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. By the way, the other Vinogradov, I. M., will also make an appearance, albeit somewhat fleeting, in this story.

Let

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

the von Mangoldt function, and define

$$\psi(x; q, a) = \sum_{\substack{n \leq x \\ n \equiv a \pmod q}} \Lambda(n) \tag{2}$$

which essentially counts the primes not exceeding x in the residue class a moduluo p with weight $\log p$. The higher powers of primes contribute, hopefully, a relatively small amount to the total, and anyway

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

where

$$\vartheta(x; q, a) = \sum_{\substack{p \leq x \\ p \equiv a \pmod q}} \log p. \tag{3}$$

All the main theorems stated here can be restated with $\psi(x; q; a)$ replaced by $\vartheta(x; 1, a)$ or

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

Note that

$$\pi(x; q, a) = \frac{\vartheta(x; q, a)}{\log x} + \int_2^x \frac{\vartheta(u; q, a)}{u \log^2 u} du \tag{4}$$

The main reason for preferring Λ is that it arises naturally as the coefficient in the Dirichlet series expansion of the logarithmic derivative of

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s},$$

viz.

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

when $\Re s > 1$.

The best general estimate we have for an individual pair q, a, which is uniform in q, is the

Siegel [1935]–Walfisz [1936] Theorem. *Suppose that $A > 0$ is a fixed real number. When $(a, q) = 1$ and $q \le (\log x)^A$ we have*

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

where c_1 is an absolute positive constant.

Let χ denote a Dirichlet character modulo q and put

$$\psi(x; \chi) = \sum_{n \le x} \chi(n)\Lambda(n). \tag{5}$$

Then, by orthogonality

$$\psi(x; q, a) = \frac{1}{\phi(q)} \sum_{\chi \pmod q} \overline{\chi}(a)\psi(x; \chi), \tag{6}$$

and clearly

$$\psi(x; \chi) = \sum_{a=1}^{q} \chi(a)\psi(x; q, a). \tag{7}$$

The proof of the above also establishes the

Siegel–Walfisz Theorem variant. *Suppose that $A > 0$ is a fixed real number. When $q \le (\log x)^A$ and χ is a Dirichlet character modulo q we have*

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

where c_1 is an absolute positive constant and $\delta(\chi)$ is 1 or 0 according as χ is principal or non-principal.

Good references for these two results 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 we know (Titchmarsh [1930]) that

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

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 for all q. We can compare this with

Bombieri's version of the Bombieri [1965]-A.I. Vinogradov [1965,1966] theorem. *For any fixed positive number A,*

$$\sum_{q\le Q}\max_{(a,q)=1}\sup_{y\le x}\left|\psi(y;q,a)-\frac{y}{\phi(q)}\right| \ll_A x(\log x)^{-A} + x^{1/2}Q(\log xQ)^3.$$

Bombieri had a somewhat inflated logarithmic factor compared with the above, but in applications that is usually of no significance. Vinogradov had an x^ε. We see that the above is practically as good, when we average over q, as having GRH for all χ to all moduli $q \le x^{1/2}(\log x)^{4-A}$. Consequently this theorem has many applications. Also, apart from the log power there is no known way of improving the crucial term $x^{1/2}Q(\log xQ)^4$ even if one assumes GRH.

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

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

All proofs of the above start off the same way. One observes that, by (6),

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

and so it suffices to bound

$$\sum_{q\le Q}\frac{1}{\phi(q)}\sum_{\chi \pmod q}\sup_{y\le x}|\psi(y;\chi)-\delta(\chi)y| \tag{8}$$

This already throws away some likely cancellation in the summation over χ, cancellation which almost certainly any improvements on Bombieri–Vinogradov will have to make some use of. When χ is induced by the primitive character χ^*, so that the conductor q^* divides q we have

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

The error term here is $\ll(\log q)\log y$ and so (8) is

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

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

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

Now

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

and

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

[To see this write $1/\phi(q)=\frac{1}{q}\sum_{r\mid q}\frac{\mu(r)}{\phi(r)}$, and put $q=rm$. Then the sum is $\sum_{r\le Q}\mu(r)r^{-2}\sum_{m\le Q/r}\frac{1}{m}$.] Hence, on replacing q^* by q (9) is

