5. THE LARGE SIEVE

The key new ingredient which gave rise to the B-VMVT 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 $\{a_n : 1 \le n \le N\}$ of complex numbers a_n with the property that for each prime p the support of a_n lies in $g(p)$ residue classes modulo p, and we may certainly suppose that $g(p) \ge 1$ always. We can think of the a_n as being the characteristic function of a set which has had $p - g(p)$ residue classes removed for each p.

Let

$$Z(q, rh) = \text{card}\{n \in \mathcal{A} : n \equiv h \pmod q\} \text{ and } Z = Z(0,1).$$

We might hope that for each prime p the support of a_n is fairly uniformly distributed into the $g(p)$ residue classes. Let $\mathcal{R}(p)$ be the set of $g(p)$ residue classes modulo p which contain the support of the a_n and consider the "variance"

$$V(p) = \sum_{r \in \mathcal{R}(p)} \left| Z(p,r) - \frac{Z}{g(p)} \right|^2 .$$

Let

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

By the orthogonality of the additive characters modulo p, for any h modulo p

$$Z(p,h) = \frac{1}{p} \sum_{a=1}^{p} e(-ha/p) \sum_{n=1}^{N} a_n e(an/p) = \frac{1}{p} \sum_{a=1}^{p} e(-ha/p) S(a/p)$$

and (Parseval's identity)

$$\sum_{r \in \mathcal{R}(p)} |Z(p,r)|^2 = \sum_{h=1}^{p} |Z(p,h)|^2 = \frac{1}{p} \sum_{a=1}^{p} |S(a/p)|^2.$$

We also have

$$\begin{aligned} V(p) &= \sum_{r\in\mathcal{R}(p)} |Z(p,r)|^2 - 2\Re \sum_{r\in\mathcal{R}(p)} Z(p,r)\overline{Z}/g(p) + |Z|^2/g(p) \\ &= \frac{1}{p}\sum_{a=1}^{p} |S(a/p)|^2 - |Z|^2/g(p) \\ &= \frac{1}{p}\sum_{a=1}^{p-1} |S(a/p)|^2 - |Z|^2\frac{p-g(p)}{pg(p)} \end{aligned}$$

Thus

$$pV(p) + |Z|^2\frac{p-g(p)}{g(p)} = \sum_{a=1}^{p-1} |S(a/p)|^2.$$

Summing over a suitable set $\mathcal{P}$ of primes p gives

$$\sum_{p\in\mathcal{P}} pV(p) + |Z|^2 \sum_{p\in\mathcal{P}} \frac{p-g(p)}{g(p)} = \sum_{p\in\mathcal{P}}\sum_{a=1}^{p-1} |S(a/p)|^2.$$

Since the variance is non-negative, but might be small in the ideal situation of uniform distribution we conclude that

$$|Z|^2 \sum_{p\in\mathcal{P}} \frac{p-g(p)}{g(p)} \le \sum_{p\in\mathcal{P}}\sum_{a=1}^{p-1} |S(a/p)|^2.$$

Suppose we can obtain an non-trivial upper bound for the right hand side, such as

$$\sum_{p\in\mathcal{P}}\sum_{a=1}^{p-1} |S(a/p)|^2 \le Y \sum_{n=1}^{N} |a_n|^2.$$

By the way, we know from the Cauchy-Schwarz inequality that such bounds exist and indeed this sum could be written in terms of a Hermitian matrix, so Y could be taken to be its largest eigenvalue. Suppose further that a_n is the characteristic function of an interesting set. Then $Z = |Z| = \sum_{n=1} |a_n|^2$ so we have

$$|Z| \le Y / \sum_{p\in\mathcal{P}} \frac{p-g(p)}{g(p)}.$$

Suppose we also write $\omega(p) = p - g(p)$ so that $\omega(p)$ is the number of residue classes removed. Then the sum here is

$$\sum_{p\in\mathcal{P}} \frac{\omega(p)}{p-\omega(p)}.$$

Doesn't this look a bit like the sum in the Selberg sieve? It certainly looks like an upper bound sieve estimate. For ω small, say 1 or 2 it only saves a $\log\log$ so is not very good. But if $\omega = (p-1)/2$, say, then the sum will save a power of N maybe. In other words it looks interesting when we are removing a large number of residue classes.

