THE BOMBIERI–VINOGRADOV THEOREM

R. C. Vaughan

1. The Main Theorem

The Bombieri-A. I. Vinogradov Theorem is concerned with the distribution of primes into arithmetic progressions. Define the von Mangoldt function by

$$\Lambda(n) = \begin{cases} \log p & \text{when } n = p^k \text{ for some } p \text{ and } k \geq 1, \\ 0 & \text{otherwise,} \end{cases}$$

$$\psi(x;q,a) = \sum_{\substack{n \leq x \\ n \equiv a \pmod q}} \Lambda(n)$$

$$\vartheta(x;q,a) = \sum_{\substack{p \leq x \\ p \equiv a \pmod q}} \log p$$

$$\pi(x;q;a) = \sum_{\substack{p \leq x \\ p \equiv \pmod q}} 1$$

The theorems stated here can be restated with $\psi(x;q;a)$ replaced by $\vartheta(x;,1,a)$ or $\pi(x;q;a)$. Note that

$$\vartheta(x;q,a) = \psi(x;q,a) + O(x^{\frac{1}{2}})$$

$$\pi(x;q,a) = \frac{\vartheta(x;q,a)}{\log x} + \int_2^x \frac{\vartheta(u;q,a)}{u \log^2 u} du$$

I prefer Λ as it arises naturally as the coefficient in the Dirichlet series for the logarithmic derivative of

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \ \textit{viz.} \ -\frac{\zeta'}{\zeta}(s) = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}$$

when $\Re s > 1$.

The best general estimate for a pair q, a, uniform in q.

Siegel[1935]–Walfisz[1936]. *There is a $c_1 > 0$ so that for any $A > 0$, when $q \le (\log x)^A$ and $(a, q) = 1$,*

$$\psi(x; q, a) = \frac{x}{\phi(q)} + O_A \Big(\exp\big(- c_1 \sqrt{\log x}\big)\Big)$$

Let χ denote a Dirichlet character modulo q and put

$$\psi(x; \chi) = \sum_{n \le x} \chi(n) \Lambda(n)$$

$$\psi(x; q, a) = \frac{1}{\phi(q)} \sum_{\chi \pmod q} \overline{\chi}(a) \psi(x; \chi)$$

Siegel–Walfisz variant. *There is a $c_1 > 0$ so that for any $A > 0$, when $q \le (\log x)^A$ and χ has modulus q,*

$$\psi(x; \chi) - \delta(\chi) x \ll_A x \exp\big(- c_1 \sqrt{\log x}\big)$$

where $\delta(\chi) = 1$ when $\chi = \chi_0$ and $= 0$ otherwise.

Good references for these are Davenport [2000] or Estermann [1952] or Montgomery and Vaughan [2006].

When $\Re s > 1$ we define

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}.$$

This has an analytic continuation to $\mathbb{C}$, and is entire except when χ is principal, in which case it is analytic except at 1 where it has a simple pole with residue $\frac{\phi(q)}{q}$. Indeed,

$$L(s,\chi_0) = \zeta(s) \prod_{p|q} \left(1 - \frac{1}{p^s}\right).$$

The Generalised Riemann Hypothesis (GRH) is the statement that $L(s,\chi) \neq 0$ when $\Re s > \frac{1}{2}$.

If GRH holds for $L(s,\chi)$, then (Titchmarsh [1930])

$$\psi(x;\chi) - \delta(\chi)x \ll x^{\frac{1}{2}} (\log qx)^2,$$

and so GRH for all χ modulo q implies that

$$\psi(x;q,a) = \frac{x}{\phi(q)} + O\big(x^{\frac{1}{2}} (\log x)^2\big)$$

uniformly in q and x.

$$\text{GRH: } \psi(x;q,a) = \frac{x}{\phi(q)} + O\big(x^{\frac{1}{2}}(\log x)^2\big)$$

But without any unproved hypothesis:

Bombieri [1965]–Vinogradov [1965,1966]. *For any $A > 0$,*

$$\sum_{q \le Q} \max_{(a,q)=1} \sup_{y \le x} \left| \psi(y;q,a) - \frac{y}{\phi(q)} \right| \ll_A x(\log x)^{-A} + x^{1/2} Q (\log xQ)^6$$

We see that it is practically as good, when we average over q, as having GRH for all χ to all moduli $q \le x^{1/2}(\log x)^{6-A}$. Consequently this theorem has many applications.

By the way, the crude estimate $(x/q+1)\log x$ for each term in the sum gives the trivial bound

$$x(\log xQ)^2 + Q \log x$$

which is better than the theorem when $Q > x^{\frac{1}{2}}$, so we can suppose in any proof that

$$Q \le x^{\frac{1}{2}}.$$

$$\sum_{q \le Q} \max_{(a,q)=1} \sup_{y \le x} \left| \psi(y;q,a) - \frac{y}{\phi(q)} \right| \\ \ll_A x(\log x)^{-A} + x^{1/2} Q (\log xQ)^6$$

All proofs of the above start off the same way. By

$$\psi(y;q,a) = \frac{1}{\phi(q)} \sum_{\chi \pmod q} \overline{\chi}(a) \psi(y;\chi),$$

we have

$$\left| \psi(y;q,a) - \frac{y}{\phi(q)} \right| \le \frac{1}{\phi(q)} \sum_{\chi \pmod q} |\psi(y;\chi) - \delta(\chi) y|$$

and so it suffices to bound

$$\sum_{q \le Q} \frac{1}{\phi(q)} \sum_{\chi \pmod q} \sup_{y \le x} |\psi(y;\chi) - \delta(\chi) y|$$

This probably already throws away some likely cancellation in the summation over χ, which one will almost certainly need to make use of for any improvements.

$$\psi(y;\chi) = \sum_{n\le y} \chi(n)\Lambda(n)$$

When χ is induced by the primitive character χ^*, so that the conductor q^* divides q we have

$$\psi(y;\chi) = \psi(y;\chi^*) + O\left(\sum_{p|q, p\nmid q^*} (\log p) \sum_{k\le (\log y)/\log p} 1 \right).$$

The error term here is $\ll (\log q)\log y$ and so

$$\begin{aligned} &\sum_{q\le Q} \frac{1}{\phi(q)} \sum_{\chi \pmod q} \sup_{y\le x} |\psi(y;\chi) - \delta(\chi)y| \\ &= \sum_{q\le Q} \frac{1}{\phi(q)} \sum_{q^*|q} \ \sideset{}{^*}\sum_{\chi^* \pmod{q^*}} \sup_{y\le x} |\psi(y;\chi^*) - \delta(\chi^*)y| \\ &\qquad + O\left(Q(\log Q)(\log x)\right) \end{aligned}$$

where $\sum^*$ indicates that the sum is restricted to primitive characters. The error term here is more than acceptable, and on interchanging the order of summation and replacing q by q^*r, the main term becomes

$$\sum_{q^*\le Q} \sum_{r\le Q/q^*} \frac{1}{\phi(q^*r)} \ \sideset{}{^*}\sum_{\chi \pmod{q^*}} \sup_{y\le x} |\psi(y;\chi) - \delta(\chi)y|$$

$$\sum_{q^*\le Q}\sum_{r\le Q/q^*}\frac{1}{\phi(q^*r)}\quad\sideset{}{^*}\sum_{\chi \pmod{q^*}}\sup_{y\le x}|\psi(y;\chi)-\delta(\chi)y| \tag{9}$$