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

Let $R=(\log x)^{6+A}$. Then, by the variant Siegel–Walfisz theorem we have

$$\sum_{q\le R}\frac{\log 2Q}{\phi(q)}\sideset{}{^*}\sum_{\chi \pmod q}\sup_{y\le x}|\psi(y;\chi)-\delta(\chi)y|\ll_A(\log x)Rx\exp\big(-c_2\sqrt{\log x}\big)$$

where c_2 is a positive constant. We can suppose that $x > x_0(A)$. Then we distinguish two cases. If $y \le \sqrt{x}$, then we get the conclusion at once. If $\sqrt{x} \le y \le x$, then the conditions of the Siegel-Walfsiz theorem are satisfied, possibly with a slightly large value of A. Hence

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

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

By definitiion $\delta(\chi) = 0$ for primitive characters with conductor $q > 1$. Thus it remains (!) to deal with the sum

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

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*

$$T(x,Q) \ll \left(x + x^{5/6}Q + x^{1/2}Q^2\right)(\log xQ)^3.$$

We remark in passing that by working harder it is possible to replace the middle term by $x^{4/5}Q$.

The desired conclusion now follows from the above by partial summation. To see this, let

$$f(q) = \frac{1}{\phi(q)} \sideset{}{^*}\sum_{\chi \pmod q} \sup_{y \le x} |\psi(y;\chi)|.$$

Then the sum in question is

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

By the Basic Mean Value Theorem this is

$$\begin{aligned}
&\ll Q^{-1}\left(x+x^{5/6}Q+x^{1/2}Q^2\right)(\log Qx)^3\\
&+\int_R^Q\left(\frac{1+\log(2Q/t)}{t^2}\right)\left(x+x^{5/6}t+x^{1/2}t^2\right)(\log Qx)^3dt\\
&\ll\left(xR^{-1}(\log Qx)^4+x^{5/6}\log^2(2Q/R)+x^{1/2}Q\right)(\log Qx)^3.
\end{aligned}$$

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

$$\ll x(\log x)^{-A}+x^{1/2}Q(\log Qx)^3$$

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. This had been 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 $p\le x$ 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

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

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

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

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!

By the orthogonality of the additive characters modulo p,

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

and so

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

Hence, by the local variant of Parseval's identity

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

Therefore, from above,

$$Z^2\sum_{p\le Q}\frac{\rho(p)}{p} \le \sum_{p\le Q}\sum_{b=1}^{b-1}\left|\sum_{n\in\mathcal{A}} e(bn/p)\right|^2. \tag{11}$$

Let

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

Then any non-trivial value for $\lambda(N,Q)$ for which

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

holds for any complex numbers a_n, is called "The Large Sieve". That such $\lambda(N,Q)$ exist is clear *via* the Cauchy-Schwarz inequality applied to $S(a/q)$. 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". By the way, $\|\alpha\|$ is the metric on $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, that is

$$\|\alpha\| = \min_{n\in\mathbb{Z}}|\alpha - n|.$$

It is useful to observe that if $(a,q)=(b,r)=1$, $q\le Q$, $r\le Q$ and $a/q\ne b/r$, then $Q^{-2}\le 1/(qr)\le |a/q-b/r|$ and so one can take

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

The first modern version of the large sieve is due to Roth [1965], who obtained

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

Bombieri [1965] then obtained

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

Gallagher [1967] gave a quite short proof that $\lambda(N,Q)=\pi N+Q^2$ is permissible, and then there was a lot of work by a number of authors improving the constants. Finally Montgomery and Vaughan [1973,1974], with an added wrinkle by Paul Cohen, and Selberg [1991] gave proofs that the bound holds with

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

and it had already been shown by Bombieri and Davenport [1968] that this is best possible even when applied to $\lambda(N,Q)$. For an overall account of this work see the survey article by Montgomery [1978].

To see the strength of these bounds one can observe that when applied to (11) we find that

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

and so, in particular

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

To give an example, suppose we remove every quadratic non-residue to every modulus $p\le Q$. Obviously the perfect squares will remain, so $Z\gg N^{\frac12}$. On the other hand, when $p>2$ we have $\rho(p)=(p-1)/2$, 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 $Q=N^{\frac12}$, then we find that

$$Z\ll N^{\frac12}\log N$$

which is not too bad really. For the most refined of the versions of the bounds of the kind (11) see Montgomery [1968] and Montgomery and Vaughan [1973]. In some sense they are the duals of the Selberg sieve as applied to an interval.

