RECENT PROGRESS IN THE GOLDBACH PROBLEM

R. D. JAMES

1. Introduction. The problem under consideration had its origin in a letter written by Goldbach to Euler in 1742 [4].[1] In the letter, Goldbach made two conjectures concerning the representation of integers as a sum of primes. They are equivalent to

(A) *every even integer greater than 2 is a sum of two primes,*
and
(B) *every integer greater than 5 is a sum of three primes.*

The two conjectures are, of course, equivalent. If $2n-2=p_1+p_2$, then $2n=p_1+p_2+2$ and $2n+1=p_1+p_2+3$. Conversely, if $2n=p_1+p_2+p_3$ one of the primes must be 2 and $2n-2=p_1+p_2$.

An impressive collection of numerical evidence indicating the truth of the conjectures has accumulated in the years since Goldbach's letter was written, but is it not known to this day whether the conjectures are true or false. What progress has been made towards the solution of the problem has been through two principal methods of attack.

The first of these is the sieve method (see §§2 and 3) due originally to Brun [1], and improved by Rademacher [16], Esterman [5], Ricci [17], [18], and Buchstab [2], [3]. The best result by this method, due to Buchstab in 1940, is

every sufficiently large even integer is a sum of two integers, each having at most four prime factors.

The sieve method has also been used in combination with results on the density of sequences of integers. (See §4.) Contributions have been made by Schnirelmann [20], Landau [12], [13], Heilbronn, Landau, Scherk [7], and Ricci [19]. Schnirelmann proved that

every integer >1 is a sum of a finite number of primes.

The best result, due to Ricci in 1937, a minor improvement on the Landau, Heilbronn, Scherk result of 1936, is

every sufficiently large integer is a sum of at most 67 primes.

This result did not stand for long. It was superseded almost immediately by the Vinogradov result described in the next paragraph.

The second principal method is the analytic method developed by Hardy and Littlewood [6] in their series of papers on Partitio Numerorum. In 1937, Vinogradov [23] established an asymptotic formula for the number of representations of an odd integer as a sum of three primes. From this it follows that

every sufficiently large odd integer is a sum of three primes.

The work of Page [15] and Siegel [21] is an important part of the Vinogradov method. Vinogradov's result is close to Conjecture B and is now known in the literature as the Goldbach-Vinogradov Theorem. The same method also leads to a proof of the result

almost all even integers are sums of two primes.

The precise meaning of "almost all" is explained in §5.

Linnik [14] and Tchudakoff [22] have given different proofs of Vinogradov's result, but the Goldbach-Vinogradov Theorem still stands as the nearest approach of modern mathematics to the Goldbach conjecture of 1742.

2. **The sieve method. Brun's contribution.** The method of Brun is an adaptation of the ancient Sieve of Eratosthenes, in which the mechanical process of striking out the multiples of primes is replaced by an algebraic recursion formula. Brun's significant contribution is the way in which the formula is suitably modified. This important point will be explained later at the appropriate time.

Let d be a given positive integer, and let p_1, p_2, $\cdots$, p_k be the primes in order, which do not divide d and do not exceed y, where y is any positive number. Let a_i; b_i; $\cdots$; a_k, b_k be a set of integers with $0 \leq a_i < p_i$, $0 \leq b_i < p_i$, $a_i \neq b_i$, $i = 1, 2, \cdots, k$. Let a be any positive integer and x any positive number. Then, following Rademacher, Buchstab, and Landau, let $F(x; d, y) = F(x; d, y; a_i, b_i; p_i)$ denote the number of integers n which satisfy the conditions

(2.1)
$$n \leq x, \qquad n \equiv a \ (\text{mod } d),$$
$$(n - a_i)(n - b_i) \not\equiv 0 \ (\text{mod } p_i), \qquad i = 1, \cdots, k.$$

The arguments a, a_i, b_i, p_i, need not be written in the function since the results will hold for every a and every set a_i, b_i of the type described. $F(x; d, 1)$ is nothing but the number of integers $n \leq x$, $n \equiv a \ (\text{mod } d)$ and will be abbreviated $F(x; d)$.

The connection between $F(x; d, y)$ and the Goldbach problem is

indicated by the following considerations. Let $d=2$, $a=1$, $y=x^{1/u}$, where x is even integer and u is an integer not less than 2. Let $a_i = 0$, $b_i \equiv x$ (mod p_i) if $p_i \nmid x$; $b_i \not\equiv x$ (mod p_i) if $p_i | x$. Then the function $F(x; 2, x^{1/u})$ is the number of odd integers $n \leq x$ such that neither n nor $x - n$ is divisible by any prime not exceeding $x^{1/u}$. Hence all the prime factors of n and $x - n$ are greater than $x^{1/u}$ and there cannot be more than $u-1$ of them. If $u=2$ each of n and $x - n$ is either a prime or equal to 1. Thus, if it could be shown that $F(x; 2, x^{1/2}) \geq 2$, it would follow that there is at least one representation $x = n + (x-n)$ in which each of n and $x - n$ is a prime. Buchstab's result, quoted in the introduction, comes from establishing a similar inequality for $F(x; 2, x^{1/5})$.

