MÉMOIRES DE LA S. M. F.

STEPHEN WAINGER

On Vinogradov's estimate of trigonometric sums and the Goldbach-Vinogradov theorem

Mémoires de la S. M. F., tome 25 (1971), p. 173-181

<http://www.numdam.org/item?id=MSMF_1971__25__173_0>

Colloque Th. Nombres [1969, Bordeaux]
Bull. Soc. math. France,
Mémoire 25, 1971, p. 173 à 181

ON VINOGRADOV'S ESTIMATE OF TRIGONOMETRIE SUMS AND THE GOLDBACH-VINOGRADOV THEOREM

by

Stephen WAINGER

-:-:-:-

The purpose of this lecture is to give an expository account of the principle ideas involved in Vinogradov's estimate for

$$S_n(\theta) = \sum_{p \leq n} e^{2\pi i p\theta}$$

(in the above summation and throughout p is a prime).

However, before we begin to estimate $S_n(\theta)$, we indicate why the estimates for for $S_n(\theta)$ are so important for the proof of the Goldbach - Vinogradov theorem.

THEOREM. (Goldbach-Vinogradov). - Every sufficiently large odd integer is the sum of three primes.

We define $r(n)$ to be the number of ways to write n as a sum of three primes. Then the first step of the proof is to observe

1) $$r(n) = \int_0^1 e^{-2\pi i n\theta} [S_n(\theta)]^3 \, d\theta$$

To see 1), observe

$$[S_n(\theta)]^3 = \sum_{p_1 \leq n} \sum_{p_2 \leq n} \sum_{p_3 \leq n} e^{2\pi i (p_1+p_2+p_3)\theta}$$

so interchanging the order of summation and integration, we observe that the right hand side of 1), is

$$\sum_{p_1 \leq n} \sum_{p_2 \leq n} \sum_{p_3 \leq n} \int_0^1 e^{-2\pi i (n-(p_1+p_2+p_3))\theta} d\theta$$

The above integral is 1 if $n = p_1+p_2+p_3$ and 0 otherwise. So the triple sum above just counts the number of ways to write n as $p_1+p_2+p_3$. This is $r(n)$, and the proof of 1) is complete.

The idea of (essentially) Hardy and Littlewood is that

$$r(n) = \int_{U_n} e^{-2\pi i n\theta}[S_n(\theta)]^3\, d\theta + \int_{V_n} e^{-2\pi i n\theta}[S_n(\theta)]^3\, d\theta = M_n + E_n .$$

U_n is the union of some very small disjoint intervals which we shall describe later. What is important for us now is that for every θ in V_n,

$$\theta = \frac{a}{q} + \frac{\varepsilon}{q^2} \qquad \text{with} \quad (a,q) = 1 , \quad |\varepsilon| \le 1$$

and what is very important

2) $$(\log n)^u \le q \le \frac{n}{(\log n)^u} .$$

where u is some fixed large integer like 100.

Now one may show that

3) $$M_n > \frac{cn^2}{(\log n)^3}$$

for $n > N_o$ and n odd. (We give some indication of this at the end of our lecture).

The main difficulty in the proof is to prove

4) $$E_n = o\left(\frac{n^2}{(\log n)^3}\right) .$$

Formulas 3) and 4) of course give the Goldbach-Vinogradov Theorem.

The method of Hardy and Littlewood also dictates the general approach of estimating E_n. Namely

5) $$|E_n| \le \sup_{\theta \in V_n} |S_n(\theta)| \int_0^1 |S_n(\theta)|^2\, d\theta .$$

By Pavseval's equality the integral in 5) is $\pi(n)$ (the number of primes equal to or less than n). So using the simple estimate

$$\pi(n) \le \frac{c\,n}{\log n} , \qquad \text{we have}$$

6) $$E_n \le \frac{c\,n}{\log n} \sup_{\theta \in V_n} |S_n(\theta)| .$$

Notice then to prove 4), we need to show

7) $$|S_n(\theta)| = o\left(\frac{n}{\log^2 n}\right) .$$

Of course trivially

$$|S_n(\theta)| \leq K\frac{n}{\log n} ,$$

so we need only a very modest improvement of the trivial estimate, but this small improvement was very difficult. Vinogradov's method proves the following :