If we are not that concerned about the logarithmic power we need only the very simplest bound. To start with we state a lemma which, in fact is a statement from linear algebra. It says that if $\mathcal{M}$ is an $N\times R$ matrix, then the two Hermitian matrices $\mathcal{M}\mathcal{M}^*$ and $\mathcal{M}^*\mathcal{M}$, where here (and only here) the asterisk denotes the complex conjugate transpose, have the same largest eigenvalue. By the way, quite a number of the underlying ideas in this area are related to, or suggested by, ideas from linear algebra.

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*

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

Then

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

holds for all complex numbers w_n.

Proof. We have

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

Hence, by Cauchy's inequality,

$$LHS^2 \leq \left(\sum_{m=1}^{N} |w_m|^2 \right) \sum_{m=1}^{N} \left| \sum_{r=1}^{R} c_{mr} \overline{z}_r \right|^2$$

where

$$z_r = \sum_{n=1}^{N} c_{nr} w_n.$$

On hypothesis this does not exceed

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

By definition of z_r this is

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

By the way I. M. Vinogradov makes repeated use of the Duality Lemma in many special cases in his work on exponential sums, but always obtained directly *via* the Cauchy-Schwarz inequality and without, apparently, being aware that it was a special case of a general theorem!

Below is a very simple proof of Roth's bound for the large sieve which would serve to establish Bombieri's theorem with a slightly inflated logarithmic power. This is certainly adequate for most applications of Bombieri;s theorem.

A Large Sieve Inequality. *Suppose that* $0 < \delta \le \frac{1}{2}$ *and the* x_r, $r = 1 \dots, R$ *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

$$\sum_{n=M+1}^{M+N} \left|\sum_{r=1}^{R} b_r e(nx_r)\right|^2 = \sum_{r=1}^{R}\sum_{s=1}^{R} b_r \bar{b}_s \sum_{n=M+1}^{M+N} e\big(n(x_r - x_s)\big). \tag{12}$$

The diagonal terms $r = s$ contribute

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

and when $r \ne s$ the sum over n satisfies

$$\left|\sum_{n=M+1}^{M+N} e\big(n(x_r - x_s)\big)\right| \le \frac{1}{|\sin \pi(x_r - x_s)|} \le \frac{1}{2\|x_r - x_s\|}. \tag{13}$$

Hence the non-diagonal terms contribute at most

$$\begin{aligned}&\sum_{r=1}^{R} \sum_{s=1, s\ne r}^{R} \frac{1}{2}(|b_r|^2 + |b_s|^2)\frac{1}{2\|x_r - x_s\|}\\ =&\sum_{r=1}^{R} |b_r|^2 \sum_{s=1, s\ne r}^{R} \frac{1}{2\|x_r - x_s\|}.\end{aligned}$$

Given an r we can add integers to the x_s with $s \ne r$ 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.

In fact, in our proof of Bombieri's theorem we will assume the slightly stronger statement that

$$\sum_{r=1}^{R} |S(x_r)|^2 \leq \lambda_1(N, \delta) \sum_{n=M+1}^{M+N} |a_n|^2$$

holds with

$$\lambda_1(N, \delta) \ll N + \frac{1}{\delta}$$

and

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

holds with

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

One quite simple wrinkle which can be employed to establish this is to insert a factor $f(n)$ in (12) which majorises the characteristic function of $[M+1, M+N]$ and has other desirable properties. One 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

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

in place of (13). This gives

$$\lambda_1(N, \delta) \ll N + \frac{1}{\delta}.$$

Selberg's argument which leads to an optimal $\lambda_1(N, \delta)$ is a more sophisticated variant of this idea. For more details see my notes on the large sieve or Montgomery's expository article (Montgomery [1978]).

Bombieri's attack on the BMVT was organised as follows.

(i) He used the large sieve inequality, which is a bound for averages of sums of values of additive characters (exponential sums) to obtain bounds for averages of sums of values of Dirichlet characters (character sums).