Thus any non-trivial value for $Y(N,Q)$ for which

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

holds for any complex numbers a_n, has become called "The Large Sieve". That such $Y(N,Q)$ exist is clear *via* the Cauchy-Schwarz inequality applied to $S(a/q)$. More generally one can ask for values of $Y_0(N,\delta)$ 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

$$\sum_{r=1}^{R}|S(x_r)|^2 \le Y_0(N,\delta)\sum_{n=M+1}^{M+N}|a_n|^2$$

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

$$Y(N,Q) = Y_0(N,Q^{-2}).$$

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

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

Bombieri [1965] then obtained

$$Y(N,Q) = N + CQ^2,$$

Gallagher [1967] gave a quite short proof that $Y(N,Q) = \pi N + Q^2$ 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 [1977], and Selberg [1991, but known by him before 1977] gave proofs that the bound holds with

$$Y_0(N,\delta) = N - 1 + \delta^{-1},$$

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

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.

Lemma 1, 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.

Theorem 2, A Large Sieve Inequality 0. *Suppose that* $0 < \delta \leq \frac{1}{2}$ *and the* x_r, $r = 1 \dots, R$ *satisfy* $\|x_r - x_s\| \geq \delta$ *whenever* $r \neq s$. *Then*

$$\sum_{r=1}^{R} |S(x_r)|^2 \leq Y_0(N,\delta) \sum_{n=M+1}^{M+N} |a_n|^2$$

holds with

$$Y_0(N,\delta) = N + \frac{1}{\delta} \log \frac{3}{\delta}$$

and

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

holds with

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

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 $r \neq s$ the sum over n satisfies

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

Hence the non-diagonal terms contribute at most

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

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.

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

Theorem 3, A Large Sieve Inequality 1. *The inequality*

$$\sum_{r=1}^{R}|S(x_r)|^2\leq Y_1(N,\delta)\sum_{n=M+1}^{M+N}|a_n|^2$$

holds with

$$Y_1(N,\delta)\ll N+\frac{1}{\delta}$$

and

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

holds with

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

Proof. The reason for the log in the previous result is because the characteristic function of $[M+1,M+N]$ has jump discontinuities. The solution is to majorise it by a smooth upper bound. One quite simple way to do this is to insert a factor $f(n)$ in (12) which majorises the characteristic function of $[M+1,M+N]$ is

$$f(x)=\max\left(0,\frac{2(1-|n-N_0-M|)}{N}\right),$$

where $N_0 = \lceil N/2 \rceil$. When $x \in [M+1, M+N]$ we have $-N/2 \le 1 - N_0 \le x - N_0 - M \le N - N_0 \le N/2$. Thus on multiplying out and interchanging the order of summation the innermost sum over n becomes a Fejèr kernel

$$\begin{aligned}\sum_n f(n)e\big(n(x_r - x_s)\big) &= \frac{2}{N}e\big((N_0 - M)(x_r - x_s)\big)\sum_{h=-N}^{N}(N - |h|)e\big(h(x_r - x_s)\big)\\ &= \frac{2}{N}e\big((N_0 - M)(x_r - x_s)\big)\Big)\left|\sum_{j=0}^{N-1} e\big(j(x_r - x_s)\big)\right|^2\\ &= 2e\big((N_0 - M)(x_r - x_s)\big)\big)\frac{\sin^2 \pi N(x_r - x_s)}{N \sin^2 \pi(x_r - x_s)}\end{aligned}$$

and satisfies

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

in place of (13). Thus we find that

$$\sum_{n=M+1}^{M+N}\left|\sum_{r=1}^{R} b_r e(nx_r)\right|^2 \ll \sum_{r=1}^{R}|b_r|^2\sum_{s=1}^{R}\min\left(N, \frac{1}{N\|x_r - x_s\|^2}\right).$$

By the spacing hypothesis for the x_r it follows that this is

$$\ll \sum_{r=1}^{R}|b_r|^2\left(N + \sum_{k=1}^{\infty}\min\left(N, \frac{1}{N(k\delta)^2}\right)\right).$$

If $N\delta > 1$ then this is

$$\ll \sum_{r=1}^{R}|b_r|^2 N(1 + N^{-2}\delta^{-2}) \ll N$$

and if $N\delta \le 1$, then it is

$$\ll \sum_{r=1}^{R}|b_r|^2\left(N + \sum_{k \le N^{-1}\delta^{-1}} + \sum_{k > N^{-1}\delta^{-1}}\frac{1}{N(k\delta)^2} \ll N + \delta^{-1}\right).$$

Thus in every case we may take

$$Y_1(N, \delta) \ll N + \frac{1}{\delta}.$$

Selberg's argument which leads to an optimal $Y_1(N, \delta)$ is a more sophisticated variant of this idea. For more details see Montgomery's expository article (Montgomery [1978]).

A this stage anyone who is not familiar with Dirichlet characters, and especially primitive characters should read the appendix.

