000	2489	3478	0.7156	6464	3481	1.8569
80000	2823	3942	0.7161	7296	3927	1.8579
90000	3139	4378	0.7170	8083	4374	1.8480
100000	3422	4798	0.7132	8888	4808	1.8486
110000	3721	5229	0.7116	9681	5242	1.8468
120000	4027	5649	0.7129	10500	5682	1.8479
130000	4347	6090	0.7138	11304	6117	1.8480
140000	4652	6516	0.7139	12086	6533	1.8500
150000	4966	6945	0.7150	12828	6956	1.8442
160000	5250	7347	0.7146	13628	7362	1.8511
170000	5522	7767	0.7110	14397	7763	1.8546
180000	5847	8192	0.7138	15134	8184	1.8492

TABLE 4

N	$P_3(N)$	$\bar{\pi}_3(N)$	$P_3(N)/\bar{\pi}_3(N)$	$P_{-3}(N)$	$\bar{\pi}_{-3}(N)$	$P_{-3}(N)/\bar{\pi}_{-3}(N)$
10000	711	616	1.1542	850	620	1.3710
20000	1302	1136	1.1461	1569	1139	1.3775
30000	1851	1633	1.1335	2238	1637	1.3671
40000	2378	2112	1.1259	2903	2108	1.3771
50000	2920	2575	1.1340	3550	2577	1.3776
60000	3428	3041	1.1273	4168	3030	1.3756
70000	3967	3490	1.1367	4796	3466	1.3837
80000	4463	3937	1.1336	5442	3935	1.3830
90000	4941	4373	1.1299	6049	4374	1.3829
100000	5426	4806	1.1290	6664	4819	1.3829
110000	5917	5233	1.1307	7253	5247	1.3823
120000	6410	5665	1.1315	7874	5673	1.3880
130000	6873	6105	1.1258	8491	6097	1.3927
140000	7337	6532	1.1232	9073	6524	1.3907
150000	7823	6940	1.1272	9663	6950	1.3904
160000	8302	7361	1.1278	10236	7363	1.3902
170000	8781	7768	1.1304	10799	7765	1.3907
180000	9240	8195	1.1275	11354	8200	1.3846

TABLE 5

N	$P_4(N)$	$\bar{\pi}_4(N) = \bar{\pi}_1(N)$	$P_4(N)/\bar{\pi}_4(N)$	$P_1(N)$	$P_1(N)/\bar{\pi}_1(N)$	$P_1(N)/P_4(N)$
10000	870	619	1.4055	841	1.3586	0.967
20000	1554	1136	1.3680	1559	1.3724	1.003
30000	2216	1633	1.3570	2268	1.3889	1.023
40000	2838	2117	1.3406	2952	1.3944	1.040
50000	3459	2583	1.3391	3613	1.3988	1.045
60000	4083	3038	1.3440	4252	1.3996	1.041
70000	4690	3485	1.3458	4888	1.4026	1.042
80000	5281	3933	1.3427	5513	1.4017	1.044
90000	5903	4364	1.3527	6084	1.3941	1.031
100000	6517	4808	1.3554	6656	1.3844	1.021
110000	7099	5247	1.3530	7239	1.3796	1.020
120000	7700	5675	1.3568	7795	1.3736	1.012
130000	8300	6103	1.3600	8369	1.3713	1.008
140000	8893	6531	1.3617	8944	1.3695	1.006
150000	9442	6941	1.3603	9505	1.3694	1.007
160000	10008	7361	1.3596	10072	1.3683	1.006
170000	10565	7770	1.3597	10658	1.3717	1.009
180000	11143	8178	1.3626	11223	1.3723	1.007

We now offer a theoretical argument in favour of these asymptotic equations for all a. We will specifically carry it through for $a = 1$, but the argument is easily generalized. The case $a = 1$ is the only one which Hardy and Littlewood treated in detail. Their computation, however, was deep and function-theoretic. In contrast, the present argument is elementary, [9]. It will be assumed that the reader is acquainted with the $n^2 + 1$ sieve which is described in detail in [3].