The fundamental recursion formula is not difficult to prove. Clearly, the number of integers satisfying (2.1) is equal to the number of integers satisfying

$$(2.2) \quad \begin{aligned} n &\leq x, \quad n \equiv a \text{ (mod } d), \\ (n - a_i)(n - b_i) &\not\equiv 0 \text{ (mod } p_i), \quad\quad i = 1, 2, \cdots, k - 1, \end{aligned}$$

diminished by the number of integers satisfying

$$(2.3) \quad\quad\quad n \leq x, \quad n \equiv a \text{ (mod } d),$$

$$(2.4) \quad\quad\quad (n - a_k)(n - b_k) \equiv 0 \text{ (mod } p_k),$$

$$(n - a_i)(n - b_i) \not\equiv 0 \text{ (mod } p_i) \quad\quad i = 1, 2, \cdots, k - 1.$$

Since $a_k \neq b_k$, $p_k \nmid d$, the conditions (2.3), (2.4) are equivalent to

$$n \leq x, \quad n \equiv a' \text{(mod } dp_k); \quad n \leq x, \quad n \equiv a'' \text{ (mod } dp_k).$$

It follows that

$$(2.5) \quad \begin{aligned} F(x; d, p_k; a) &= F(x; d, p_{k-1}; a) \\ &\quad - F(x; dp_k, p_{k-1}; a') - F(x; dp_k, p_{k-1}; a''), \end{aligned}$$

where the arguments a, a', a'' are temporarily indicated in the function. Since the results are to hold for every a, a', a'' it is usual to write (2.5) symbolically in the form

$$(2.6) \quad\quad F(x; d, p_k) = F(x; d, p_{k-1}) - 2F(x; dp_k, p_{k-1}),$$

where, of course, if $k=1$, $F(x; d, p_1) = F(x; d) - 2F(x; dp_1)$. Equation (2.6) is the fundamental recursion formula.

From (2.6) it is but a short step to

$$F(x; d, p_k) = F(x; d) - 2\sum_{r=1}^{k} F(x; dp_r, p_{r-1})$$

and

$$F(x; d, p_k) = F(x; d) - 2\sum_{r=1}^{k} F(x; d p_r)$$

(2.7)
$$+ 4\sum_{r=1}^{k}\sum_{s=1}^{r-1} F(x; d p_r p_s) - \cdots$$

$$+ (-2)^k F(x; d p_k p_{k-1} \cdots p_2 p_1).$$

Since $F(x; d)$ is the number of integers $n \leq x$, $n \equiv a \pmod{d}$, $F(x; d)$ $= x/d + \theta$, where $|\theta| \leq 1$. Hence, from (2.7),

$$F(x; d, p_k) = (x/d)\{1 - 2\sum (1/p_r)$$

$$+ 4\sum\sum (1/p_r p_s) - \cdots$$

$$+ (-2)^k (1/p_k \cdots p_1)\} + R,$$

or

(2.8) $$F(x; d, p_k) = (x/d)\prod_{r=1}^{k} (1 - 2/p_r) + R,$$

where $|R|$ does not exceed the number of terms in the brackets, that is, $|R| \leq 2^{k+1} - 1$.

Unfortunately, formula (2.8) is of no help as it stands. The difficulty is that, if p_1, p_2, $\cdots$, p_k are the primes up to $x^{1/u}$ then $\prod(1 - 2/p_r) \sim c_1 u^2/\log^2 x$, and $k = \pi(x^{1/u}) \sim u x^{1/u}/\log x$, so that R is of higher order than the principal term in (2.8). This is where Brun's contribution comes in. He replaces the equation (2.8) by an inequality in which the number of terms is sufficiently reduced to make R of lower order than the principal term.