(ii) He used the bounds for character sums to obtain mean value theorems for Dirichlet polynomials, that is sums of the kind

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

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

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

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

(iv) Bombieri finally uses the zero density estimates to bound $\psi(y;\chi)$ *via* the explicit formula

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

Gallagher [1968] found a way of omitting step (iii), i.e. of going directly from Dirichlet polynomials to $\psi(y;\chi)$, and Vaughan [1980] found a way of also omitting step (ii), i.e. of going directly from character sums to $\psi(y;\chi)$. The remainder of this exposition is based on this latter argument. Davenport [2000] also gives an account of the Bombieri–Vinogradov theorem similar to that in these notes. The first edition of Davenport [1967] has a different proof, closer to Bombieri's original.

3. Bounds for character sums

We first record some useful facts about the simplest character sums and Gauss sums. Given a character χ modulo q, we define the Gauss sum by

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

This can be thought of as an inner product between additive and multiplicative characters, and is the principal medium for translation between additive and multiplicative characters.

Lemma 1. *Suppose that 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, $|\tau(\chi)| = \sqrt{q}$.

Proof. The case $(n,q)=1$ is trivial. The case $(n,q)>1$, which we now assume, is not quite. Choose m and d so that $(m,d)=1$ and $m/d=n/q$. Then

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

We now show that the inner sum vanishes. Suppose that $d\mid q$, $d<q$. Since χ is primitive, there exist integers m and n such that $m\equiv n \pmod d$, $\chi(m)\neq\chi(n)$, $\chi(mn)\neq 0$. Choose c so that $(c,q)=1$, $cm\equiv n \pmod q$. As k runs through a complete residue system $\pmod{q/d}$, the numbers $n=hc+kcd$ run through all residues $\pmod q$ for which $n\equiv h \pmod d$. Thus the sum S in question is

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

Since $\chi(c)\neq 1$, it follows that $S=0$. To evaluate $|\tau(\chi)|$ we take the square of the modulus of both sides of the first part of the lemma, and sum over n to see that

$$\varphi(q)|\tau(\chi)|^2=\sum_{n=1}^{q}\left|\sum_{a=1}^{q}\chi(a)e(an/q)\right|^2=\sum_{a=1}^{q}\sum_{b=1}^{q}\chi(a)\overline{\chi}(b)\sum_{n=1}^{q}e((a-b)n/q).$$

The innermost sum on the right is 0 unless $a\equiv b \pmod q$, in which case it is equal to q. Thus $\varphi(q)|\tau(\chi)|^2=\varphi(q)q$, and hence $|\tau(\chi)|=\sqrt{q}$.

One use for the above is

The Pólya [1918]–I. M. Vinogradov [1918] inequality. *Suppose that χ is a non-principal character modulo q. Then*

$$\sum_{x<n\le y}\chi(n)\ll q^{\frac12}\log q$$

uniformly in x and y with $x\le y$.

Proof. (Schur [1919] and Vinogradov [1919]). We first prove this when χ is primitive. By the orthogonality of the additive characters modulo q and Lemma 1 we have

$$\begin{aligned}\sum_{x<n\le y}\chi(n)&=\sum_{x<n\le y}\sum_{m=1}^{q}\chi(m)\frac1q\sum_{h=1}^{q}e\big(h(m-n)/q\big)\\&=\frac1q\sum_{h=1}^{q}\sum_{m=1}^{q}\chi(m)e(hm/q)\sum_{x<n\le y}e(-hn/q)\\&=\frac1q\sum_{h=1}^{q-1}\overline{\chi}(h)\tau(\chi)\sum_{x<n\le y}e(-hn/q)\end{aligned}$$

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

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

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

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

and the primitive case follows.

To deduce the imprimitive case, let χ^* be the primitive character which induces χ and let r be the conductor of χ^*. Then

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

A Large Sieve for Characters. *Suppose that*

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

Then

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

holds with

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

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

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

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 that $a_1, \dots, a_M$, $b_1, \dots, b_N$ are complex numbers. Then*

$$\sum_{q \le Q} \frac{q}{\phi(q)} \sideset{}{^*}\sum_{\chi} \left| \sum_{m=1}^{M} \sum_{n=1}^{N} a_m b_n \chi(mn) \right| \ll \sqrt{(M+Q^2)(N+Q^2) \sum_{m=1}^{M} |a_m|^2 \sum_{n=1}^{N} |b_n|^2}.$$

