An elementary method in prime number theory

by

R. C. VAUGHAN (London)

1. Introduction. In number theory one often desires to estimate sums of the form

\[ \sum f(p) \]

(or equivalently \( \sum A(n)f(n) \) with \( A \) von Mangoldt's function) where, for example, \( f(n) \) is an exponential function \( e(P(n)) = e^{i\pi F(n)} \), or a character \( \chi(n) \). The techniques for the estimation of such sums, whether analytic ([1], [3], [6]) or elementary ([7], [8], [9]) invariably relate such sums to bilinear forms of the kind either

(I) \[ \sum_{m} \sum_{n} a_{m}f(mn) \]

or

(II) \[ \sum_{m} \sum_{n} a_{m}b_{n}f(mn) \].

Suppose that the range for \( n \) in (I) is \([1, X]\), or equivalently that \( f(n) = 0 \) when \( n > X \). Then as a fairly general principle the estimates for (I) are good provided that \( m \ll M \) with \( M \) small compared with \( X \), and those for (II) are good provided that \( m \ll M, n \ll N \) with both \( M \) and \( N \) small compared with \( X \).

One method (I. M. Vinogradov [9]) of carrying out this procedure is to use the sieve of Eratosthenes to write

\[ f(1) = \sum_{p \leq X} f(p) = \sum_{m \leq M} \sum_{n \leq N} \mu(m)f(mn), \]

where \( P \) is the product of those prime numbers not exceeding \( \sqrt{X} \). The right hand side of this is of type (I), but has the defect of including \( m \) that are close to \( X \). In order to treat these \( m \), Vinogradov has to introduce a combinatorial argument, which for the sharpest estimates is quite involved, that allows him to relate this portion of the expression to bilinear forms of type (II).
In [7] and [8] an elementary method was introduced and developed which avoids these combinatorial difficulties. Consider the identity
\[
(2) \quad \sum_{n} g(1, n) + \sum_{m \leq n} \sum_{d | m} \mu(d) g(m, n) = \sum_{d | m} \sum_{r} \sum_{n} \mu(d) g(dr, n),
\]
which holds for any double sequence \( g(m, n) \) for which the right hand side converges absolutely, and is an immediate consequence of the relation
\[
\sum_{d | m} \mu(d) = 0 \quad (1 < m \leq n).
\]
Let
\[
g(m, n) = \begin{cases} A(n) f(mn) & (v < n \leq X/n), \\ 0 & \text{otherwise}. \end{cases}
\]
Then, by (2),
\[
(3) \quad \sum_{n \leq X} A(n) f(n) = S_1 - S_2 - S_3,
\]
where
\[
(4) \quad S_1 = \sum_{d | m} \sum_{d | k} \mu(d) (\log k) f(dk),
\]
\[
(5) \quad S_2 = \sum_{r | m} \sum_{r | k} \left( \sum_{d | m} \mu(d) A(n) \right) f(kr),
\]
\[
(6) \quad S_3 = \sum_{n < u \leq X/n} \sum_{d | u} \mu(d) A(n) f(mn).
\]
Now
\[
(7) \quad S_1 \leq \int_{1}^{X} S_1(a) \frac{da}{a}
\]
where
\[
(8) \quad S_1(a) = \sum_{d | m \leq a(X/a)} \sum_{0 < h < d} \mu(d) f(dk).
\]
Clearly both \( S_1(a) \) and \( S_2 \) are of type (I) above, whilst \( S_3 \) is of type (II). Thus suitable choices for \( u \) and \( v \) will often ensure that the corresponding estimates are good. We further remark that on some occasions the sum \( S_1 \) may be more sharply estimated by breaking it into two parts and treating the second part as a type (II) sum.

In [7] and [8] the above method was applied in the case \( f(n) = e(\alpha n) \). The purpose here is to show how the method can be applied to give a proof of the Bombieri–I. I. Vinogradov theorem concerning the average error term in the distribution of prime numbers in arithmetic progressions. Two essential ingredients (see [1], [3] or [6]) in the proof of this theorem are the Siegel–Walfisz theorem and a mean value theorem giving a bound for
\[
(9) \quad T(Y, Q) = \sum_{\psi(q) \leq Y} \frac{q}{\phi(q)} \sum_{\chi \leq Y} \max_{\nu \leq \nu} \psi(X, \chi),
\]
where \( \sum_{\chi} \) denotes summation over the primitive characters \( \chi \) modulo \( q \). The technique described above enables one to give an elementary proof of the following theorem.

**Theorem.** Suppose that \( q \gg 1 \), \( Y \gg 2 \), \( L = \log X \). Then
\[
T(Y, Q) \ll \varepsilon(Y + \varepsilon^2 + Y^{12} Q^4).
\]

As an easy consequence of this and the Siegel–Walfisz theorem one has the corollary.

**Corollary (Bombieri–Vinogradov).**
\[
\sum_{\psi(q) \leq Y} \max_{\chi \leq Y} \left| \psi(X, \chi, a) - \frac{X}{\phi(q)} \right| \ll Y (\log Y)^{-d} + Y^{12} Q L^4.
\]

**2. Lemmata.** The first lemma is an immediate consequence of the large sieve inequality (see, for instance, Gallagher [9] or (1.4) of [4]), the proofs of which are entirely elementary and Cauchy's inequality.