It would take too long to reproduce the details here. The interested reader may find them in Brun's original memoir or in Rademacher's paper. The final result is that, if k_1, k_2, $\cdots$, k_t are integers satisfying $1 \leq k_t \leq \cdots \leq k_1 \leq k$, then

(2.9) $$F(x; d, p_k) = (x/d)E - R$$

where $E = 1 - 2\sum(1/p_{r_1}) + 4\sum\sum(1/p_{r_1}p_{r_2}) - \cdots$, the summation indices satisfy

$$1 \leq r_1 \leq k,$$

$$1 \leq r_{2j} \leq k_j,$$

$$1 \leq r_{2j+1} \leq k_j,$$ $$j = 1, 2, \cdots, t,$$

$$1 \leq r_{2t+1} < r_{2t} < \cdots < r_2 < r_1,$$

and R does not exceed the number of terms in the product

$$(1 - 2\sum (1/p_{r_1}))(1 - 2\sum (1/p_{r_2})) \cdots ,$$

that is, $R \leq (2k+1)(2k_1+1)^2 \cdots (2k_t+1)^2$.

By suitably choosing the integers $k_1, k_2, \cdots, k_t$ it can be shown, for example, that

$$(2.10) \qquad E > \frac{3}{10} \prod_{r=1}^{k} (1 - 2/p_r), \qquad R = O(p_k^{79/10}).$$

This is Rademacher's result, and illustrates the point that reducing the number of terms only changes the principal term to a numerical constant times its former value, but reduces the remainder term by a considerable amount.

It follows from (2.9) and (2.10) that, if $p_1, p_2, \cdots, p_k$ are the primes up to $x^{1/8}$, then $F(x; 2, x^{1/8}) > c_2 x/\log^2 x - c_3 x^{79/80}$. Hence every sufficiently large even integer x is a sum $x = n + (x-n)$ of two odd integers, each with at most seven prime factors. The reduction from seven to four prime factors is accomplished by the method of Buchstab, which is discussed in the next section.

It should be noted in passing that the Brun method will yield upper bounds for $F(x; d, p_k)$ as well. It is only necessary to stop the series for E with r_{2t} instead of r_{2t+1}. Such upper bounds are needed in Buchstab's work (§3) and in the density-sieve method (§4).

3. **The sieve method. Buchstab's contribution.** Buchstab's results are based on a modification of the fundamental recursion formula (2.6), and uses a different estimate for the number of integers satisfying (2.3), (2.4). (See also James [8].)

The number of integers n such that

$$(3.1) \qquad \begin{array}{cc} n \leq x, & n \equiv a \pmod{d}, \\ n \equiv a_k \pmod{p_k}, & (n - a_i)(n - b_i) \not\equiv 0 \pmod{p_i} \end{array}$$

is equal to the number of integers m such that

$$a_k + mp_k \leq x, \qquad a_k + mp_k \equiv a \pmod{d},$$
$$(a_k + mp_k - a_i)(a_k + mp_k - b_i) \not\equiv 0 \pmod{p_i},$$

or

$$(3.2) \qquad \begin{array}{cc} m \leq (x - a_i)/p_k, & m \equiv a' \pmod{d}, \\ (m - a_i')(m - b_i') \not\equiv 0 \pmod{p_i}. \end{array}$$

The number of integers satisfying (3.2) and hence (3.1) is $F((x-a_k)/p_k; d, p_{k-1})$. Similarly the number of integers satisfying

(3.1) with a_k replaced by b_k is $F((x-b_k)/p_k; d, p_{k-1})$, and therefore

(3.3)
$$F(x; d, p_k) = F(x; d, p_{k-1})$$
$$- F((x - a_k)/p_k; d, p_{k-1}) - F((x - b_k)/p_k; d, p_{k-1}).$$

Since $0 \leqq a_k < p_k$, $0 \leqq b_k < p_k$, the difference between the last two terms on the right of (3.3) and $2F(x/p_k; d, p_{k-1})$ is at most 2. Therefore

(3.4) $F(x; d, p_k) = F(x; d, p_{k-1}) - 2F(x/p_k; d, p_{k-1}) + \theta_k,$

where $0 \leqq \theta_k \leqq 2$. This is the recursion formula in the Buchstab method.

Let u and v be integers such that $2 \leqq u < v$, and let $p_t, p_{t+1}, \cdots, p_{k-1}, p_k$ be the primes in order for which

$$p_t \leqq x^{1/v} < p_{t+1} < \cdots < p_k \leqq x^{1/u} < p_{k+1}.$$

By repeated application of (3.4) it follows that

(3.5)
$$F(x; d, x^{1/u}) = F(x; d, x^{1/v})$$
$$- 2 \sum_{i=t}^{k} F(x/p_i; d, p_{i-1}) + \sum_{i=t}^{k} \theta_i.$$

If a lower bound for $F(x; d, x^{1/v})$ and upper bounds for each term in the sum are known, formula (3.5) provides a lower bound for $F(x; d, x^{1/u})$.