Now

$$\frac{1}{\phi(q^*r)}\le\frac{1}{\phi(q^*)\phi(r)}$$

and

$$\sum_{r\le Q}\frac{1}{\phi(r)}\ll\log 2Q.$$

To see this put $1/\phi(r)=\frac{1}{r}\sum_{m|q}\frac{\mu(m)}{\phi(m)}$, and let $r=mn$.

$$\sum_{r\le Q}\frac{1}{\phi(r)}=\sum_{m\le Q}\frac{\mu(m)}{m^2}\sum_{n\le Q/m}\frac{1}{n}.$$

Hence (9) (for simplicity we write q for q^*) is

$$\ll\sum_{q\le Q}\frac{\log 2Q}{\phi(q)}\quad\sideset{}{^*}\sum_{\chi \pmod{q}}\sup_{y\le x}\left|\psi(y;\chi)-\delta(\chi)y\right|.$$

We have reduced Bombieri–Vinogradov to bounding

$$\sum_{q\le Q}\frac{\log 2Q}{\phi(q)}\sideset{}{^*}\sum_{\chi \pmod q}\sup_{y\le x}|\psi(y;\chi)-\delta(\chi)y|\,.$$

Let $R=\log^{6+A}x$. We use the Siegel–Walfisz variant on

$$\sum_{q\le R}\frac{\log 2Q}{\phi(q)}\sideset{}{^*}\sum_{\chi \pmod q}\sup_{y\le x}|\psi(y;\chi)-\delta(\chi)y|$$

We can suppose $x>x_0(A)$. Let χ have modulus $q\le R$. We distinguish two cases. If $y\le\sqrt{x}$, then $\psi(y;\chi)-\delta(\chi)y\ll\sqrt{x}$. If $\sqrt{x}\le y\le x$, then by the Siegel-Walfsiz variant, possibly with a slightly large value of A ,

$$\psi(y;\chi)-\delta(\chi)y\ll x\exp\big(-c_2\sqrt{\log x}\big)$$

where c_2 is a positive constant. Hence

$$\begin{aligned}&\sum_{q\le R}\frac{\log 2Q}{\phi(q)}\sideset{}{^*}\sum_{\chi \pmod q}\sup_{y\le x}|\psi(y;\chi)-\delta(\chi)y|\\&\qquad\ll_A x(\log x)^{-A}\end{aligned}$$

Everything so far is classical and could have been done in 1935.

$$R = \log^{6+A} x$$

It remains to deal with

$$\sum_{R<q\le Q} \frac{\log 2Q}{\phi(q)} \sideset{}{^*}\sum_{\chi \pmod q} \sup_{y\le x} |\psi(y;\chi) - \delta(\chi)y|$$

By definitiion $\delta(\chi) = 0$ for primitive characters with conductor $q > 1$. Thus the above becomes

$$\sum_{R<q\le Q} \frac{\log 2Q}{\phi(q)} \sideset{}{^*}\sum_{\chi \pmod q} \sup_{y\le x} |\psi(y;\chi)| \,. \tag{10}$$

The essential extra ingredient is the following

Basic Mean Value Theorem. *Let*

$$T(x,Q) = \sum_{q\le Q} \frac{q}{\phi(q)} \sideset{}{^*}\sum_{\chi \pmod q} \sup_{y\le x} |\psi(y;\chi)|$$

where $\sum^*$ *indicates that the sum is over primitive characters modulo* q*, and suppose that* $Q \ge 1$*,* $x \ge 2$*. Then*

$$T(x,Q) \ll \left(x + x^{5/6}Q + x^{1/2}Q^2\right)(\log xQ)^5.$$

♠♠♠♠♠

$$\sum_{R<q\le Q} \frac{\log 2Q}{\phi(q)} \sideset{}{^*}\sum_{\chi \pmod q} \sup_{y\le x} |\psi(y;\chi)| \tag{10}$$

$$T(x,Q) = \sum_{q\le Q} \frac{q}{\phi(q)} \sideset{}{^*}\sum_{\chi \pmod q} \sup_{y\le x} |\psi(y;\chi)|$$

$$T(x,Q) \ll \left(x + x^{5/6}Q + x^{1/2}Q^2\right) (\log xQ)^5.$$

The desired result follows by partial summation. Let

$$f(q) = \frac{\log 2Q}{\phi(q)} \sideset{}{^*}\sum_{\chi \pmod q} \sup_{y\le x} |\psi(y;\chi)|.$$

Then (10) is

$$\begin{aligned}
\sum_{R<q\le Q} f(q) &= \sum_{R<q\le Q} qf(q) \left(\frac{1}{Q} + \int_q^Q \frac{dt}{t^2}\right) \\
&= Q^{-1} \sum_{R<q\le Q} qf(q) + \int_R^Q \sum_{R<q\le t} qf(q) \frac{dt}{t^2} \\
&\le (\log 2Q) \left(Q^{-1}T(x,Q) + \int_R^Q T(x,t)\frac{dt}{t^2}\right)
\end{aligned}$$

♠♠♠♠♠

$$\sum_{R<q\le Q} \frac{\log 2Q}{\phi(q)} \sum_{\chi \pmod q}^{*} \sup_{y\le x} |\psi(y;\chi)| \tag{10}$$

$$\le (\log 2Q)\left(Q^{-1}T(x,Q) + \int_R^Q T(x,t)\frac{dt}{t^2}\right)$$

The Basic Mean Value Theorem states

$$T(x,Q) \ll \left(x + x^{5/6}Q + x^{1/2}Q^2\right)(\log xQ)^5$$

This gives (remember we can suppose $Q \le x^{\frac{1}{2}}$)

$$\ll Q^{-1}\left(x + x^{5/6}Q + x^{1/2}Q^2\right)(\log x)^6$$
$$+ \int_R^Q t^{-2}\left(x + x^{5/6}t + x^{1/2}t^2\right)(\log x)^6 dt$$
$$\ll \left(xR^{-1} + x^{5/6}\log(2Q/R) + x^{1/2}Q\right)(\log x)^6.$$

We recall our choice $R = (\log x)^{6+A}$ to conclude that

$$(10) \ll x(\log x)^{-A} + x^{1/2}Q(\log x)^6$$

as required.

2. Background to the Basic Mean Value Theorem: The Large Sieve

The key new ingredient which gave rise to the BMVT was the large sieve, invented by Linnik [1941,1942] in work on the least quadratic non–residue $n(p)$ modulo a prime p. He was able to show that for any fixed positive number δ there are at most

$$\ll \log\log x$$

primes $p \leq x$ such that $n(p) > p^\delta$.