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 that $x \ge 2$, Then on the premises of Lemma 2,*

$$\sum_{q \le Q} \frac{q}{\phi(q)} \sideset{}{^*}\sum_{\chi} \sup_{y \le x} \left| \sum_{m=1}^{M} \sum_{\substack{n=1 \\ mn \le y}}^{N} a_m b_n \chi(mn) \right| \ll (\log xMN) \sqrt{(M+Q^2)(N+Q^2) \sum_{m=1}^{M} |a_m|^2 \sum_{n=1}^{N} |b_n|^2}.$$

Proof. Let

$$C = \int_{-\infty}^{\infty} \frac{\sin \alpha}{\alpha} d\alpha.$$

We only need to know that C exists and $C > 0$ which is trivial from the observation that the integral can be written as

$$2 \sum_{n=1}^{\infty} (-1)^{n-1} \int_0^{\pi} \frac{\sin \alpha}{\pi(n-1) + \alpha} d\alpha$$

and the terms in the series oscillate in sign and their absolute values form a decreasing sequence tending to 0, so Leibnitz' test may be applied.

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 that the integral is real, and

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

Thus changing variables gives the value 1 when $0 \le \beta < \gamma$ and 0 when $\beta > \gamma$.

By integration by parts, provided that $\lambda > 0$ and $A > 0$, one has

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

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

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

and

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

We have

$$\min_{m,n} \left|\log\left(\lfloor y \rfloor + \frac{1}{2}\right) - \log mn\right| = \min\left(\log \frac{\lfloor y \rfloor + \frac{1}{2}}{\lfloor y \rfloor}, \log \frac{\lfloor y \rfloor + 1}{\lfloor y \rfloor + \frac{1}{2}}\right) \gg \frac{1}{y}.$$

Thus

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

Hence

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

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

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

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

By Lemma 2,

$$\sum_{q \le Q} \frac{q}{\phi(q)} \sideset{}{^*}\sum_{\chi^*} \left| \sum_{m=1}^{M} \sum_{n=1}^{N} a_m m^{i\alpha} b_n n^{i\alpha} \chi(mn) \right|$$
$$\ll \sqrt{(M+Q^2)(N+Q^2) \sum_{m=1}^{M} |a_m|^2 \sum_{n=1}^{N} |b_n|^2}$$

and

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

5. Dealing with the von Mangoldt function

The general philosophy is that we have good information about various kinds of bilinear forms, at least on average. Thus we want to convert our sums involving $\Lambda(n)$ into double sums. One, possibly naive, way of doing this is *via* the formula

$$\Lambda(n) = \sum_{lm=n} \mu(l) \log m$$

so that, for example,

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

and we would think of $\mu(l)$ and $\log m$ as being values of the variables in the bilinear form and f(lm) as being the coefficient of the bilinear form. The first person to successfully attack such a problem was I. M. Vinogradov [1937] in his proof that every sufficiently large odd number is the sum of three primes. He needed to bound

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

when α does not have a good rational approximation with a relatively small denominator. His first step is not dissimilar to that mentioned above in the case of

$\Lambda(n)$. It was while examining Vinogradov's methods 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.

In considering bilinear forms

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

which might arise one has to have some idea of which ones can be sensibly 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. In these one of the variables is smooth, ideally always 1, such as

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

and it is possible to perform the summation over n with effect. Usually the only constraint is that the sum over m should not be too long, i.e. ideally we want to ensure that the m are restricted to a fairly short interval.

Type II. In these we are not lucky enough to find that one of the variables is congenial. One needs to use quite general bounds, such as those provided by the large sieve. To illustrate this let us look at the bound provided by Lemma 3. For sake of argument, lets suppose that $MN\asymp x$, and

$$\sum_m|a_m|^2\ll M,\quad \sum_n|b_n|^2\ll N.$$

Then Lemma 3 gives the bound

$$\ll\sqrt{(M+Q^2)(N+Q^2)MN}\ll x+xQM^{-\frac12}+xQN^{-\frac12}+x^{\frac12}Q^2$$

and 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".