Formula (3.5) is one of the important parts of Buchstab's contribution. The larger v is the easier it is to get a good lower bound for $F(x; d, x^{1/v})$ and reasonably good upper bounds for the terms in the sum, using the Brun methods. This leads to a better lower bound for $F(x; d, x^{1/u})$ than can be obtained with the sieve method by direct calculation.

Another important part of the method is the way in which (3.5) is used. It can be shown by the Brun method that there exists a pair of non-negative step functions $g_0(u)$, $G_0(u)$ defined for $u \geqq 2$ such that

(3.6) $g_0(u)cx/\log^2 x \leqq F(x; d, x^{1/u}) \leqq G_0(u)cx/\log^2 x$

for all sufficiently large x, where c is an absolute constant. In Buchstab's first paper [2], for example, he starts with

(3.7)
$$g_0(u) = \begin{cases} 0, & 2 \leqq u < 10, \\ 98, & u = 10; \end{cases}$$
$$G_0(u) = 101.6, \quad 2 \leqq u \leqq 10.$$

It then follows from (3.6) that $g_0(u)$ and $G_0(u)$ may be replaced by

$g_1(u)$ and $G_1(u)$, respectively, where

$$g_1(u) < g_0(v) - 2 \int_{u-1}^{v-1} G_0(z)(z+1)z^{-2}dz,$$

$$G_1(u) > G_0(v) - 2 \int_{u-1}^{v-1} g_0(z)(z+1)z^{-2}dz,$$

for $2 \leqq u < v$.

The process may be repeated as long as the new bounds are better than the old. Starting with the rough estimate (3.7), Buchstab [2] obtains the following:

(3.8)

$$g_3(u) = \begin{cases} 0, & 2 \leq u < 6, \\ 0.3, & 6 \leq u < 8, \\ 53.51, & 8 \leq u < 9, \\ 75.58, & 9 \leq u < 10, \\ 98.0, & u = 10; \end{cases}$$

$$G_3(u) = \begin{cases} 67.58, & 2 \leq u \leqq 7, \\ 72.86, & 7 < u \leqq 8, \\ 85.1, & 8 < u \leqq 9, \\ 101.6, & 9 < u \leqq 10. \end{cases}$$

In a second note, Buchstab [3] starts with the set (3.8) with the addition of

$$g_3(u) = \begin{cases} 98.0, & 10 \leqq u < 15, \\ 244.9997, & u = 15; \end{cases}$$

$$G_3(u) = \begin{cases} 144.1328, & 10 < u \leqq 12, \\ 196.0022, & 12 < u \leqq 14; \end{cases}$$

and works down to $g_i(5) = 0.96438$. From this follows the result quoted in the introduction, the best so far obtained by this method, that every sufficiently large even integer x is a sum $x = n + (x - n)$, in which each of n and $x - n$ has at most four prime factors.

It is not entirely improbable that the newer computation methods may be used to work further down, perhaps even to $g_i(2) > 0$. If this should be the case, Goldbach's first conjecture and other similar ones would be proved, at least for sufficiently large integers.

4. The density-sieve method. While it is true that the results of

the density sieve method were soon superseded, the method itself is still of interest.

It is well known that, if $\pi(x)$ denotes, as usual, the number of primes not greater than x, then $\pi(x)/x \to 0$ as $x \to \infty$. On the other hand, Schnirelmann [20] proved that, if $M(x)$ denotes the number of integers $n \leq x$ which are a sum of two primes, then $M(x)/x \geq 1/h$, $x \geq 4$, where h is a constant greater than 1. In other words, the set of integers $x \geq 4$, which are sums of two primes, has positive density. It follows from standard theorems on the density of sequences of integers that every integer $x > 1$ is a sum of a finite number of primes. This is the Schnirelmann theorem. (See Landau [13].)

The Schnirelmann result depends on the sieve method in the following way. Let $A(m)$ denote the number of representations of an integer m as a sum of two primes, and let $M(x)$ be the function defined above. Then, by the Cauchy-Schwarz inequality,

$$\left(\sum_{m=4}^{x} A(m) \right)^2 \leq M(x) \sum_{m=4}^{x} A^2(m)$$

and

(4.1) $$M(x) \geq \left(\sum_{m=4}^{x} A(m) \right)^2 \Big/ \left(\sum_{m=4}^{x} A^2(m) \right).$$

What is needed then is a lower bound for the numerator, and an upper bound for the denominator, of the term on the right of (4.1). A suitable lower bound is easily found by noting that

(4.2) $$\sum A(m) \geq \pi^2(x/2) > c_1 x^2/\log^2 x.$$

The sieve method enters into the problem of finding an upper bound. Clearly, the number of solutions of $m = p + p'$, where p and p' are primes, each greater than $m^{1/2}$, does not exceed $F(m; 1, m^{1/2})$. The number of solutions where $p \leq m^{1/2}$ or $p' \leq m^{1/2}$ does not exceed $2m^{1/2}$. Hence, an upper bound for $F(m; 1, m^{1/2})$ provides an upper bound for $A(m)$ and finally for $\sum A^2(m)$.