To give some idea of the background and explain what is otherwise a rather obscure terminology, consider a set $\mathcal{A}$ of integers in $[1, N]$ of cardinality Z, and define

$$Z(p,a) = \operatorname{card}\{n \in \mathcal{A} : n \equiv a \pmod p\}.$$

Now look at

$$V(p) = \sum_{a=1}^{p} \left| Z(p,a) - \frac{Z}{p} \right|^2$$

$$Z(p,a) = \mathrm{card}\{n \in \mathcal{A} : n \equiv a \pmod{p}\}.$$

$$V(p) = \sum_{a=1}^{p} \left| Z(p,a) - \frac{Z}{p} \right|^2$$

Here Z/p is the "expected" number of elements counted by $Z(p,a)$. Suppose further that for each $p \le Q$ there are $\rho(p)$ residue classes modulo p that contain no element of $\mathcal{A}$. In other words we think of $\mathcal{A}$ as arising from sifting out $\rho(p)$ residue classes from the integers in $[1, N]$ for each prime p. Then $Z(p,a) = 0$ for $\rho(p)$ values of a, and so

$$Z^2 p^{-2} \rho(p) \le V(p)$$

and hence

$$Z^2 \sum_{p \le Q} \frac{\rho(p)}{p} \le \sum_{p \le Q} pV(p).$$

Thus any non-trivial upper bound for the right hand side is likely to give non-trivial information about the size Z of the sifted set $\mathcal{A}$. Moreover there is no restriction on the size of $\rho(p)$. In particular it can grow with p, so the number of sifted classes can be large!

$$Z^2 \sum_{p \le Q} \frac{\rho(p)}{p} \le \sum_{p \le Q} pV(p).$$

By the orthogonality of the additive characters mod p,

$$Z(p,a) = \frac{1}{p} \sum_{b=1}^{p} e(-ab/p) \sum_{n \in \mathcal{A}} e(bn/p),$$

$$Z(p,a) - \frac{Z}{p} = \frac{1}{p} \sum_{b=1}^{p-1} e(-ab/p) \sum_{n \in \mathcal{A}} e(bn/p).$$

Hence, by the local variant of Parseval's identity

$$pV(p) = p \sum_{a=1}^{p} \left| Z(p,a) - \frac{Z}{p} \right|^2 = \sum_{b=1}^{b-1} \left| \sum_{n \in \mathcal{A}} e(bn/p) \right|^2$$

Therefore

$$Z^2 \sum_{p \le Q} \frac{\rho(p)}{p} \le \sum_{p \le Q} \sum_{b=1}^{b-1} \left| \sum_{n \in \mathcal{A}} e(bn/p) \right|^2$$

$$Z^2 \sum_{p \le Q} \frac{\rho(p)}{p} \le \sum_{p \le Q} \sum_{b=1}^{b-1} \left| \sum_{n \in \mathcal{A}} e(bn/p) \right|^2$$

Let

$$S(\alpha) = \sum_{n=M+1}^{M+N} a_n e(n\alpha),$$

and suppose

$$\sum_{q \le Q} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \left| S\left(\frac{a}{q}\right) \right|^2 \le \lambda(N, Q) \sum_{n=M+1}^{M+N} |a_n|^2$$

holds for any complex numbers a_n. This has become "The Large Sieve". More generally one can ask for values of $\lambda_0(N, \delta)$ such that whenever $x_1, \dots, x_R$ are R real numbers with $\|x_r - x_s\| \ge \delta$ whenever $r \ne s$ we have

$$\sum_{r=1}^{R} |S(x_r)|^2 \le \lambda_0(N, \delta) \sum_{n=M+1}^{M+N} |a_n|^2$$

for any complex numbers a_n. Such inequalities are also now called "The Large Sieve".

If $(a, q) = (b, r) = 1$, $q \leq Q$, $r \leq Q$ and $a/q \neq b/r$, then $Q^{-2} \leq 1/(qr) \leq |a/q - b/r|$ so the Farey fractions of order Q are at least $1/Q^2$ modulo 1. Thus one can take

$$\lambda(N, Q) = \lambda_0(N, Q^{-2}).$$

The first modern version of the large sieve is Roth [1965],

$$\lambda(N, Q) \ll N + Q^2 \log Q.$$

Bombieri [1965] then obtained

$$\lambda(N, Q) = N + CQ^2,$$

Gallagher [1967] gave a short proof that $\lambda(N, Q) = \pi N + Q^2$ is permissible, and a number of papers improved the constants. Finally Montgomery and Vaughan [1973], with a wrinkle by Paul Cohen, and Selberg [1991] proved

$$\lambda_0(N, \delta) = N - 1 + \delta^{-1}$$

It had already been shown by Bombieri and Davenport [1968] that this is best possible. For an overall account of this work see the survey article by Montgomery [1978].

$$Z^2 \sum_{p \le Q} \frac{\rho(p)}{p} \le \sum_{p \le Q} \sum_{b=1}^{b-1} \left| \sum_{n \in \mathcal{A}} e(bn/p) \right|^2$$

$$\sum_{q \le Q} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \left| \sum_{n=M+1}^{M+N} a_n e(an/q) \right|^2 \le \lambda(N,Q) \sum_{n=M+1}^{M+N} |a_n|^2$$

$$\lambda(N,Q) \ll N + Q^2$$

Thus $Z^2 \sum_{p \le Q} \frac{\rho(p)}{p} \le \lambda(N,Q)Z$ and so

$$Z \ll \frac{N+Q^2}{\sum_{p \le Q} \frac{\rho(p)}{p}}.$$

Example. Suppose we remove every quadratic non-residue modulo $p \le Q$. The squares remain, so $Z \gg N^{\frac{1}{2}}$. But when $p > 2$, $\rho(p) = (p-1)/2$, so

$$\sum_{p \le Q} \frac{\rho(p)}{p} \gg \sum_{3 \le p \le Q} 1 \gg Q/\log Q.$$

Hence if we take $Q = N^{\frac{1}{2}}$, then we find that $Z \ll N^{\frac{1}{2}} \log N$ which is not too bad really.

For the application to BMVT we need only the simplest bound. First we give a lemma which in fact is a statement in linear algebra. It says that if $\mathcal{M}$ is an $N \times R$ matrix, then the Hermitian matrices $\mathcal{M}\mathcal{M}^*$ and $\mathcal{M}^*\mathcal{M}$, where here (and only here) the * denotes the complex conjugate transpose, have the same largest eigenvalue.

Duality Lemma. *Suppose that c_{nr}, $n = 1, \ldots, N$,$r = 1, \ldots, R$ are complex numbers and λ is a real number such that for all complex numbers z_r we have*

$$\sum_{n=1}^{N} \left| \sum_{r=1}^{R} c_{nr} z_r \right|^2 \leq \lambda \sum_{r=1}^{R} |z_r|^2.$$

Then, for all complex numbers w_n,

$$\sum_{r=1}^{R} \left| \sum_{n=1}^{N} c_{nr} w_n \right|^2 \leq \lambda \sum_{n=1}^{N} |w_n|^2$$

Duality Lemma. *Suppose that c_{nr}, $n = 1, \dots, N$,$r = 1, \dots, R$ are complex numbers and λ is a real number such that for all complex numbers z_r we have*

$$\sum_{n=1}^{N} \left| \sum_{r=1}^{R} c_{nr} z_r \right|^2 \le \lambda \sum_{r=1}^{R} |z_r|^2.$$

Then, for all complex numbers w_n,

$$\sum_{r=1}^{R} \left| \sum_{n=1}^{N} c_{nr} w_n \right|^2 \le \lambda \sum_{n=1}^{N} |w_n|^2$$

Proof. By the Cauchy-Schwarz inequality applied to

$$LHS = \sum_{m=1}^{N} w_m \sum_{r=1}^{R} c_{mr} \sum_{n=1}^{N} \overline{c}_{nr} \overline{w}_n,$$

$$LHS^2 \le \sum_{m=1}^{N} |w_m|^2 \sum_{m'=1}^{N} \left| \sum_{r=1}^{R} c_{m'r} \overline{z}_r \right|^2, \; z_r = \sum_{n=1}^{N} c_{nr} w_n.$$

On hypothesis this is

$$\le \sum_{m=1}^{N} |w_m|^2 \lambda \sum_{r=1}^{R} |z_r|^2 = (LHS) \lambda \sum_{m=1}^{N} |w_m|^2.$$

Now we establish a suitable version of the large sieve.