THEOREM. - *If* θ *is in* V_n , *then*

8) $$|S_n(\theta)| \leq c\, n(\log n)^{9/2} \{(\frac{1}{q} + \frac{q}{n})^{1/2} + e^{-c\sqrt{\log n}}\},$$

for $$\theta = \frac{a}{q} + \frac{\varepsilon}{q^2} \quad \text{with} \quad (a,q) = 1 , \quad 1 \leq a < q , \quad |\varepsilon| \leq 1 .$$

Recall that for $\theta \in V_n$

$$(\log n)^u \leq q \leq \frac{n}{(\log n)^u} .$$

Thus Vinogradov's estimate for $|S_n(\theta)|$ gives 7) and the Goldbach-Vinogradov Theorem.

We begin by isolating one of Vinogradov's basic ideas. Namely in a double sum

$$\sum_n \sum_m a_n b_m e^{2\pi i n m \theta} ,$$

there must be a lot of cancellation for most θ's , even if the a_n and b_m are quite irregular (like say a_n = number of divisors of n and $b_m = \mu(m)$ where μ is the Moebius function). To fix ideas consider

9) $$s = \sum_{n=1}^{q} a_n \sum_{m=1}^{q} b_m e^{2\pi i n m \frac{a}{q}} ,$$

with $(a,q) = 1$. If we neglect the cancellation due to $\exp(2i\pi nm\frac{a}{q})$, we may trivially use Schwarz's inequality to obtain

$$|s| \leq \sum_{n=1}^{q} |a_n| \sum_{m=1}^{q} |b_m| \leq q\, A\, B$$

where $$A = (\sum_{n=1}^{q} |a_n|^2)^{1/2} \quad \text{and}$$

$$B = (\sum_{m=1}^{q} |b_m|^2)^{1/2} .$$

However, this estimate may be improved Namely

IO) $$|s| \leq \sqrt{q}\, A\, B .$$

We now prove (IO) (à la Vinogradov). Using Schwarz's inequality on the outer sum in (9), we see

$$|s|^2 \le A^2 \sum_{n=1}^{q} \Big| \sum_{m=1}^{q} b_m e^{2\pi i n m \frac{a}{q}} \Big|^2$$

$$= A^2 \sum_{n=1}^{q} \sum_{m_1=1}^{q} \sum_{m_2=1}^{q} b_{m_1} \overline{b_{m_2}} e^{2\pi i n (m_1 - m_2)\frac{a}{q}}$$

$$= A^2 \sum_{m_1=1}^{q} \sum_{m_2=1}^{q} b_{m_1} \overline{b_{m_2}} \sum_{n=1}^{q} e^{2\pi i n (m_1 - m_2)\frac{a}{q}} .$$

Now since $(a,q) = 1$, one sees that the inner sum (which is the sum of a geometric progression) is q if $m_1 = m_2$ and 0 otherwise. So we have

$$|s|^2 \le q A^2 B^2 ,$$

which is (I0).

To prove 8) of course we must also consider irrational θ's but this same slightly tricky use of Schwarz's inequality can still be used to obtain cancellation in that case.

To prove 8) we must use the well known Möbius function μ. We recall

$$\mu(1) = 1$$

$$\mu(p_1,\dots,p_r) = (-1)^r \quad \text{if} \quad p_i \neq p_j$$

and

$$\mu(n) = 0 \quad \text{if} \quad p^2 \text{ divides } n$$

for some prime p.

We recall without proof the fundamental property of μ that we need :

II)
$$\sum_{d|a} \mu(d) = \begin{cases} 1 & \text{if } a = 1 \\ 0 & \text{if } a > 1 \end{cases}$$

Let $R = \prod_{p \le \sqrt{n}} p$. Then

$$S_n(\theta) = O(\sqrt{n}) + \sum_{\sqrt{n} \le p \le n} e^{2\pi i p \theta}$$

$$= O(\sqrt{n}) + \sum_{\substack{\sqrt{n} < m \le n \\ (R,m) = 1}} e^{2\pi i m \theta}$$

$$= O(\sqrt{n}) + \sum_{\substack{1 \le m \le n \\ (R,m) = 1}} e^{2\pi i m \theta}$$

$$= O(\sqrt{n}) + \sum_{1 \le m \le n} \sum_{d|(R,m)} \mu(d) e^{2\pi i m \theta}$$

by II). [(x,y) = greatest common divisor of x and y]. Interchanging the order of summation, one sees

$$S_n(\theta) = O(\sqrt{n}) + \sum_{d|R} \mu(d) \sum_{\substack{1 \le m \le n \\ d|m}} e^{2\pi i n\theta} .$$