The details are given by Landau [13] and will be omitted here. The result is

(4.3) $$F(m; 1, m^{1/2}) < c_2(m/\log^2 m) \prod_{p \mid m} (1 + 1/p),$$

where c_2 is a constant. It then follows that $\sum A^2(m) < c_3 x^3/\log^4 x$, and then from (4.1) and (4.2) that $M(x)/x \geq 1/h$, $h = c_3/c_1^2$.

In the Heilbronn, Landau, Scherk paper [7], $M(x)$ is the number of integers $m \leq x$ for which $30m = p + p'$, and the inequality (4.3) is

replaced by

$$\limsup_{m\to\infty,\,m\equiv 0(\mathrm{mod}\,30)} \frac{F(m;\,30,\,m^{1/2})}{(m/\log^2 m)\prod_{p\mid m}(1+1/(p-2))} \leqq c_4$$

where c_4 is a constant. There are similar changes for $\sum A(m)$ and $\sum A^2(m)$. By a careful consideration of the numerical constants involved, it is shown that

$$\liminf_{x\to\infty} M(x)/x \geqq \frac{\liminf x^{-4}\log^4 x(\sum A(m))^2}{\limsup x^{-3}\log^4 x(\sum A^2(m))} \geqq \frac{1}{34}.$$

Finally, from a general result of Khintchine [10] it follows that every sufficiently large integer u may be written in the form

$$u = \sum_{i=1}^{34} m_i + r, \qquad\qquad 0 \leqq r \leqq 33.$$

Hence, for each j such that $2\leqq j\leqq 31$,

$$v = 30u + j = \sum_{i=1}^{34}(p_i + p_i') + s$$

where $2\leqq s\leqq 1021$.

It is easily verified that every s in the given range is a sum of at most three primes, and therefore every sufficiently large integer v is a sum of at most 71 primes. Ricci [19] reduced the number 71 to 67, but the improvement is one of detail and no new principle is involved.

5. **The analytic method.** The method pioneered by Hardy and Littlewood [6] is based on the idea of a generating function and on the Cauchy residue theorem. Thus, if $\{a_j\}$ is an infinite sequence of positive integers with $a_j < a_{j+1}$, the function $f(z) = \sum_{j=1}^\infty z^{a_j}$, $|z| < 1$, is the generating function of the sequence. If $A_s(N)$ denotes the number of representations of an integer N as a sum of s integers of the sequence, then

$$f^s(z) = \sum_{j=1}^\infty A_s(j)z^j,$$

and

(5.1) $A_s(N) = (1/2\pi i)\int_C f^s(z)z^{-N-1}dz,$

where C is the circle $|z| = r < 1$.

The question of convergence of the series, but not the difficulty of the proof, is avoided in the Vinogradov [23] version by considering not $f(z)$ but $F(x) = \sum_{a_j \le N} e(a_j x)$ where $e(a_j x)$ denotes $e^{2\pi i a_j x}$, and x is real. Then (5.1) is replaced by

$$A_s(N) = \int_0^1 F^s(x)e(-Nx)dx.$$

If the sequence a_j is the sequence of primes, the number of representations of an integer N as a sum of three primes is

(5.2) $$A(N) = \int_0^1 F^3(x)e(-Nx)dx,$$

where

(5.3) $$F(x) = \sum_{p \le N} e(px).$$

Formula (5.2) is Vinogradov's starting point, and since his work is not easily accessible, some details may be of interest.

As is well known, the first step is to divide the interval $(0, 1)$ into sub-intervals whose centers are the points $x = a/q$, $(a, q) = 1$, $0 \le a \le q$, $1 \le q \le Q$, corresponding to a Farey series of some order Q. The sub-intervals corresponding to the smaller values of q form the intervals of Class I (major arcs), and the remaining sub-intervals, the intervals of Class II (minor arcs). If n is written for $\log N$ and if h is a number greater than 3, the classes for the Goldbach problem are defined by

I: $x = a/q + y$, $1 \le q \le [n^{3h}]$, $|y| \le n^{3h}/N$,

II: $x = a/q + y$, $[n^{3h}] < q \le N/n^{3h}$, $|y| \le n^{3h}/qN$.