A Large Sieve Inequality. *Suppose that* $0 < \delta \le \frac{1}{2}$ *and the* x_r, $r = 1 \ldots, R$ *satisfy* $\|x_r - x_s\| \ge \delta$ *whenever* $r \ne s$. *Then*

$$\sum_{r=1}^{R} |S(x_r)|^2 \le \lambda_0(N,\delta) \sum_{n=M+1}^{M+N} |a_n|^2$$

holds with

$$\lambda_0(N,\delta) = N + \frac{1}{\delta}\log\frac{3}{\delta}$$

and

$$\sum_{q\le Q} \sum_{\substack{a=1\\(a,q)=1}}^{q} \left|S\left(\frac{a}{q}\right)\right|^2 \le \lambda(N,Q) \sum_{n=M+1}^{M+N} |a_n|^2$$

holds with

$$\lambda(N,Q) = N + Q^2 \log 3Q^2.$$

Proof. By the Duality Lemma it suffices to bound

$$\sum_{n=M+1}^{M+N} \left|\sum_{r=1}^{R} b_r e(nx_r)\right|^2 \tag{12}$$

$$\sum_{n=M+1}^{M+N} \left| \sum_{r=1}^{R} b_r e(nx_r) \right|^2$$

$$= \sum_{r=1}^{R} \sum_{s=1}^{R} b_r \bar{b}_s \sum_{n=M+1}^{M+N} e\big(n(x_r - x_s)\big) \tag{13}$$

The diagonal terms $r = s$ contribute

$$N \sum_{r=1}^{R} |b_r|^2$$

and when $r \neq s$ the sum over n in modulus is

$$\leq \frac{1}{|\sin \pi(x_r - x_s)|} \leq \frac{1}{2\|x_r - x_s\|}$$

Hence the non-diagonal terms contribute at most

$$\sum_{r=1}^{R} \sum_{s=1, s\neq r}^{R} \frac{1}{2}(|b_r|^2 + |b_s|^2) \frac{1}{2\|x_r - x_s\|}$$

$$= \sum_{r=1}^{R} |b_r|^2 \sum_{s=1, s\neq r}^{R} \frac{1}{2\|x_r - x_s\|}.$$

To summarise so far, the dual form is bounded by

$$N \sum_{r=1}^{R} |b_r|^2 + \sum_{r=1}^{R} |b_r|^2 \sum_{s=1, s\neq r}^{R} \frac{1}{2\|x_r - x_s\|}$$

Now given an r we can add integers to the x_s with $s \neq r$ so that the the resulting x'_s lie in $[x_r - \frac{1}{2}, x_r + \frac{1}{2}]$. For convenience write $x'_r = x_r$. Now the numbers x'_s are all spaced δ apart. Moreover if x_- and x_+ are the smallest and largest values of the x'_s, then $x_- + 1 - \delta \geq x_+ \geq x_- + (R-1)\delta$. Thus $R\delta \leq 1$ and

$$\sum_{s=1, s\neq r}^{R} \frac{1}{2\|x_r - x_s\|} \leq 2 \sum_{k \leq 1/\delta} \frac{1}{2k\delta} \leq \frac{1}{\delta} \log \frac{3}{\delta}.$$

This establishes the theorem.

♠♠♠♠♠

One useful wrinkle is to insert a factor $f(n)$ in

$$\sum_{n=M+1}^{M+N} \left| \sum_{r=1}^{R} b_r e(n x_r) \right|^2 \tag{12}$$

which majorises the indicator function of $[M+1, M+N]$ and has other desirable properties. A simple example is

$$\max\left(0, 2(1 - |n - N_0 - M|/N)\right),$$

where $N_0 = \lceil N/2 \rceil$. Then on multiplying out and interchanging the order of summation the exponential sum over n becomes a Fejèr kernel and satisfies

$$\ll \min\left(N, \frac{1}{N\|x_r - x_s\|^2}\right)$$

This gives

$$\lambda_0(N, \delta) \ll N + \frac{1}{\delta}.$$

Selberg's argument which leads to an optimal $\lambda_0(N, \delta)$ is a more sophisticated variant of this idea.

Bombieri's attack on BMVT was organised as follows.

(i) He used the large sieve, i.e. a bound for averages over additive characters (exponential sums) to give bounds for averages over multiplicative characters (character sums).

(ii) He used the bounds for character sums to obtain mean value theorems for Dirichlet polynomials

$$\sum_n c_n \chi(n) n^{-s}.$$

(iii) The bounds for Dirichlet polynomials were used to obtain density estimates for zeros of L–functions

$$N(Q,T,\sigma) = \sum_{q \le Q} \sum_{\chi \pmod q}^{*} \sum_{\substack{\rho=\beta+i\gamma \\ |\gamma|\le T, \beta>\sigma}} 1$$

where ρ is used to denote a typical zero of $L(s,\chi)$.

(iv) Bombieri finally uses (iii) to bound $\psi(y;\chi)$ *via*

$$\psi(y;\chi) = \delta(\chi) y - \sum_\rho \frac{x^\rho}{\rho} + \text{smaller terms}.$$

Gallagher [1968] found a way of omitting step (iii), and Vaughan [1980] found a way of also omitting step (ii).

3. Bounds for character sums

Given a χ modulo q, the Gauss sum is defined by

$$\tau(\chi) = \sum_{a=1}^{q} \chi(a) e(a/q).$$

Lemma 1. *Suppose χ is a character modulo q and either $(n,q)=1$ or χ is primitive. Then*

$$\sum_{a=1}^{q} \chi(a) e(na/q) = \overline{\chi}(n)\tau(\chi).$$

When χ is primitive, $|\tau(\chi)| = \sqrt{q}$.

Pólya–I. M. Vinogradov inequality. *Suppose χ is non-principal modulo q. Then, uniformly in x and y,*

$$\sum_{x<n\le y} \chi(n) \ll q^{\frac{1}{2}} \log q$$

♠♠♠♠♠

Lemma 1. *If* $\chi \mod q$ *is primitive or* $(n,q)=1$*, then*

$$\sum_{a=1}^{q} \chi(a) e(na/q) = \overline{\chi}(n)\tau(\chi).$$

When χ *is primitive,* $|\tau(\chi)| = \sqrt{q}$.