Theorem 4, A Large Sieve for Characters. *Suppose that*

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

Then

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

holds with

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

Proof. By Appendix Lemma ?, 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 that lemma,

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

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 5. *Suppose that $a_1, \ldots, a_M$, $b_1, \ldots, b_N$ 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 6. *Suppose that $x\ge 2$, Then on the premises of Lemma 5,*

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

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 $Y > 0$ and $A > 0$, one has

$$\int_A^\infty \frac{\sin Y\alpha}{\alpha} d\alpha \ll \frac{1}{YA}$$

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

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

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 5,

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

and

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

Appendix on properties of Dirichlet characters

This is an adumbration of parts of Chapters 4 and 9 of MNT.

It is often useful to represent the characteristic function of a reduced residue class (mod q) as a linear combination of totally multiplicative functions $\chi(n)$ each one supported on the reduced residue classes and having period q. These are the *Dirichlet characters*. In the fancy language of abstract algebra we are examining the structure of homomorphisms from the units modulo q to an isomorphism of this group on the unit circle in the complex plane. Fundamental is that the homomorphisms themselves form a group which is also isomorphic to the original group.

Since $\chi(n)$ has period q we may think of it as mapping from residue classes, and since $\chi(n) \neq 0$ if and only if $(n, q) = 1$, we may think of χ as mapping from the multiplicative group of reduced residue classes to the multiplicative group $\mathbb{C}^\times$ of non-zero complex numbers. As χ is totally multiplicative, $\chi(mn) = \chi(m)\chi(n)$ for all m, n, we see that the map $\chi : \left(\mathbb{Z}/q\mathbb{Z}\right)^\times \longrightarrow \mathbb{C}^\times$ is a homomorphism. The method we use to describe these characters applies when $\left(\mathbb{Z}/q\mathbb{Z}\right)^\times$ is replaced by an arbitrary finite abelian group G, so we consider the slightly more general problem of finding all homomorphisms $\chi : G \to \mathbb{C}^\times$ from such a group G to $\mathbb{C}^\times$. We call these homomorphisms the characters of G, and let $\widehat{G}$ denote the set of all characters of G. We let χ_0 denote the *principal character*, whose value is identically 1. We note that if $\chi \in \widehat{G}$ then $\chi(e) = 1$ where e denotes the identity in G. Let n denote the order of G. If $g \in G$ and $\chi \in \widehat{G}$, then $g^n = e$, and hence $\chi(g^n) = 1$. Consequently $\chi(g)^n = 1$, and so we see that all values taken by characters are n^{th} roots of unity. In particular, this implies that $\widehat{G}$ is finite, since there can be at most n^n such maps. If χ_1 and χ_2 are two characters of G, then we can define a product character $\chi_1\chi_2$ by $\chi_1\chi_2(g) = \chi_1(g)\chi_2(g)$. For $\chi \in \widehat{G}$, let $\overline{\chi}$ be the character $\overline{\chi(g)}$. Then $\chi \cdot \overline{\chi} = \chi_0$, and we see that $\widehat{G}$ is a finite abelian group with identity χ_0. The following lemmas prepare for a full description of $\widehat{G}$ in Theorem 9 below.

Lemma 7. *Suppose that G is cyclic of order n, say $G = (a)$. Then there are exactly n characters of G, namely $\chi k(a^m) = e(km/n)$ for $1 \leq k \leq n$. Moreover,*

$$\sum_{g \in G} \chi(g) = \begin{cases} n & \text{if } \chi = \chi_0, \\ 0 & \text{otherwise,} \end{cases} \tag{12}$$

and

$$\sum_{\chi \in \widehat{G}} \chi(g) = \begin{cases} n & \text{if } g = e, \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

In this situation, $\widehat{G}$ is cyclic, $\widehat{G} = (\chi_1)$.

Proof. Suppose that $\chi \in \widehat{G}$. As we have observed, $\chi(a)$ is a nth root of unity, say $\chi(a) = e(k/n)$ for some k, $1 \le k \le n$. Hence $\chi(a^m) = \chi(a)^m = e(km/n)$. Since the characters are now known explicitly, the remaining assertions are easily verified.

Next we describe the characters of the direct product of two groups in terms of the characters of the factors.