**Lemma 1.** Suppose that \( a_m (m = 1, \ldots, M) \) and \( b_n (n = 1, \ldots, N) \) are complex numbers. Then
\[
\sum_{\psi(q) \leq Y} \frac{q}{\phi(q)} \sum_{\chi \leq Y} \left| \sum_{m=1}^{M} \sum_{n=1}^{N} a_m b_n \chi(mn) \right| \ll \left( \sum_{M} + Q^2 \right)^{2M} \sum_{M} \left| a_m \right|^2 \sum_{N} \left| b_n \right|^2 12^4.
\]

The proof of the theorem rests on a maximal version of this.

**Lemma 2.** On the premises of Lemma 1 we have
\[
\sum_{\psi(q) \leq Y} \frac{q}{\phi(q)} \sum_{\chi \leq Y} \max_{X \leq T} \left| \sum_{m=1}^{M} \sum_{n=1}^{N} a_m b_n X(mn) \right| \ll \left( \sum_{M} + Q^2 \right)^{2M} \sum_{M} \left| a_m \right|^2 \sum_{N} \left| b_n \right|^2 12^4 \log YMN.
\]

**Proof.** Let
\[
C = \int_{-\infty}^{\infty} \frac{\sin \alpha}{\alpha} \, d\alpha,
\]

γ > 0, and δ(β) = 1 when 0 < β < γ, δ(β) = 0 when β > γ. Then C > 0 and it is easily seen that for A ≥ 1, β ≥ 0, β ≠ γ, we have

\[ δ(β) = \int \frac{sin(γa)}{a} da + O \left( \frac{1}{A|γ - β|} \right). \]

Let γ = log(\lfloor Y \rfloor + \frac{1}{2}), β = log mn. Thus

\[ \sum_{mn \leq X} \sum_{a,b} a^m b^n \chi(mn) \sin(γa) = \int \sum_{m} \sum_{n} a^m b^n \chi(mn) \sin(γa) da + O \left( \frac{X}{A} \sum_{m} \sum_{n} |a^m b^n| \right). \]

The desired conclusion now follows easily from Lemma 1 on taking A = YM N.

3. Proof of the theorem. If Q^2 > Y, then the theorem follows at once from Lemma 2 on taking M = 1, a_1 = b_1 = A(n). Hence it can be assumed that Q^2 ≤ Y.

Let

\[ u = v = \min(Q^2, Y^{1/α}, Y^{-1}). \]

By applying Lemma 2 as in the case Q^2 > Y it is easily seen that

\[ \sum_{0 < q < \varphi(q)^{1/2}} |\varphi(X, χ)| \leq (u^2Q + uQ^2) L^2. \]

Hence, on writing f(n) = χ(n) in (3), to prove the theorem it suffices, by (9), to show that for j = 1, 2, 3 the sum

\[ T_j = \sum_{0 < q < \varphi(q)} \sum_{m < X} \max_{u < x < X} |s_j| \]

satisfies

\[ T_j \leq L^2(Y + Y^{1/2}Q + Y^{1/3}Q^2). \]

By (6),

\[ T_j \leq \sum_{M \in \mathcal{M}} T_3(M) \]

where

\[ \mathcal{M} = \{ k \in \mathbb{N} : k = 0, 1, \ldots, \log Y / log 2 \} \]

and

\[ T_3(M) = \sum_{0 < q < \varphi(q)} \sum_{m < X} \max_{u < x < X} \left( \sum_{M < m < M + 1} \sum_{n < Y / M} \left( \sum_{d | n} \mu(d) \right) A(n) \chi(mn) \right). \]

By Lemma 2,

\[ T_1(M) \leq \left( \sum_{\varphi(q)} (Y \varphi(q)^{-1} + Q^2) \sum_{m < X} \delta(m)^2 \sum_{n < Y / M} \sum_{d | n} \mu(d) A(n) \chi(mn) \right)^{1/2} \log Y \]

\[ \leq \log Y \left( Y + Y^{1/2}M^{-2}Q + Y M^{-1/2}Q + Y^{1/3}Q^2 \right). \]

This easily gives (11) with j = 3.

By (6),

\[ T_2 \leq T_2^{1/2} + T_2^{1/2} \]

where

\[ T_2 = \sum_{0 < q < \varphi(q)} \sum_{x < X} \max_{u < x < X} \left( \sum_{M < m < M + 1} \sum_{n < Y / M} \mu(d) A(n) \chi(mn) \right). \]

and

\[ T_2^{1/2} \leq \sum_{0 < q < \varphi(q)} \max_{x < X} \left( \sum_{M < m < M + 1} \sum_{n < Y / M} \mu(d) A(n) \chi(mn) \right). \]

This is expressed directly via the Pólya-Vinogradov inequality (observe that Schur's proof [5] is elementary). Therefore

\[ T_2 \leq \left( Y + Y^{1/2}Q + Y^{1/3}Q^2 \right). \]

The sum T_2 is estimated directly via the Pólya-Vinogradov inequality (observe that Schur's proof [5] is elementary). Therefore

\[ T_2 \leq \left( Y + Y^{1/2}Q + Y^{1/3}Q^2 \right). \]

and with (10) and (12) this implies (11) with j = 2.

By (7) and (8) it is easily seen that T_1 can be estimated in the same way as T_2.

References