Now setting $m = dr$, we have

I2) $$S_n(\theta) = O(\sqrt{n}) + \sum_{d|R} \mu(d) \sum_{1 \le r \le \frac{n}{d}} e^{2\pi i dr\theta} .$$

Now at first glance things may apprear very nice because the inner sum above is a geometric progession. However $\mu(d)$ is not zero for more than $O(n)$ values of d, so we are very far from the proof. What we must do is replace the double sum in I2) by another double sum with shorter ranges of summation. However, we are willing to pay a price. Namely we do not need to have such regular coefficies as $a_r = 1$ (a_r is the coefficient of $\exp 2\pi i dr\theta$) in the sum I2) if we use the idea of the proof of I0).

There are several methods of "shortening" the double sum at the price of introducing rough coefficients. We shall outline here the method given in Prachar Primzahlverteilung, where the complete details may be found.

Note : When we use the words relatively easy the reader should not try to stop and fill in the gaps as they are only really easy for people who are quite familiar with certain special methods of number theory. For complete details one should consult Prachar Primzahlverteilung.

Interchanging the order of summation in I2), we have

I3) $$S_n(\theta) = O(\sqrt{n}) + \sum_{1 \le r \le n} \sum_{\substack{d|R \\ 1 \le d \le \frac{n}{r}}} \mu(d)\, e^{2\pi i dr\theta} = L_n + F_n$$

where

$$L_n = \sum_{1 \le r \le e^{\sqrt{\log n}}} \sum_{\substack{d|R \\ 1 \le d \le \frac{n}{r}}} \mu(d)\, e^{2\pi i dr\theta}$$

and

$$F_n = O(\sqrt{n}) + \sum_{e^{\sqrt{\log n}} \le r \le n} \sum_{\substack{d|R \\ 1 \le d \le \frac{n}{r}}} \mu(d)\, e^{2\pi i dr\theta} .$$

Now F_n is relatively easy to dispose of because in thie double sum above

$$d \le \frac{n}{r} \le \frac{n}{e^{\sqrt{\log n}}} .$$

This range of summation on d is sufficiently small that one may interchange the order of summation in F_n and use the fact that the sum on r is a geometric

progression. Thus we must consider L_n. Instead of considering L_n, we consider

$$G_n = \sum_{1 \le r \le e^{\sqrt{\log n}}} \sum_{\substack{d|r \\ 1 \le d \le n/r \\ d \text{ has at least one prime factor} \ge e^{2\sqrt{\log n}}}} \mu(d)\ e^{2\pi i d r}$$

Thus G_n differs from L_n in that now we are not summing over any d's all of whose prime factors are less than $e^{2\sqrt{\log n}}$. A well known theorem of Rankin implies that there are so few such d's that

$$\sum_{1 \le r \le e^{\sqrt{\log n}}} \sum_{\substack{d|r \\ 1 \le d \le n/r \\ d \text{ has all prime factors} \le e^{2\sqrt{\log n}}}} 1 = O(n\, e^{-c\sqrt{\log n}}) .$$

Thus it suffices to consider G_n. In G_n we now write d as $p.j$ where p is a prime $e^{2\sqrt{\log n}} < p \le \sqrt{n}$ (since $p|d$, $d|R$ and $R = \prod_{p \le \sqrt{n}} p$).

Define

$$H_n = \sum_{1 \le r \le e^{\sqrt{\log n}}} \sum_{e^{2\sqrt{\log n}} \le p \le \sqrt{n}} \sum_{\substack{j|R \\ 1 \le pj \le \frac{n}{r}}} \mu(j)\ e^{2\pi i j p r \theta}$$

H_n is not equal to G_n because we have not taken into account the number of ways d may be written as jp. However, it is not too difficult to see that it suffices to consider H_n.