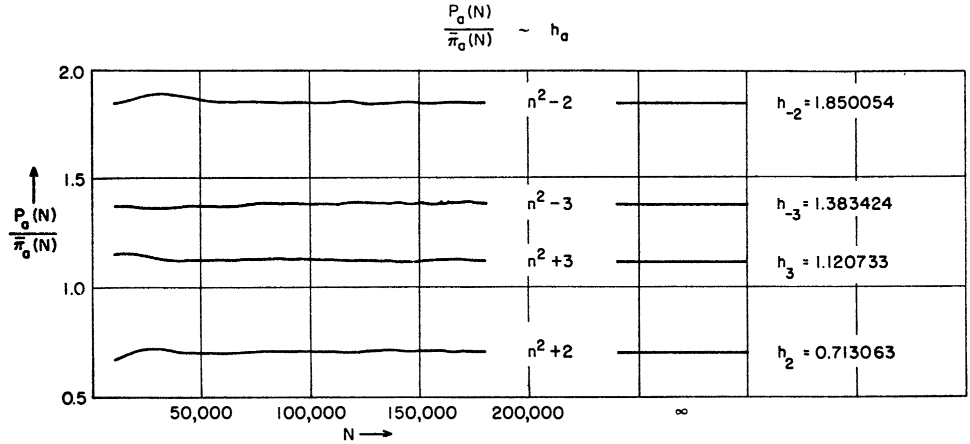

FIG. 1.—The Hardy Littlewood Conjecture.

Consider the infinite product (3) for h_1, *not* in the form in which it was given by Hardy and Littlewood, (2),

$$h_1 = \left(1 + \frac{1}{3-1}\right)\left(1 - \frac{1}{5-1}\right)\left(1 + \frac{1}{7-1}\right)\left(1 + \frac{1}{11-1}\right)\left(1 - \frac{1}{13-1}\right)\cdots,$$

since this masks its true nature; but in the equivalent form

$$h_1 = \frac{1}{\left(1-\frac{1}{3}\right)} \cdot \frac{\left(1-\frac{2}{5}\right)}{\left(1-\frac{1}{5}\right)} \cdot \frac{1}{\left(1-\frac{1}{7}\right)} \cdot \frac{1}{\left(1-\frac{1}{11}\right)} \cdot \frac{\left(1-\frac{2}{13}\right)}{\left(1-\frac{1}{13}\right)} \cdots$$

or, even better, as

$$(23) \quad h_1 = \frac{\left(1-\frac{1}{2}\right)}{\left(1-\frac{1}{2}\right)} \cdot \frac{1}{\left(1-\frac{1}{3}\right)} \cdot \frac{\left(1-\frac{2}{5}\right)}{\left(1-\frac{1}{5}\right)} \cdot \frac{1}{\left(1-\frac{1}{7}\right)} \cdot \frac{1}{\left(1-\frac{1}{11}\right)} \cdot \frac{\left(1-\frac{2}{13}\right)}{\left(1-\frac{1}{13}\right)} \cdots.$$

Now for a suitably large N let w^* be the greatest prime satisfying $w \leqq N$ and let p^* be the greatest prime of the form $4m + 1$ which satisfies $p \leqq N$. We write the corresponding *partial* product of (23), which approximates h_1, as follows:

$$(24) \qquad h_1 \approx N \cdot \frac{N\left(1-\frac{1}{2}\right)\left(1-\frac{2}{5}\right)\left(1-\frac{2}{13}\right)\cdots\left(1-\frac{2}{p^*}\right)}{N^2\left(1-\frac{1}{2}\right)\left(1-\frac{1}{3}\right)\left(1-\frac{1}{5}\right)\cdots\left(1-\frac{1}{w^*}\right)}.$$

Now this approximation to h_1 is in turn seen to be approximated (and we will inquire later as to the degree of the approximation) by N times the ratio of the primes which remain in two sieves, the Eratosthenes sieve (for all primes) from $n = 1$ to $n = N^2$ in the denominator and the $n^2 + 1$ sieve from $n^2 + 1 = 2$ to $n^2 + 1 = N^2 + 1$ in the numerator.

Without attempting precision at this point—that is, without bounding the error—we note that in the Eratosthenes sieve one first strikes out the multiples of 2. This leaves $N^2(1 - \frac{1}{2})$ numbers (with an error of 0 or $\frac{1}{2}$). One then strikes out the remaining multiples of 3 leaving $N^2(1 - \frac{1}{2})(1 - \frac{1}{3})$ numbers (again except for a possible end-effect correction.) Continuing with the primes 5, 7, $\cdots$, w^* creates the denominator of (24). The latter therefore equals

$$\pi(N^2) - \pi(N) + E(N),$$

the number of primes up to N^2 minus the number of primes up to N, with an end-effects error, $E(N)$, which is not yet bounded. We note that

$$\pi(N^2) - \pi(N) \sim \frac{N}{2}\pi(N) \sim N\bar{\pi}_1(N)$$

by the prime number theorem.