Proof. The case $(n,q)=1$ is easy, so suppose $(n,q)>1$. Choose m, d so $(m,d)=1$ and $m/d = n/q$. Then

$$\sum_{a=1}^{q} \chi(a) e(an/q) = \sum_{h=1}^{d} e(hm/d) \sum_{\substack{a=1 \\ a \equiv h \pmod d}}^{q} \chi(a)$$

Suppose $d \mid q$, $d < q$. As χ is primitive, there are m, n so that $m \equiv n \pmod d$, $\chi(m) \neq \chi(n)$, $(mn,q)=1$. Choose c so $(c,q)=1$, $cm \equiv n \pmod q$. As k runs over all $k \pmod{q/d}$, the numbers $a = hc + kcd$ run over all $a \pmod q$ with $a \equiv h \pmod d$. Thus

$$S = \sum_{k=1}^{q/d} \chi(hc + kcd) = \chi(c) S.$$

Since $\chi(c) \neq 1$, it follows that the inner sum is 0. To evaluate $|\tau(\chi)|$, square the modulus of both sides of the first part, sum over n and apply orthonogonality.

♠♠♠♠♠

Pólya–I. M. Vinogradov inequality. *Suppose χ is non-principal modulo q. Then, uniformly in x and y,*

$$\sum_{x<n\le y}\chi(n) \ll q^{\frac{1}{2}}\log q$$

Proof. (Schur and Vinogradov). First suppose χ primitive. By orthogonality and Lemma 1

$$\begin{aligned}\sum_{x<n\le y}\chi(n) &= \sum_{x<n\le y}\sum_{m=1}^{q}\chi(m)\frac{1}{q}\sum_{h=1}^{q}e\big(h(m-n)/q\big)\\ &= \frac{1}{q}\sum_{h=1}^{q}\sum_{m=1}^{q}\chi(m)e(hm/q)\sum_{x<n\le y}e(-hn/q)\\ &= \frac{1}{q}\sum_{h=1}^{q-1}\overline{\chi}(h)\tau(\chi)\sum_{x<n\le y}e(-hn/q)\end{aligned}$$

as the sum over m is 0 when $h=q$. The sum over n is

$$\ll \sin^{-1}\pi h/q \ll \|h/q\|^{-1}$$

and so by the last part of Lemma 1 our sum is

$$\ll q^{-\frac{1}{2}}\sum_{h=1}^{q-1}\|h/q\|^{-1} \ll q^{\frac{1}{2}}\sum_{h\le q/2}\frac{1}{h}.$$

♠♠♠♠♠

Pólya–I. M. Vinogradov inequality. *Suppose χ is non-principal modulo q. Then, uniformly in x and y,*

$$\sum_{x<n\leq y} \chi(n) \ll q^{\frac{1}{2}} \log q$$

To deduce the imprimitive case, let χ^* be the primitive character inducing χ and let r be its conductor. Then

$$\begin{aligned}
\sum_{x<n\leq y} \chi(n) &= \sum_{\substack{x<n\leq y \\ (n,q/r)=1}} \chi^*(n) \\
&= \sum_{m|q/r} \mu(m)\chi^*(m) \sum_{x/m<l<y/m} \chi^*(l) \\
&\ll d(q/r) r^{\frac{1}{2}} \log r \ll q^{\frac{1}{2}} \log q.
\end{aligned}$$

A Large Sieve for Characters. *Suppose that*

$$S(\chi) = \sum_{n=M+1}^{M+N} a_n \chi(n).$$

Then, with $\lambda(N,Q) = N + Q^2 \log 3Q^2$*,*

$$\sum_{q \le Q} \frac{q}{\phi(q)} \sideset{}{^*}\sum_{\chi \pmod q} |S(\chi)|^2 \le \lambda(N,Q) \sum_{n=M+1}^{M+N} |a_n|^2$$

Proof. By Lemma 1, with χ replaced by $\overline{\chi}$,

$$S(\chi) = \sum_{n=M+1}^{M+N} a_n \sum_{a=1}^{q} \frac{\overline{\chi}(a)}{\tau(\overline{\chi})} e(na/q) = \sum_{a=1}^{q} \frac{\overline{\chi}(a)}{\tau(\overline{\chi})} S(a/q)$$

Hence, by the last part of Lemma 1,

$$\sideset{}{^*}\sum_{\chi \pmod q} |S(\chi)|^2 \le \frac{1}{q} \sum_{\chi \pmod q} \left| \sum_{a=1}^{q} \overline{\chi}(a) S(a/q) \right|^2$$

and by Parseval's identity this is

$$\frac{\phi(q)}{q} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \left| S\left(\frac{a}{q}\right) \right|^2 .$$

We now proceed to convert this into a form more suitable for our application. The first step is a simple application of the Cauchy-Schwarz inequality.

Lemma 2. *Suppose $a_1, \dots, a_M$, $b_1, \dots, b_N$ are complex numbers. Then*

$$\sum_{q\le Q} \frac{q}{\phi(q)} \sum_{\chi}{}^{*} \left| \sum_{m=1}^{M} \sum_{n=1}^{N} a_m b_n \chi(mn) \right| \ll$$

$$\left((M + Q^2 \log Q)(N + Q^2 \log Q) \sum_{m=1}^{M} |a_m|^2 \sum_{n=1}^{N} |b_n|^2 \right)^{\frac{1}{2}}$$

Proof. At once from our version of the large sieve for characters and the Cauchy-Schwarz inequality.

The next step is to insert a maximal condition into this. There are various ways of doing this. The one chosen here is motivated by something anyone who has seen a proof of the prime number theorem will be familiar with. This is the use of a formula of the kind

$$\sum_{n\le x} c_n = \frac{1}{2\pi i} \int_{c-i\infty}^{c+\infty} \sum_{n=1}^{\infty} \frac{c_n}{n^s} \frac{x^s}{x} ds$$

which is valid when $c > 0$, $x \notin \mathbb{N}$ and, for example, the series $\sum_n |c_n|$ converges. The effect of this formula is to replace the condition $n \le x$ by a twisting factor n^{-s}. To simplify matters we use a "real" version of this, i.e. a version on the line $c = 0$.

Lemma 3. *Suppose $a_1, \dots, a_M,\ b_1, \dots, b_N$ are complex numbers. Then*

$$\sum_{q \le Q} \frac{q}{\phi(q)} \sum_{\chi}{}^{*} \sup_{y \le x} \left| \sum_{m=1}^{M} \sum_{\substack{n=1 \\ mn \le y}}^{N} a_m b_n \chi(mn) \right| \ll \log xMN$$

$$\cdot \left((M + Q^2 \log Q)(N + Q^2 \log Q) \sum_{m=1}^{M} |a_m|^2 \sum_{n=1}^{N} |b_n|^2 \right)^{\frac{1}{2}}$$

Proof. Let $C = \int_{-\infty}^{\infty} \frac{\sin \alpha}{\alpha} d\alpha$. We only need to know that $C > 0$. Let $\gamma > 0$ and define

$$\delta(\beta) = \begin{cases} 1 & 0 \le \beta < \gamma, \\ 0 & \beta > \gamma. \end{cases}$$