If x belongs to an interval of Class I, the sum $F(x)$ is written in the form

$$F(a/q + y) = \sum_{r=0}^{q-1} \sum_{p \le N, p \equiv r \pmod{q}} e(p(a/q + y))$$

$$= \sum_{r=0, (r,q)=1}^{q-1} e(ra/q) \sum_{p \le N, p \equiv r \pmod{q}} e(py) + O\left(\sum_{p \mid q} 1\right).$$

The inner sum is approximated by $(1/\phi(q))J(y)$, where

$$J(y) = \int_M^N (e(yt)/\log t)dt, \qquad M = N^{1/2}.$$

In Vinogradov's paper the lower limit in $J(y)$ is 2, but it is better to

use M (see James and Weyl [9]). From the results of Page [15] and Siegel [21], the error made in approximating to the inner sum is $O(N/\phi(q)n^{15h+1})$.

Since

$$\sum_{r=0,\,(r,q)=1}^{q-1} e(ra/q) = \mu(q),$$

it follows in turn that

$$F(a/q + y) = (\mu(q)/\phi(q))J(y) + \text{an error term},$$
$$F^3(a/q + y) = (\mu(q)/\phi^3(q))J^3(y) + \text{an error term}.$$

The part of the integral in (5.2) over an interval of Class I is

$$\int_{-n^{3h}/N}^{n^{3h}/N} F^3(a/q + y)e(-N(a/q + y))dy$$

$$= (\mu(q)/\phi^3(q))e(-Na/q)\int_{-n^{3h}/N}^{n^{3h}/N} J^3(y)e(-Ny)dy$$

$$+ \text{an error term}.$$

The limits in the integral may be extended from $-\infty$ to ∞ without making the error term any worse. Hence

$$\sum\int_{\mathrm{I}} F^3(x)e(-Nx)dx$$

$$(5.4) \qquad = \sum_{q=1}^{[n^{3h}]} \sum_{a=0,\,(a,q)=1}^{q-1} (\mu(q)/\phi^3(q))e(-Na/q)\int_{-\infty}^{\infty} J^3(y)e(-Ny)dy$$

$$+ \text{an error term}.$$

If the series for q is extended from 1 to ∞ it becomes the so-called singular series, $S(N)$, and, if $R(N)$ is written for the integral in (5.4), then

$$(5.5) \qquad \sum\int_{\mathrm{I}} F^3(x)e(-Nx)dx = S(N)R(N) + \text{an error term}.$$

The singular series is a purely arithmetical function of an integer N (it is zero when N is even), whereas $R(N)$ takes only the size of N into account, being defined for all real values of N. In fact $R(N) \sim N^2/2n^3$ as $N \to \infty$.

One of Vinogradov's fundamental contributions is an estimate for $F(x)$ when x belongs to an interval of Class II. Let D denote the

product $p_1 p_2 \cdots p_k$ of all primes not exceeding $N^{1/2}$. Then

$$F(x) = \sum_{p \leq N} e(px) = \sum_{m=2, (m,D)=1}^{N} e(mx) + O(N^{1/2}).$$

If $\mu(d)$ is the Möbius function and

$$F_d(x) = \sum_{m=2, d|m}^{N} e(mx),$$

then (the sieve again)

(5.6) $$F(x) = \sum_{d \leq N, d|D} \mu(d) F_d(x) + O(N^{1/2}).$$

It follows from (5.6) and Vinogradov's important results on exponential sums that, if x belongs to an interval of Class II, then

(5.7) $$F(x) = O(N/n^{h-2}),$$

where h is an arbitrary number greater than 3. Hence

$$\sum \int_{II} F^3(x) e(-Nx) dx = O(N/n^{h-2}) \int_0^1 |F(x)|^2 dx$$

(5.8)

$$= O(N/n^{h-2}) \pi(N) = O(N^2/n^{h-1}).$$

This is of lower order than the principal term in (5.5), and hence, by (5.2), (5.5), and (5.8), $A(N) = S(N)R(N) + O(N^2/n^{h-1})$.

There is an asymptotic expansion (James and Weyl [9])

$$R(N) \sim (N^2/n^3)(1/2 + b_1/n + b_2/n^2 + \cdots),$$

where the error in stopping with the term b_{j-1}/n^{j-1} is less than C_j/n^j, where C_j does not depend on N. Hence

(5.9) $$A(N) \sim S(N)(N^2/n^3)(1/2 + b_1/n + b_2/n^2 + \cdots).$$

The singular series has a product representation (Hardy and Littlewood [6], Landau [11])

$$S(N) = \prod_{p} (1 + 1/(p-1)^3) \prod_{p|N} (1 - 1/(p^2 - 3p + 3)).$$

For even N the first factor of the second product is zero, but for odd N it can be shown that $S(N) > C$, where C is a positive constant. By (5.9), $A(N) > 0$ if N is sufficiently large, so that every sufficiently large odd integer is a sum of three primes.