Note that in H_n all the sums are much shorter than n.

$$\left(j \le \frac{n}{rp} \le \frac{n}{e^{2\sqrt{\log n}}} \quad (\text{since } p > e^{2\sqrt{\log n}})\right).$$

But H_n is a triple sum instead of a double sum.

Now to change this to a double sum we set $Z = pr$, and we have

$$H_n = \sum_{e^{2\sqrt{\log n}} \le Z \le \sqrt{n}\, e^{\sqrt{\log n}}} d_1(Z) \cdot \sum_{\substack{j|R \\ 1 \le j \le \frac{n}{Z}}} \mu(j)\ e^{2\pi i j Z \theta}$$

where $d_1(Z)$ is the number of ways of writing $Z = pr$. $d_1(Z)$ is of course very irregular; but $_M(d_1(Z)) \le d(Z)$ (the number of divisors of Z), and very good estimates for $\sum_{Z=1}^{M} |d(Z)|^2$ are known

$$\left(\sum_{Z=1}^{M} |d(Z)|^2 = O\{M (\log M)^3\}\right) .$$

Notice in the last expression for H_n both sums are much shorter than n

($j \le \frac{n}{e^{2\sqrt{\log n}}}$ since $j \le n/Z$ and $Z \ge e^{2\sqrt{\log n}}$) . Thus we have achieved our objective : we have a double sum with shortened ranges of summation and of course have paid the stated price. If we now break the Z-sum in H_n into a most $\log n$ dyadic blocs and use Shwarz's inequality as in the proof of IO) , we may obtain 8) .

At this point you basically understand the idea of the proof of 4). However, we shall proceed a little farther for completeness. (Note that the trivial estimate for H_n is

$$|H_n| \le n \sum_{Z \le \sqrt{n}} e^{\sqrt{\log n}}\, d(Z) / Z$$

and this last quantity is $> n \log n$).

We divide H_n into dyadic blocks $K(u)$ where

$$K(u) = \sum_{u \le Z \le u'} d_1(Z) \cdot \sum_{\substack{j \mid R \\ 1 \le j \le \frac{n}{Z}}} (j)\, e^{2\pi i j Z\theta}$$

with $u' \le 2u$. We must consider at most $\log n$ such blocks $K(u)$. (A factor of $\log n$ at this point makes little difference since one may just choose u a little larger in 5) to obtain 7)).

Then

$$|K(u)|^2 \le \sum_{u < Z < u'} |d(Z)|^2 \sum_{u < Z < u'} \Big| \sum_{\substack{j \mid R \\ 1 \le j \le \frac{n}{Z}}} \mu(j)\ e^{2\pi i j Z\theta} \Big|^2$$

$$\le c\, u(\log u)^3 \sum_{u < Z < u'} \sum_{\substack{j_1 \mid R \\ 1 \le j_1 \le \frac{n}{Z}}} \sum_{\substack{j_2 \mid R \\ 1 \le j_2 \le \frac{n}{Z}}} \cdot\ \mu(j_1)\ \mu(j_2)\ e^{2\pi i (j_1 - j_2) Z\theta}$$

$$\le c\, u(\log u)^3 \sum_{1 \le j_1 \le \frac{n}{u}} \sum_{1 \le j_2 \le \frac{n}{u}} \Big| \sum_{u < Z < u'} e^{2\pi i Z (j_1 - j_2)\theta} \Big|$$

$$\le c\, u(\log u)^3 \sum_{1 \le j_1 \le \frac{n}{u}} \sum_{1 \le j_2 \le \frac{n}{u}} \min\Big(u, \frac{1}{|\sin (j_1 - j_2)\theta|}\Big).$$

If one now observes that for any integer ℓ , the fractional parts of the numbers $m\theta$

$$\ell q - \frac{q}{2} < m < \ell q + \frac{q}{2}$$

are of the form

$$\frac{s}{q} + \lambda_\ell + \frac{t}{2q}$$

with λ_ℓ arbitrary and $|\varepsilon| \le 1$ and with each s arising from only one value of m, the completion of the proof of 8) is then a matter of arithmetic.