Then

$$\int_{-\infty}^{\infty} e^{i\beta\alpha} \frac{\sin \gamma\alpha}{C\alpha} d\alpha = \delta(\beta)$$

since pairing α and $-\alpha$ shows theintegral is real, and

$$\cos \beta\alpha \sin \gamma\alpha = \tfrac{1}{2} \sin((\gamma + \beta)\alpha) + \tfrac{1}{2} \sin((\gamma - \beta)\alpha) \quad (14)$$

so changing variables gives 1 when $0 \le \beta < \gamma$ and 0 when $\beta > \gamma$.

$$\int_{-\infty}^{\infty} e^{i\beta\alpha}\frac{\sin\gamma\alpha}{C\alpha}d\alpha = \delta(\beta) = \begin{cases} 1 & 0 \le \beta < \gamma, \\ 0 & \beta > \gamma. \end{cases}$$

$$\cos\beta\alpha\sin\gamma\alpha = \tfrac{1}{2}\sin((\gamma+\beta)\alpha) + \tfrac{1}{2}\sin((\gamma-\beta)\alpha) \quad (14)$$

When $\lambda > 0$ and $A > 0$ integration by parts gives

$$\int_{A}^{\infty} \frac{\sin\lambda\alpha}{\alpha}d\alpha \ll \frac{1}{\lambda A}$$

and so through the relationship (14) again one has

$$\delta(\beta) = \int_{-A}^{A} e^{i\beta\alpha}\frac{\sin\gamma\alpha}{C\alpha}d\alpha + O\left(\frac{1}{A|\gamma-\beta|}\right)$$

Now we specialise $\gamma = \log\left(\lfloor y\rfloor + \frac{1}{2}\right)$, $\beta = \log mn$ so

$$\int_{-A}^{A} (mn)^{i\alpha}\frac{\sin\gamma\alpha}{C\alpha}d\alpha + O\left(\frac{1}{A\left|\log\left(\lfloor y\rfloor+\frac{1}{2}\right) - \log mn\right|}\right)$$

$$= \delta(\log mn) = \begin{cases} 1 & mn \le y, \\ 0 & mn > y \end{cases}$$

$$\int_{-A}^{A}(mn)^{i\alpha}\frac{\sin\gamma\alpha}{C\alpha}d\alpha+O\left(\frac{1}{A\left|\log\left(\lfloor y\rfloor+\frac{1}{2}\right)-\log mn\right|}\right)$$

$$=\delta(\log mn)=\begin{cases}1 & mn\le y,\\ 0 & mn>y\end{cases}$$

Now

$$\left|\log\frac{\lfloor y\rfloor+\frac{1}{2}}{mn}\right|\ge\min\left(\log\frac{\lfloor y\rfloor+\frac{1}{2}}{\lfloor y\rfloor},\log\frac{\lfloor y\rfloor+1}{\lfloor y\rfloor+\frac{1}{2}}\right)\gg\frac{1}{y}$$

Thus

$$\delta(\log mn)=\int_{-A}^{A}(mn)^{i\alpha}\frac{\sin\gamma\alpha}{C\alpha}d\alpha+O\left(\frac{y}{A}\right),$$

$$\sum_{m=1}^{M}\sum_{\substack{n=1\\ mn\le y}}^{N}a_m b_n\chi(mn)=\sum_{m=1}^{M}\sum_{n=1}^{N}a_m b_n\chi(mn)\delta(\log mn)$$

$$=\int_{-A}^{A}\sum_{m=1}^{M}\sum_{n=1}^{N}a_m m^{i\alpha}b_n n^{i\alpha}\chi(mn)\frac{\sin\gamma\alpha}{C\alpha}d\alpha$$

$$+O\left(\frac{y}{A}\sum_{m=1}^{M}\sum_{n=1}^{N}|a_m b_n|\right)$$

$$\sum_{m=1}^{M}\sum_{\substack{n=1\\ mn\le y}}^{N} a_m b_n \chi(mn) = \sum_{m=1}^{M}\sum_{n=1}^{N} a_m b_n \chi(mn)\delta(\log mn)$$

$$= \int_{-A}^{A} \sum_{m=1}^{M}\sum_{n=1}^{N} a_m m^{i\alpha} b_n n^{i\alpha} \chi(mn) \frac{\sin\gamma\alpha}{C\alpha} d\alpha$$

$$+ O\left(\frac{y}{A}\sum_{m=1}^{M}\sum_{n=1}^{N}|a_m b_n|\right).$$

The error term here is more than acceptable if we take $A = xMN$, and when $y \le x$ the integral is $\ll$

$$\int_{-A}^{A}\left|\sum_{m=1}^{M}\sum_{n=1}^{N} a_m m^{i\alpha} b_n n^{i\alpha}\chi(mn)\right| \min\left(\log x, \frac{1}{|\alpha|}\right) d\alpha$$

Now we can appeal to Lemma 2 and observe that

$$\int_{-A}^{A} \min\left(\log x, \frac{1}{|\alpha|}\right) d\alpha \ll \log xMN.$$

5. Dealing with the von Mangoldt function

General philosophy. We have good bounds for certain bilinear forms, at least on average. We want to convert sums in $\Lambda(n)$ into double sums. One simple way is *via*

$$\Lambda(n) = \sum_{lm=n} \mu(l) \log m \quad \text{so that}$$

$$\sum_{n \le x} \Lambda(n) f(n) = \sum_{l \le x} \sum_{m \le x/n} \mu(l)(\log m) f(lm).$$

$\mu(l)$ and $\log m$ are values of the variables in the form and f(lm) is the coefficient of the form. The first successful attack of this kind was by I. M. Vinogradov [1937] in his proof that every large odd number is the sum of three primes. Beginning with a step similar to above, he bounded

$$\textstyle\sum_{p \le x} e(p\alpha)$$

It was while studying Vinogradov's proof that Vaughan [1977] found a way of dealing with

$$\sum_{n \le x} \Lambda(n) e(n\alpha)$$

which was intrinsically more direct and focussed towards the available information on bilinear forms.

$$\sum_m \sum_n a_m b_n c_{mn}$$

In considering bilinear forms which might arise one has to have some idea of which ones can be dealt with. Here we should think of the c_{mn} as oscillating and potentially giving some cancellation. Typical examples are additive or multiplicative characters. It is useful to divide bilinear forms into two categories.

Type I. One of the variables is smooth, such as in

$$\sum_m \sum_n a_m c_{mn}$$

and one expects to sum over n with effect. Usually the only constraint is that the sum over m should not be too long, e.g. the m are restricted to a fairly short interval.

Type II. In these we are not lucky enough to have a smooth variable. One needs to use general bounds, such as those provided by the large sieve. To illustrate this look at Lemma 2 when $MN \asymp x$, and suppose $\sum_m |a_m|^2 \ll M$, $\sum_n |b_n|^2 \ll N$. Then Lemma 2 gives

$$\begin{aligned}
&\sum_{q\le Q} \frac{q}{\phi(q)} \sideset{}{^*}\sum_{\chi} \left| \sum_m \sum_n a_m b_n \chi(mn) \right| \\
&\ll \sqrt{(M+Q^2)(N+Q^2)MN} \\
&\ll x + xQM^{-\frac{1}{2}} + xQN^{-\frac{1}{2}} + x^{\frac{1}{2}}Q^2.
\end{aligned}$$

$$\sum_{q \le Q} \frac{q}{\phi(q)} \sum_{\chi}{}^{*} \left| \sum_{m} \sum_{n} a_m b_n \chi(mn) \right|$$
$$\ll x + xQM^{-\frac{1}{2}} + xQN^{-\frac{1}{2}} + x^{\frac{1}{2}} Q^2.$$

This is a good bound (cf BMVT) provided that M and N are both large (or equivalently M is large but not too close to x). In effect we are saying that the rectangular coefficient matrix (c_{mn}) should not be too "thin". In general a bilinear form which is not "thin" will be a "good" Type II form. It turns out that there is a way of dealing with the von Mangoldt function which gives rise solely to "good" bilinear forms of types I and II.