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 $u>0$, $v>0$, $y\ge 2$ and $f:\mathbb{N}\to\mathbb{C}$. Then*

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

where

$$S_1 = \sum_{m \le u} \mu(m) \sum_{n \le y/m} (\log n) f(mn),$$
$$S_2 = \sum_{m \le uv} c_m \sum_{n \le y/m} f(mn) \text{ where } c_m = \sum_{\substack{k \le u, l \le v \\ kl = m}} \Lambda(k)\mu(l),$$
$$S_3 = \sum_{m > u} \sum_{\substack{n > v \\ mn \le y}} \left(\sum_{\substack{k \mid m \\ k > u}} \Lambda(k) \right) \mu(n) f(mn),$$
$$S_4 = \sum_{n \le v} \Lambda(n) f(n).$$

One can see that 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. Consider the identity

$$-\frac{\zeta'}{\zeta}(s) = G(s)(-\zeta'(s)) - F(s)G(s)\zeta(s) - (-\zeta'(s) - F(s)\zeta(s)) \left(G(s) - \frac{1}{\zeta(s)} \right) + F(s)$$

where

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

and write this as

$$D_1(s) - D_2(s) - D_3(s) + D_4(s)$$

Each of the $D_j(s)$ 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,

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

By inspection of each of the Dirichlet series $D_j(s)$ we can see that each S_j satisfies

$$S_j = \sum_n \Lambda_j(n) f(n).$$

6. Proof of the Basic Mean Value Theorem

We now return to the proof of the theorem, that is, we bound

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

It is useful to deal with some special situations first. If $Q^2>x$, then using Lemma 3 directly with $M=1$, $a_1=1$, $N=\lfloor x\rfloor$, $b_n=\Lambda(n)$ gives the bound

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

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

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

Then in the same way from Lemma 3, when the supremum is restricted to $y\le u^2$, we get

$$\sum_{q\le Q}\frac{q}{\phi(q)}\sideset{}{^*}\sum_{\chi}\sup_{y\le u^2}|\psi(y;\chi)|\ll(u^2Q+uQ^2)(\log x)^3\ll(x^{2/3}Q+x^{1/3}Q^2)(\log x)^3 \tag{15}$$

which is good enough. Thus it suffices to bound

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

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

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

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 (15).

The expression T_1 is also fairly easy, since log is smooth and

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

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

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

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

$$T_1 \ll (x + uQ^{5/2})(\log xQ)^2 \ll (x + x^{1/2}Q^2)(\log xQ)^2$$

on examining the different cases $Q \le x^{1/6}$, $x^{1/6} < Q \le x^{1/3}$ and $Q > x^{1/3}$.

The expression T_3 is more complicated to deal with. 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

$$\text{card}\mathcal{M} \ll \log x.$$

Then

$$S_3(\chi) \ll \sum_{M \in \mathcal{M}} |S_3(\chi; M)|$$

where

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

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 $mn \le y$.

It follows now that

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

where

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

By Lemma 3,

$$T_3(M) \ll (\log x) \sqrt{(M + Q^2)(xM^{-1} + Q^2) \sum_{m \le 2M} (\log m)^2 \sum_{n \le x/M} \mu(n)^2}.$$

Since

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

we have

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

The estimation of T_3 is completed by summing over the elements of $\mathcal{M}$. Thus

$$T_3 \ll (\log x)^3 \left(x + xu^{-1/2}Q + x^{1/2}Q^2\right).$$

Again, separate inspection of the ranges $Q \le x^{1/6}$, $x^{1/6} < Q \le x^{1/3}$, $Q > x^{1/3}$ establishes that

$$T_3 \ll (\log x)^3 \left(x + x^{5/6}Q + x^{1/2}Q^2\right).$$

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

$$S_2(\chi) = \sum_{m \le u^2} \sum_{n \le y/m} c_m \chi(mn)$$

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

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

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

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

is treated in the same way as T_3. Note that

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

Appendix on primitive characers

Suppose that $d \mid q$ and that $\chi^\star$ is a character (mod d), and set

$$\chi(n) = \begin{cases} \chi^\star(n) & (n,q) = 1; \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Then $\chi(n)$ is multiplicative and has period q, so $\chi(n)$ is a Dirichlet character (mod q). In this situation we say that $\chi^\star$ *induces* χ. If q is composed entirely of primes dividing d then $\chi(n) = \chi^\star(n)$ for all n, but if there is a prime factor of q not found in d then $\chi(n)$ does not have period d. Nevertheless, χ and $\chi^\star$ are nearly the same. Our immediate task is to determine when one character induces another.