For reasons which are not too clear, the more recent papers mark a return to the original Hardy-Littlewood method (Hardy and Little-

wood [6], Landau [11]). Hardy and Littlewood start with the function

$$(5.10) \qquad\qquad f(z) = \sum_{p \geq 2} \log p\, z^p$$

and so does Tchudakoff [22]. Linnik [14], on the other hand, begins with

$$(5.11) \qquad\qquad F(z) = \sum_{m=1}^{\infty} \Lambda(m) z^m,$$

where $\Lambda(m) = \log p$ if $m = p^t$, $t \geq 1$, and $\Lambda(m) = 0$, otherwise.

Each of the functions $f(z)$ and $F(z)$ is easier to treat by classical methods than is $\sum z^p$, but there is no essential difference. Starting from (5.10) the first step is to follow the plan of the Vinogradov method, that is, to write $z = e^{-1/N} e(a/q + y)$, $e^{-1/N}$ being a convergence factor, and

$$f(z) = \sum_{r=0,\,(r,q)=1}^{q-1} \;\; \sum_{p \geq 2,\, p \equiv r \,(\mathrm{mod}\ q)} \log p\, e^{-p/N} e(p(a/q + y))$$

$$+ \text{ an error term.}$$

It is well known that (Landau [11])

$$\sum_{\chi} \overline{\chi}(r)\chi(p) = \begin{cases} \phi(q) & \text{if } \quad p \equiv r \ (\mathrm{mod}\ q), \\ 0 & \text{if } \quad p \not\equiv r \ (\mathrm{mod}\ q), \end{cases}$$

where the summation is over all characters mod q. Hence $f(z)$ is equal to

$$(1/\phi(q)) \sum_{\chi} \sum_{r} e(ra/q)\overline{\chi}(r) \sum_{p} \chi(p) \log p\, e^{-p/N} e(py)$$

plus an error term. Also, the inner sum may be replaced by

$$\sum_{m=2}^{\infty} \Lambda(m)\chi(m) e^{-m/N} e(my)$$

with an error $O(N^{1/2}/n^{3/2})$, $n = \log N$. Thus

$$f(z) = (1/\phi(q)) \sum_{\chi} \sum_{r} e(ra/q)\overline{\chi}(r) \sum_{m} \Lambda(m)\chi(m) e^{-m/N} e(my)$$

$$+ O(q^{1/2} N^{1/2}/n^{3/2}).$$

Linnik, starting with (5.11) instead of (5.10), arrives at the same expression for $F(z)$, but with a slightly different error term.

By the classical formula of Mellin,

$$\sum_{m} \Lambda(m)\chi(m)e^{-m/N}e(my)$$

$$= (1/2\pi i) \int_{2-\infty i}^{2+\infty i} w^{-s}\Gamma(s)(L'(s,\chi)/L(s,\chi))ds,$$

where $w=1/N-2\pi iy$, $|y| \leq n^{3h}/N$ (major arc). It is here that the original Hardy-Littlewood paper ran into trouble. Clearly, the zeros of $L(s,\chi)$ play an important part in the evaluation of the integral, and at that time information concerning the density of these zeros was not precise enough to be of help. The remainder of Hardy and Little-wood's paper depends on an unproved hypothesis concerning the zeros. No further progress was made until the appearance of the papers of Page and Siegel.

Linnik and Tchudakoff base their work on the approximate functional equations for the L-functions, analogous to the famous approximate functional equation for the zeta function, and on improved estimates of the number of zeros of the L-functions in the critical strip. Their work involves some of the most delicate and intricate analysis written on the subject. Since Tchudakoff's work is given in detail (and in English) it is not necessary to reproduce it here. The final result is that, if $Q(N) = \int_C f^3(z)z^{-N-1}dz = \sum_{p+p'+p''=N} \log p \log p'$ $\log p''$, where C is the circle $|z| = e^{-1/N}$, then

$$Q(N) = S(N)((N+1)(N+2)/2) + \text{an error term.}$$

From this it follows that $A(N)\sim S(N)(N^2/2n^3)$, but not the more complete result (5.9).

No doubt something is gained by appealing to classical results in the theory of the distribution of primes, but Vinogradov's appears to be the more direct approach.