Lemma 4. *Suppose that $u > 0$, $v > 0$, $y \geq 2$ and $f : \mathbb{N} \to \mathbb{C}$. Then*

$$\sum_n \Lambda(n) f(n) = S_1 - S_2 - S_3 + S_4$$

where

$$S_1 = \sum_{m \leq u} \mu(m) \sum_{n \leq y/m} (\log n) f(mn)$$

$$S_2 = \sum_{m \leq uv} c_m \sum_{n \leq \frac{y}{m}} f(mn) \text{ where } c_m = \sum_{\substack{k \leq u, l \leq v \\ kl = m}} \Lambda(k)\mu(l)$$

$$S_3 = \sum_{m > u} \sum_{\substack{n > v \\ mn \leq y}} \Lambda(m) \left(\sum_{\substack{l \mid n \\ l \leq u}} \mu(l) \right) f(mn)$$

$$S_4 = \sum_{n \leq v} \Lambda(n) f(n)$$

If u and v are allowed to grow, but not too fast, then S_1 and S_2 will be good bilinear forms of type I and S_3 will be a good bilinear form of type II. Presumably the number of terms in S_4 will be relatively small so it can be bounded trivially.

Proof of Lemma 4. Consider the identity

$$
\begin{aligned}
-\frac{\zeta'}{\zeta}(s) &= G(s)(-\zeta'(s)) - F(s)G(s)\zeta(s) \\
&\quad - \left(-\frac{\zeta'}{\zeta}(s) - F(s)\right)(\zeta(s)G(s) - 1) + F(s)
\end{aligned}
$$

where

$$
F(s) = \sum_{n \leq u} \Lambda(n) n^{-s}, \quad G(s) = \sum_{n \leq v} \mu(n) n^{-s},
$$

and write this as

$$
-\frac{\zeta'}{\zeta}(s) = D_1(s) - D_2(s) - D_3(s) + D_4(s)
$$

Each of the four terms $D_j(s)$ can be written as a Dirichlet series. Let $\Lambda_j(n)$ be the coefficient of n^{-s} in $D_j(n)$. Then, by the identity theorem for Dirichlet series,

$$
\Lambda(n) = \Lambda_1(n) - \Lambda_2(n) - \Lambda_3(n) + \Lambda_4(n)
$$

$$
\begin{aligned}
-\frac{\zeta'}{\zeta}(s) &= G(s)(-\zeta'(s)) - F(s)G(s)\zeta(s) \\
&\quad - \left(-\frac{\zeta'}{\zeta}(s) - F(s)\right)(\zeta(s)G(s) - 1) + F(s) \\
F(s) &= \sum_{n \le u} \Lambda(n) n^{-s}, \quad G(s) = \sum_{n \le v} \mu(n) n^{-s} \\
&D_1(s) - D_2(s) - D_3(s) + D_4(s) \\
\Lambda(n) &= \Lambda_1(n) - \Lambda_2(n) - \Lambda_3(n) + \Lambda_4(n)
\end{aligned}
$$

Collecting together terms in $G(s)(-\zeta'(s)) = D_1(s)$ gives

$$\Lambda_1(n) = \sum_{lm=n, l \le v} \mu(l) \log m$$

Likewise

$$\Lambda_2(n) = \sum_{klm=n, k \le u, l \le v} \Lambda(k)\mu(l)$$

$$\Lambda_3(n) = \sum_{lm=n, l>u, m>v} \Lambda(l) \left(\sum_{k|m, k \le v} \mu(k) \right)$$

and $\Lambda_4(n)$ is $\Lambda(n)$ or 0 according as $n \le u$ or $n > u$. Then multiplying by $f(n)$ and summing over all n gives the lemma.

$$T(x,Q)=\sum_{q\le Q}\frac{q}{\phi(q)}\sideset{}{^*}\sum_{\chi \pmod q}\sup_{y\le x}|\psi(y;\chi)|$$

We now return to bounding T. Repeated use is made of

Lemma 3. *Suppose $a_1,\dots,a_M,\ b_1,\dots,b_N$ are complex numbers. Then*

$$\sum_{q\le Q}\frac{q}{\phi(q)}\sideset{}{^*}\sum_{\chi}\sup_{y\le x}\left|\sum_{m=1}^{M}\sum_{\substack{n=1\\ mn\le y}}^{N}a_mb_n\chi(mn)\right|\ll\log xMN$$

$$\cdot\left(M+Q^2\log Q)(N+Q^2\log Q)\sum_{m=1}^{M}|a_m|^2\sum_{n=1}^{N}|b_n|^2\right)^{\frac12}$$

It is useful to deal with some special situations first. If $Q^2>x$, then $M=a_1=1$, $N=\lfloor x\rfloor$, $b_n=\Lambda(n)$ gives

$$T(x,Q)\ll Q^2(\log Q)^2\sqrt{\sum_{n\le x}\Lambda(n)^2}\ll x^{\frac12}Q^2\log^3Qx.$$

Thus we can suppose that $Q^2\le x$.

$$T(x,Q) = \sum_{q\le Q} \frac{q}{\phi(q)} \sideset{}{^*}\sum_{\chi \pmod q} \sup_{y\le x} |\psi(y;\chi)|$$

Lemma 3. *Suppose $a_1,\dots,a_M$, $b_1,\dots,b_N$ are complex numbers. Then*

$$\sum_{q\le Q} \frac{q}{\phi(q)} \sideset{}{^*}\sum_{\chi} \sup_{y\le x} \left| \sum_{m=1}^{M} \sum_{\substack{n=1\\ mn\le y}}^{N} a_m b_n \chi(mn) \right| \ll \log xMN$$

$$\cdot \left((M+Q^2\log Q)(N+Q^2\log Q) \sum_{m=1}^{M} |a_m|^2 \sum_{n=1}^{N} |b_n|^2 \right)^{\frac{1}{2}}$$

We now assume $Q^2 \le x$. Let

$$u = v = \min\left(Q^2, x^{1/3}, xQ^{-2}\right)$$

Then $M = a_1 = 1$, $N = \lfloor u^2 \rfloor$, $b_n = \Lambda(n)$ gives

$$\sum_{q\le Q} \frac{q}{\phi(q)} \sideset{}{^*}\sum_{\chi} \sup_{y\le u^2} |\psi(y;\chi)|$$
$$\ll (u^2 Q + uQ^2)(\log x)^3 \ll (x^{2/3}Q + x^{1/3}Q^2)(\log x)^3$$

which is good enough.