Lemma A.1. *Let χ be a character* (mod q)*. We say that d is a* quasiperiod *of χ if $\chi(m) = \chi(n)$ whenever $m \equiv n \pmod{d}$ and $(mn, q) = 1$. The least quasiperiod of χ is a divisor of q.*

Proof. Let d be a quasiperiod of χ, and put $g = (d, q)$. We show that g is also a quasiperiod of χ. Suppose that $m \equiv n \pmod{g}$ and that $(mn, q) = 1$. Since g is a linear combination of d and q, and $m - n$ is a multiple of g, it follows that there are integers x and y such that $m - n = dx + qy$. Then $\chi(m) = \chi(m - qy) = \chi(n + dx) = \chi(n)$. Thus g is a quasiperiod of χ.

With more effort it can be shown that if d_1 and d_2 are quasiperiods of χ then (d_1, d_2) is also a quasiperiod, and hence the least quasiperiod divides all other quasiperiods, and in particular it divides q (since q is a quasiperiod of χ).

The least quasiperiod d of χ is called the *conductor* of χ. Suppose that d is the conductor of χ. If $(n, d) = 1$ then $(n + kd, d) = 1$. Also, if $(r, d) = 1$ then there exist values of $k \pmod{r}$ for which $(n + kd, r) = 1$. Hence there exist integers k for which $(n + kd, q) = 1$. For such a k put $\chi^\star(n) = \chi(n + kd)$. Although there are many such k, there is only one value of $\chi(n + kd)$ when $(n + kd, q) = 1$. We extend the definition of $\chi^\star$ by setting $\chi^\star(n) = 0$ when $(n, d) > 1$. It is readily seen that $\chi^\star$ is multiplicative and that $\chi^\star$ has period d. Thus $\chi^\star$ is a character modulo d. Moreover, if χ_0 is the principal character modulo q, then $\chi(n) = \chi^\star(n)\chi_0(n)$. Thus $\chi^\star$ induces χ. Clearly $\chi^\star$ has no quasiperiod smaller than d, for otherwise χ would have a smaller quasiperiod, contradicting the minimality of d. In addition, $\chi^\star$ is the only character (mod d) that induces χ, for if there were another, say χ_1, then for any n with $(n, d) = 1$ we would have $\chi^\star(n) = \chi^\star(n + kd) = \chi(n + kd) = \chi_1(n + kd) = \chi_1(n)$, on choosing k as above.

A character χ modulo q is said to be *primitive* when q is the least quasiperiod of χ. Such χ are not induced by any character having a smaller conductor. We summarize our discussion as follows.

Theorem A.2. *Let χ denote a Dirichlet character modulo q and let d be the conductor of χ. Then $d \mid q$, and there is a unique primitive character $\chi^\star$ modulo d that induces q.*

We now give two useful criteria for primitivity.

Theorem A.3. *Let χ be a character modulo q. Then the following are equivalent:*

(1) *χ is primitive.*
(2) *If $d \mid q$ and $d < q$ then there is a c such that $c \equiv 1 \pmod{d}$, $(c, q) = 1$, $\chi(c) \neq 1$.*
(3) *If $d \mid q$ and $d < q$, then for every integer a,*

$$\sum_{\substack{n=1 \\ n \equiv a \pmod{d}}}^{q} \chi(n) = 0.$$

Proof. (1) $\Rightarrow$ (2). Suppose that $d \mid q$, $d < q$. Since χ is primitive, there exist integers m and n such that $m \equiv n \pmod{d}$, $\chi(m) \neq \chi(n)$, $\chi(mn) \neq 0$. Choose c so that $(c, q) = 1$, $cm \equiv n \pmod{q}$. Thus we have (2).

(2) $\Rightarrow$ (3). Let c be as in (2). As k runs through a complete residue system (mod q/d), the numbers $n = ac + kcd$ run through all residues (mod q) for which $n \equiv a \pmod{d}$. Thus the sum S in question is

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

Since $\chi(c) \neq 1$, it follows that $S = 0$.

(3) $\Rightarrow$ (1). Suppose that $d \mid q$, $d < q$. Take $a = 1$ in (3). Then $\chi(1) = 1$ is one term in the sum, but the sum is 0, so there must be another term $\chi(n)$ in the sum such that $\chi(n) \neq 1$, $\chi(n) \neq 0$. But $n \equiv 1 \pmod{d}$, so d is not a quasiperiod of χ, and hence χ is primitive.