In the $n^2 + 1$ sieve we first factor a 2 from all numbers where $n = 2m + 1$ leaving $N(1 - \frac{1}{2})$ of the numbers (except for an end-effect error). We then factor a 5 where $n = 5m + 2$ and where $n = 5m + 3$. This leaves $N(1 - \frac{1}{2})(1 - \frac{2}{5})$ numbers (except for the end-effect error). Continuing with all primes of the form $4m + 1$; 13, 17, $\cdots$, p^* generates the numerator. The latter therefore equals

$$P(N) - P(\sqrt{N - 1}) + e(N),$$

the number of primes of the form $n^2 + 1$ up to $N^2 + 1$ minus the number of such primes up to N with an end effect $e(N)$.

Therefore, we *may* write

$$h_1 = \lim_{N\to\infty} \frac{P(N) - P(\sqrt{N - 1}) + e(N)}{\bar{\pi}_1(N) + E(N)/N}, \tag{25}$$

while what we would *like* to write is

$$h_1 = \lim_{N\to\infty} \frac{P(N)}{\bar{\pi}_1(N)}.$$

Now by Merten's Theorem the denominator of (24) is asymptotic to $N^2e^{-\gamma}/\log N$ where γ is Euler's constant [10]. Therefore the end effect, $E(N)/N$, is *not* negligible compared with $\bar{\pi}_1(N)$. Instead we have

$$\frac{E(N)/N}{\bar{\pi}_1(N)} \sim 0.1229 = 2e^{-\gamma} - 1. \tag{26}$$

If we could show

$$\frac{e(N)}{P(N) - P(\sqrt{N - 1})} \sim 2e^{-\gamma} - 1 \tag{27}$$

all would be well, but the difficulty of the problem is such that we cannot even prove that the left side of (27) is *bounded* from above. If we could do that, we would at least have $P(N) \to \infty$ but even this "weak" result eludes us.

It is of interest to analyze this difficulty. Let

$$D(N) = P(N) - P(\sqrt{N - 1}) \tag{28}$$

and

$$S(N) = N\left(1 - \frac{1}{2}\right)\left(1 - \frac{2}{5}\right)\cdots\left(1 - \frac{2}{p^*}\right). \tag{29}$$

Then the conjectured relation (27) is equivalent to the conjecture

$$\frac{S(N)}{D(N)} \sim 2e^{-\gamma} = 1.1229. \tag{30}$$

Now from the sieve for $n^2 + 1$, [3], we can obtain an *exact* formula for $D(N)$ by using the "integer part of x" function, $[x]$. Consider the set of numbers obtained from

$$d = 2^a \cdot 5^b \cdot 13^c \cdots p^{*z}$$

by assigning (in all possible ways) 0 and 1 to the exponents $a, b, c, \cdots$. For each such d, let A_i be the solutions of

$$A^2 \equiv -1 \pmod d$$

which satisfy

$$0 \leqq A < d.$$

Then if d is a product of α primes, we have

$$D(N) = \sum_d (-1)^\alpha \sum_i \left[\frac{N + A_i}{d}\right]. \tag{31}$$

It may be seen that if there are M primes of the form $4m + 1$ which are $\leqq N$, then there will be $2 \cdot 3^M$ terms in this sum. Even for a very modest N, say 15, we have $p^* = 13$, $M = 2$, and there are already 18 terms. Specifically,

$$\begin{aligned} D(N) = [N] &- \left[\frac{N+1}{2}\right] - \left[\frac{N+3}{5}\right] - \left[\frac{N+2}{5}\right] + \left[\frac{N+7}{10}\right] + \left[\frac{N+3}{10}\right] \\ &- \left[\frac{N+8}{13}\right] - \left[\frac{N+5}{13}\right] + \left[\frac{N+21}{26}\right] + \left[\frac{N+5}{26}\right] + \left[\frac{N+57}{65}\right] \\ &+ \left[\frac{N+47}{65}\right] + \left[\frac{N+18}{65}\right] + \left[\frac{N+8}{65}\right] - \left[\frac{N+83}{130}\right] - \left[\frac{N+73}{130}\right] \\ &- \left[\frac{N+57}{130}\right] - \left[\frac{N+47}{130}\right]. \end{aligned}$$

In general, it is easily seen, the formula for $S(N)$ may be obtained from that for $D(N)$ by deleting the A_i and the square brackets. Thus for $N = 15$ in the example, we have

$$\begin{aligned} S(N) &= N - \frac{N}{2} - \frac{2N}{5} + \frac{2N}{10} - \frac{2N}{13} + \frac{2N}{26} + \frac{4N}{65} - \frac{4N}{130} \\ &= N\left(1 - \frac{1}{2}\right)\left(1 - \frac{2}{5}\right)\left(1 - \frac{2}{13}\right). \end{aligned}$$