Thus it suffices to bound

$$\sum_{q \le Q} \frac{q}{\phi(q)} \sum_{\chi}^{*} \sup_{u^2 < y \le x} |\psi(y;\chi)|.$$

In view of Lemma 4 with $f(n) = \chi(n)$ when $n \le y$ and $f(n) = 0$ otherwise it then suffices to bound

$$T_j = \sum_{q \le Q} \frac{q}{\phi(q)} \sum_{\chi}^{*} \sup_{u^2 < y \le x} |S_j(\chi)|$$

for $j = 1, 2, 3, 4$. The case $j = 4$ is easy since

$$S_4(\chi) = \sum_{n \le u} \chi(n)\Lambda(n) = \psi(u;\chi)$$

and $u \le u^2$, and so we can appeal to the previous result.

T_1 is also fairly easy, since log is smooth and

$$\begin{aligned}
S_1(\chi) &= \sum_{m \le u} \mu(m)\chi(m) \sum_{n \le y/m} (\log n)\chi(n) \\
&= \int_1^y \sum_{m \le \min(u, y/t)} \mu(m)\chi(m) \sum_{t < n \le y/m} \chi(n) \frac{dt}{t}
\end{aligned}$$

and when $q > 1$ the Pólya-Vinogradov inequality gives

$$\ll \int_1^y u q^{1/2} \log q \frac{dt}{t} \ll u q^{1/2} (\log q)(\log y).$$

This with the trivial bound $x(\log x)^2$ for $q = 1$ gives

$$T_1 \ll (x + uQ^{5/2})(\log xQ)^2$$

Recall $Q^2 \le x$, $u = v = \min\left(Q^2, x^{1/3}, xQ^{-2}\right)$. Examining the cases $Q \le x^{1/6}$, $x^{1/6} < Q \le x^{1/3}$ and $Q > x^{1/3}$ shows that

$$T_1 \ll (x + x^{1/2} Q^2)(\log x)^2$$

in each case.

T_3 is more complicated to deal with and typically the "worst" case.

$$S_3 = \sum_{m>u} \sum_{\substack{n>v \\ mn \le y}} \Lambda(m) \left(\sum_{\substack{l|n \\ l \le u}} \mu(l) \right) f(mn)$$

In Lemma 3 we want $MN \asymp x$ but both m and n have to range over more than $x^{\frac{1}{2}}$ values. We keep control of the overall number of pairs by splitting up the range for m dyadically. Let

$$\mathcal{M} = \left\{ 2^k \lfloor u \rfloor : k = 0, 1, ...; 2^k \lfloor u \rfloor \le x/u \right\}$$

so that if $M \in \mathcal{M}$, then $u < M \le x/u$, and

$$\text{card}\mathcal{M} \ll \log x \text{ and } S_3(\chi) \ll \sum_{M \in \mathcal{M}} |S_3(\chi; M)|$$

where

$$S_3(\chi; M) = \sum_{M < m \le 2M} \sum_{\substack{u < n \le x/M \\ mn \le y}} \Lambda(m) \left(\sum_{\substack{l|n \\ l \le u}} \mu(l) \right) \chi(mn).$$

Note that here the upper limit x/M is never smaller than y/m and will only come into play after we have used Lemma 3 to remove the condition $mn \le y$.

We have

$$T_3 \leq \sum_{M \in \mathcal{M}} T_3(M),$$

$$T_3(M) = \sum_{q \leq Q} \frac{q}{\phi(q)} \sum_{\chi}{}^{*} \sup_{u^2 < y \leq x} |S_3(\chi; M)|$$

$$S_3(\chi; M) = \sum_{M < m \leq 2M} \sum_{\substack{u < n \leq x/M \\ mn \leq y}} \left(\sum_{\substack{l \mid m \\ l \leq u}} \mu(l) \right) \Lambda(n)\chi(mn).$$

$$\sum_{q \leq Q} \frac{q}{\phi(q)} \sum_{\chi}{}^{*} \sup_{y \leq x} \left| \sum_{m=1}^{M} \sum_{\substack{n=1 \\ mn \leq y}}^{N} a_m b_n \chi(mn) \right| \ll \log xMN$$

$$\cdot \left((M + Q^2 \log Q)(N + Q^2 \log Q) \sum_{m=1}^{M} |a_m|^2 \sum_{n=1}^{N} |b_n|^2 \right)^{\frac{1}{2}}$$

Thus

$$T_3(M) \ll (\log Qx)^2 \times$$

$$\left((M + Q^2)(x/M + Q^2) \sum_{m \leq 2M} d(m)^2 \sum_{n \leq x/M} \Lambda(n)^2 \right)^{\frac{1}{2}}$$

$$T_3 \le \sum_{M \in \mathcal{M}} T_3(M),$$

$$T_3(M) \ll (\log Qx)^2 \times$$
$$\left((M+Q^2)(x/M+Q^2) \sum_{m \le 2M} d(m)^2 \sum_{n \le x/M} \Lambda(n)^2 \right)^{\frac{1}{2}}$$

Since

$$\sum_{m \le z} d(m)^2 \ll z(\log 2z)^3, \quad \sum_{n \le z} \Lambda(n)^2 \ll z \log 2z,$$

$$T_3(M) \ll (\log x)^4 \left(x + xM^{-\frac{1}{2}}Q + x^{\frac{1}{2}}M^{\frac{1}{2}}Q + x^{\frac{1}{2}}Q^2 \right)$$

The estimation of T_3 is completed by summing over the $M \in \mathcal{M}$. Recall that $u < M \le x/u$. Thus

$$T_3 \ll (\log x)^5 \left(x + xu^{-\frac{1}{2}}Q + x^{\frac{1}{2}}Q^2 \right)$$

$Q^2 \le x$, $u = v = \min\left(Q^2, x^{1/3}, xQ^{-2}\right)$. Again, separate inspection of the ranges $Q \le x^{1/6}$, $x^{1/6} < Q \le x^{1/3}$, $Q > x^{1/3}$ establishes an acceptable bound.

The final sum to consider is T_2 and for this we use a hybrid method. We have

$$S_2(\chi) = \sum_{m \leq u^2} \sum_{n \leq y/m} c_m \chi(mn),\ c_m = \sum_{\substack{k \leq u, l \leq v \\ kl=m}} \Lambda(k)\mu(l)$$

We now split this sum into two parts, so that

$$S_2(\chi) = S_2'(\chi) + S_2''(\chi)$$

where $S_2'(\chi)$ contains the terms with $m \leq u$ and $S_2''(\chi)$ the terms with $u < m \leq u^2$. The sum

$$T_2' = \sum_{q \leq Q} \frac{q}{\phi(q)} \sum_{\chi}^{*} \sup_{u^2 < y \leq x} |S_2'(\chi)|$$

is then treated *via* a direct use of the Pólya-Vinogradov inequality and in a concomitant manner to T_1 and the sum

$$T_2'' = \sum_{q \leq Q} \frac{q}{\phi(q)} \sum_{\chi}^{*} \sup_{u^2 < y \leq x} |S_2''(\chi)|$$

is treated in the same way as T_3. Note that

$$|c_m| = \left| \sum_{\substack{k \leq u, l \leq u \\ kl=m}} \mu(k)\Lambda(l) \right| \leq \sum_{l|m} \Lambda(l) = \log m.$$