References

E. Bombieri [1965] On the large sieve, Mathematika **12**, 201–225.

E. Bombieri and H. Davenport [1968], *On the large sieve method*, Abh. Zahlentheorie Anal., pp. 9–22.

H. Davenport [1967], *Multiplicative number theory*, Markham, Chicago.

H. Davenport [2000], *Multiplicative Number Theory, third edition*, Springer-Verlag, Berlin.

T. Estermann [1952], *Introduction to modern prime number theory*, Cambridge University Press, Cambridge, Tract No. 41.

P. X. Gallagher [1967], *The large sieve*, Mathematika **14**, 14–20.

P. X. Gallagher [1968], *Bombieri's mean value theorem*, Mathematika **15**, 1–6.

Yu. V. Linnik [1941] The large sieve, C. R. (Dokl.) Acad. Sci. URSS, n. Ser. **30**, 292–294.

Yu. V. Linnik [1942] A remark on the least quadratic non-residue, C. R. (Dokl.) Acad. Sci. URSS, n. Ser. **36**, 119–120.

H. L. Montgomery [1968], *A note on the large sieve*, J. Lond. Math. Soc. **43**, 93–98.

H. L. Montgomery [1978], *The analytic principle of the large sieve*, Bull. Am. Math. Soc. **84**, 547–567.

H. L. Montgomery and R. C. Vaughan [1973], *The large sieve*, Mathematika **20**, 119–134.

H. L. Montgomery and R. C. Vaughan [1974], *Hilbert's inequality*, J. Lond. Math. Soc. (2) **8**, 73–82.

H. L. Montgomery and R. C. Vaughan [2006], *Multiplicative Number Theory. I. Classical Theory*, Cambridge University Press, Cambridge.

G. Pólya [1918], *Über die Verteilung der quadratischen Reste und Nichtreste*, Nachr. Akad. Wiss. Göttingen 1918, 21–29.

K. F. Roth [1965], *On the large sieves of Linnik and Renyi*, Mathematika **12**, 1–9.

I. Schur [1918], *Einige Bemerkungen zu der vorstehenden Arbeit des Herrn G. Pólya: Über die Verteilung der quadratischen Reste und Nichtreste*, Nachr. Akad. Wiss. Göttingen 1918, 30–36.

A. Selberg [1991], *Collected papers. Volume II*, Springer-Verlag, Berlin.

C. L. Siegel [1935], *Über die Klassenzahl quadratischer Zahlkörper*, Acta Arith. **1**, 83–86.

E. C. Titchmarsh [1930], *A divisor problem*, Rend. Circ. Mat. Palermo **54**, 414–429.

R. C. Vaughan [1977], *Sommes trigonométriques sur les nombres premiers,*, C. R. Acad. Sci. Paris, Série A **285**, 981-983.

R. C. Vaughan [1980], *An elementary method in prime number theory*, Acta Arith. **37**, 111–115.

A. I. Vinogradov [1965], *On the density hypothesis for Dirichlet L– series*, Izv. Akad. Nauk SSSR, Ser. Mat. **29**, 903–934 (1965).

A. I. Vinogradov [1966], *Corrections to the work of A.I. Vinogradov 'On the density hypothesis for Dirichlet L– series'*, Izv. Akad. Nauk SSSR, Ser. Mat. **30**, 719–729.

I. M. Vinogradov [1918], *Sur la distribution des résidus et des nonrésidus des puissances*, J. Soc. Phys. Math. Univ. Permi 1918, 18–28.

I. M. Vinogradov [1919], *Über die Verteilung der quadratischen Reste und Nichtreste*, J. Soc. Phys. Math. Univ. Permi 1919, 1–14.

I. M. Vinogradov [1937], *Some theorems concerning the theory of primes*, Recueil Math. (2) **44**, 179– 195.

A. Walfisz [1936], *Zur additiven Zahlentheorie. II*, Math. Z. **40**, 592–607.

Department of Mathematics, McAllister Building, Pennsylvania State University, University Park, PA 16802, U.S.A.

E-mail address: `rvaughan@math.psu.edu`