TABLE 6

N	$S(N)$	$D(N)$	$S(N)/D(N)$
100	16.261	15	1.016
200	28.252	28	1.009
300	39.800	42	0.948
400	50.696	51	0.994
500	61.344	62	0.989
600	71.763	68	1.055
700	81.656	78	1.047
800	91.345	87	1.050
900	101.075	92	1.099
1000	110.901	102	1.087
1100	119.913	112	1.071
1200	129.451	122	1.061
1300	138.223	128	1.080
1400	147.754	140	1.055
1500	156.790	150	1.045

For N small, $S(N)$ and $D(N)$ are nearly equal; e.g., $S(15) = 3.81$, $D(15) = 4$ As N increases, $S(N)$ gradually pulls ahead of $D(N)$, as is seen in Table 6.

The end effect

$$e(N) = S(N) - D(N)$$

is given by

$$e(N) = \sum_d (-1)^\alpha \sum_i \left\{ \frac{N}{d} - \left[\frac{N + A_i}{d} \right] \right\}. \tag{32}$$

Since the quantity in each brace is smaller in magnitude than unity, it is easy enough to bound $e(N)$. What is difficult to obtain is a sufficiently *good* bound—that is, to prove in general, the extensive cancellation of terms of opposite sign which occurs in the sum of (32). The essential difficulty stems from the very rapid increase in the number of terms, $2 \cdot 3^M$.

Techniques of deleting or combining terms, in sieve formulations of related problems, have been devised by Brun and others [11] but to date nothing sufficiently sharp has been developed. A general assessment of sieve techniques given by Selberg [12] is not encouraging.

Applied Mathematics Laboratory
David Taylor Model Basin
Washington 7, District of Columbia

1. G. H. HARDY & J. E. LITTLEWOOD, "Partitio numerorum III: On the expression of a number as a sum of primes," *Acta Math.*, v. 44, 1923, p. 48.
2. A. E. WESTERN, "Note on the number of primes of the form $n^2 + 1$," Cambridge Phil. Soc., *Proc.*, v. 21, 1922, p. 108–109.
3. DANIEL SHANKS, "A sieve method for factoring numbers of the form $n^2 + 1$," *MTAC*, v. 13, 1959, p. 78–86.
4. DANIEL SHANKS, "Quadratic residues and the distribution of primes," *MTAC*, v. 13, 1959, p. 272–284.
5. DANIEL SHANKS, "On the conjecture of Hardy and Littlewood concerning the number of primes of the form $n^2 + a$," *Notices*, Amer. Math. Soc., v. 6, 1959, p. 417. Abstract 559–52.

6. The numbers $b(s)$ also arise in an entirely different connection—they are related to the number of distinct *circular parity switches of order s*. See DANIEL SHANKS, "A circular parity switch and applications to number theory," *Notices*, Amer. Math. Soc., v. 5, 1958, p. 96. Abstract 543–7. It was in *this* connection that the author first noted the unusual proof of a special case of the Fermat "little" theorem—see (9a) above. Likewise it was in this connection that BERNARD ELPSAS, in a private communication to the author (Sept. 3, 1958), developed the formula (9).

7. E. LANDAU, *Aus der elementaren Zahlentheorie*, Chelsea, 1946, Part IV, Chap. 6–9.

8. FLETCHER, MILLER & ROSENHEAD, *Index of Mathematical Tables*, McGraw-Hill, 1946, p. 42, 43, p. 63. The correspondence between our notation and theirs is as follows: $L_1(s) = u_n$, $L_2(s) = p_n$, $L_{-2}(s) = q_n$, $L_3(s) = h_n$, and $L_{-3}(s) = t_n$.

9. A similar sieve argument was given for the twin prime problem in CHARLES S. SUTTON, "An investigation of the average distribution of twin prime numbers," *Jn. Math. Phys.*, v. 16, 1937, p. 1–42.

10. G. H. HARDY & E. M. WRIGHT, *An Introduction to the Theory of Numbers*, Oxford, 1938, p. 349.

11. ERNST TROST, *Primzahlen*, Basel, 1953, Chap. IX.

12. A. SELBERG, "The general sieve method and its place in prime number theory," *Proc.*, Inter. Congress Math., Cambridge, 1950, p. 286.