The analytic method also has something to contribute towards the solution of Goldbach's first conjecture. Let $B(x)$ denote the number of even integers $m \leq x$ which are *not* a sum of two primes. Then, by the Vinogradov, Linnik, or Tchudakoff methods, it can be shown that

$$\lim_{x\to\infty} B(x)/x = 0.$$

This is the precise meaning of the statement that almost all even integers are sums of two primes. As Landau [11] puts it in the introduction to his chapter on the Goldbach conjecture, "The Goldbach conjecture is false for at most 0% of all even integers. This at most 0% does not exclude, of course, the possibility that there are infinitely many exceptions."

References

1. V. Brun, *Le crible d'Eratosthène et le théorème de Goldbach*, Videnskabs-selskabet i Kristiania Skrifter I, Matematisk-Naturvidenskapelig Klasse No. 3 (1920) pp. 1–36.

2. A. A. Buchstab, *Neue Verbesserungen in der Methode des Eratosthenischen Siebes*, Rec. Math. (Mat. Sbornik) N.S. vol. 4 (1938) pp. 375–387.

3. ———, *Sur la décomposition des nombres pairs en somme de deux composantes dont chacune est formée d'un nombre borné de facteurs premiers*, C. R. (Doklady) Acad. Sci. URSS. N.S. vol. 29 (1940) pp. 544–548.

4. L. E. Dickson, *History of the theory of numbers*, vol. 1, New York, 1934.

5. T. Estermann, *Eine neue Darstellung und neue Anwendungen der Viggo Brunschen Methode*, J. Reine Angew. Math. vol. 168 (1932) pp. 106–116.

6. G. H. Hardy and J. E. Littlewood, *Some problems of Partitio Numerorum III*, Acta Math. vol. 44 (1922).

7. H. Heilbronn, E. Landau, and P. Scherk, *Alle grossen ganzen Zahlen lassen sich als Summe von höchstens* 71 *Primzahlen darstellen*, Časopis pro Pěstování Mathematiky a Fysiky vol. 65 (1936) pp. 117–141.

8. R. D. James, *On the sieve method of Viggo Brun*, Bull. Amer. Math. Soc. vol. 49 (1943) pp. 422–432.

9. R. D. James and H. Weyl, *Elementary note on prime number problems of Vinogradov's type*, Amer. J. Math. vol. 64 (1942) pp. 539–552.

10. A. Khintchine, *Zur additive Zahlentheorie*, Rec. Math. (Mat. Sbornik) vol. 39 (1932) pp. 27–34.

11. E. Landau, *Vorlesungen über Zahlentheorie*, vol. 1, Leipzig, 1927.

12. ———, *Die Goldbach Vermutung und der Schnirelmannsche Satz*, Nach. Ges. Wiss. Göttingen (1930) pp. 255–276.

13. ———, *Über einige neuere Fortschritte der additiven Zahlentheorie*, Cambridge, 1937.

14. U. V. Linnik, *A new proof of the Goldbach-Vingradov theorem*, Rec. Math. (Mat. Sbornik) N.S. vol. 19 (1946).

15. A. Page, *On the number of primes in an arithmetic progression*, Proc. London Math. Soc. (2) vol. 39 (1935) pp. 116–141.

16. H. Rademacher, *Beitrage zur Viggo Brunschen Methode in der Zahlentheorie*, Abh. Math. Sem. Hamburgischen Univ. vol. 3 (1924) pp. 12–30.

17. G. Ricci, *Sui grandi divisori primi delle coppie di interi in posti corrispondenti di due progressioni aritmetiche. Applicazione del metodo di Brun*, Annali di Matematica Pura ed Applicata (4) vol. 11 (1932–1933).

18. ———, *Ricerche aritmetiche sui polinomi*, I, II, Rend. Circ. Mat. Palermo vol. 57 (1933), vol. 58 (1934).

19. ———, *Su la congettura di Goldbach e la costante di Schnirelmann*, Annali della R. Scuola Normale Superiore di Pisa (2) vol. 6 (1937).

20. L. Schnirelmann, *Ob additiwnich swoistwach tschisel*, Izvestiya Donskowo Polytechnitscheskowo Instituta (Nowotscherkask) vol. 14 (1930) pp. 3–28.

21. C. L. Siegel, *Über die Classenzahl quadratischer Zahlkorper*, Acta Arithmetica vol. 1 (1935) pp. 83–86.

22. N. Tchudakoff, *On Goldbach-Vinogradov's theorem*, Ann. of Math. (2) vol. 48 (1947) pp. 515–545.

23. I. M. Vinogradov, *Some theorems concerning the theory of primes*, Rec. Math. (Mat. Sbornik) N.S. vol. 2 (1937) pp. 179–195.

University of British Columbia