To make the above observation note that

$$m\theta = \ell q\theta + (m-\ell q)\theta$$
$$= \Lambda_\ell + j\theta \qquad \text{with} \quad -\frac{q}{2} < j < \frac{q}{2} .$$

Hence we must just observe that

$$j_1 \frac{a}{q} \not\equiv j_2 \frac{a}{q} \pmod 1 \qquad \text{for} \quad j_1 \neq j_2 .$$

If $j_1 a - j_2 a = q s$ for s an integer, the $q \mid a(j_1-j_2)$. But $(a,q) = 1$ and $|j_1-j_2| < q$ so this is impossible.

As our main objective was to indicate the proof of 8) (and hence 4), we shall give only a short summary of 3).

First we shall give the definition of U_n

$$U_n = \bigcup_{\substack{q \le (\log n)^u \\ (a,q)=1 \ , \ 0<a<q}} J_{a,q}$$

$$J_{a,q} = \{\theta \mid |\theta - \frac{a}{q}| \le \frac{(\log n)^u}{n}\}$$

One sees easily that the intervals $J_{a,q}$ are disjoint for n large.

Then $V_n = [0,1] - U_n$. The theorem of Dirichilet on approximations of irrationals by rationals implies that each $\theta \in V_n$ has the form

$$\theta = \frac{a}{q} + \frac{\varepsilon}{q}\frac{(\log n)^u}{n} = \frac{a}{q} + \frac{\varepsilon'}{q^2}$$

with $|\varepsilon'| \le |\varepsilon| \le 1$, $(\log n)^u \le q \le \dfrac{n}{(\log n)^u}$, and $(a,q) = 1$ as required above.

To prove 3) we need information on $\pi(n,q;a)$ = the number of primes less than or equal to n which are congruent to a modulo g.

The main tool for the estimate 3) is the following.

THEOREM. If $(a,q) = 1$

14) $$\pi(n,q,a) = \frac{1}{\phi(q)} \int_1^n \frac{dt}{\log t} + O(n\, e^{-c\sqrt{\log n}})$$

uniformly for $1 \le q \le (\log n)^u$, where $\phi(q)$ is the number of integers relatively prime to q. Here and in the following $c > 0$.

We shall not go into the proof of this difficult theorem [See Prachar]. (The original proof of the Goldbach-Vinogradov Theorem had much extra complication because the uniformity q was not as large as in the above theorem).

The first step in the above proof is the following.

LEMMA. - For $0 \le a < q$, $(a,q) = 1$, and $1 \le q \le (\log n)^{\mu}$,

I5) $$S_n(\frac{a}{q}) = \frac{\mu(q)}{\phi(q)} \int_1^n \frac{dt}{\log t} + O(n\, e^{-c\sqrt{\log n}}).$$

To prove I5) observe

$$S_n(\frac{a}{q}) = \sum_{p \le n} e^{2\pi i p \frac{a}{q}} = \sum_{p \le n} e^{2\pi i p \frac{a}{q}} + \sum_{\substack{p \mid q \\ p \le n}} e^{2\pi i p \frac{a}{q}}$$

$$= \sum_{\substack{m=1 \\ (m,q)=1}}^{q} e^{2\pi i \frac{ma}{q}}\ \pi(n,q,m) + O(\log q) .$$

Now one finishes the proof of I5) essentially by using I4) . The next step is the following.

LEMMA. - Let $q \le (\log n)^{\mu}$, $(a,q) = 1$, $0 \le a < q$ and $0 \le |\beta| \le \frac{1}{2}$.
Then

I6) $$S_n(\frac{a}{q} + \beta) = \frac{\mu(q)}{\phi(q)} \sum_{1 \le j \le M} \frac{e^{2\pi i j \beta}}{\log j} + O\{N(N|\beta| + 1)e^{-c\sqrt{\log n}}\}$$

To prove I6) observe

$$S_n(\frac{a}{q}+\beta) = \sum_{j=1}^{n} e^{2\pi i j \beta}\, [S_n(\frac{a}{q}) - S_{n-1}(\frac{a}{q})].$$

One then sums by parts, uses I5) and completes the proof by another summation by parts.

The remainder of the proof of 3) is now relatively straightforward. One needs to substitute I6) into the integral for U_n and keep tract of the error terms.

Supported by an N.S.F. grant at Faculté de Sciences at Orsay.

-:-:-:-

University of Wisconsin
Department of Mathematics
Madison
Wisconsin (U.S.A.)