Lemma 8. *Suppose that G_1 and G_2 are finite abelian groups, and that $G = G_1 \otimes G_2$. If χi is a character of G_i, $i = 1, 2$, and $g \in G$ is written $g = (g_1, g_2)$, $g_i \in G_i$, then $\chi(g) = \chi_1(g_1)\chi_2(g_2)$ is a character of G. Conversely, if $\chi \in \widehat{G}$, then there exist unique $\chi i \in G_i$ such that $\chi(g) = \chi_1(g_1)\chi_2(g_2)$. The identities* 12) *and* 13) *hold for G if they hold for both G_1 and G_2.*

We see here that each $\chi \in \widehat{G}$ corresponds to a pair $(\chi_1, \chi_2) \in \widehat{G}_1 \times \widehat{G}_2$. Thus $G \cong \widehat{G}_1 \otimes \widehat{G}_2$.

Proof. The first assertion is clear. As for the second, put $\chi_1(g_1) = \chi((g_1, e_2))$, $\chi_2(g_2) = \chi((e_1, g_2))$. Then $\chi i \in \widehat{G}_i$ for $i = 1, 2$, and $\chi_1(g_1)\chi_2(g_2) = \chi(g)$. The χi are unique, for if $g = (g_1, e_2)$ then

$$\chi(g) = \chi((g_1, e_2)) = \chi_1(g_1)\chi_2(e_2) = \chi_1(g_1),$$

and similarly for χ_2. If $\chi(g) = \chi_1(g_1)\chi_2(g_2)$, then

$$\sum_{g \in G} \chi(g) = \Big(\sum_{g_1 \in G_1} \chi_1(g_1) \Big) \Big(\sum_{g_2 \in G_2} \chi_2(g_2) \Big),$$

so that 12) holds for G if it holds for G_1 and for G_2. Similarly, if $g = (g_1, g_2)$, then

$$\sum_{\chi \in \widehat{G}} \chi(g) = \Big(\sum_{\chi_1 \in \widehat{G}_1} \chi_1(g_1) \Big) \Big(\sum_{\chi_1 \in \widehat{G}_2} \chi_2(g_2) \Big),$$

so that (13) holds for G if it holds for G_1 and G_2.

Theorem 9. *Let G be a finite abelian group. Then $\widehat{G}$ is isomorphic to G, and* 12) *and* (13) *both hold.*

Proof. Any finite abelian group is isomorphic to a direct product of cyclic groups, say

$$G \cong C_{n_1} \otimes C_{n_2} \otimes \cdots \otimes C_{n_r}.$$

The result then follows immediately from the lemmas.

Though G and $\widehat{G}$ are isomorphic, the isomorphism is not canonical. That is, no particular one-to-one correspondence between the elements of G and those of $\widehat{G}$ is naturally distinguished.

Corollary 10. *The multiplicative group* $\big(\mathbb{Z}/q\mathbb{Z}\big)^\times$ *of reduced residue classes* (mod q) *has* $\varphi(q)$ *Dirichlet characters. If* χ *is such a character, then*

$$\sum_{\substack{n=1\\(n,q)=1}}^{q} \chi(n) = \begin{cases} \varphi(q) & \text{if } \chi = \chi_0, \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

If $(n,q)=1$ *then*

$$\sum_{\chi} \chi(n) = \begin{cases} \varphi(q) & \text{if } n \equiv 1 \pmod q, \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

where the sum is extended over the $\varphi(q)$ *Dirichlet characters* χ (mod q).

As we remarked at the outset, for our purposes it is convenient to define the Dirichlet characters (mod q) on all integers; we do this by setting $\chi(n)=0$ when $(n,q)>1$. Thus χ is a totally multiplicative function with period q that vanishes whenever $(n,q)>1$, and any such function is a Dirichlet character (mod q). In this book a character is understood to be a Dirichlet character unless the contrary is indicated.

Corollary 11. *If* χi *is a character* (mod q_i) *for* $i=1,2$, *then* $\chi_1(n)\chi_2(n)$ *is a character* (mod $[q_1,q_2]$). *If* $q=q_1q_2$, $(q_1,q_2)=1$, *and* χ *is a character* (mod q), *then there exist unique characters* χi (mod q), $i=1,2$, *such that* $\chi(n)=\chi_1(n)\chi_2(n)$ *for all* n.

Proof. The first assertion follows immediately from the observations that $\chi_1(n)\chi_2(n)$ is totally multiplicative, that it vanishes if $(n,[q_1,q_2])>1$, and that it has period $[q_1,q_2]$. As for the second assertion, we may suppose that $(n,q)=1$. By the Chinese Remainder Theorem we see that

$$\big(\mathbb{Z}/q\mathbb{Z}\big)^\times \cong \big(\mathbb{Z}/q_1\mathbb{Z}\big)^\times \otimes \big(\mathbb{Z}/q_2\mathbb{Z}\big)^\times$$

if $(q_1,q_2)=1$. Thus the result follows from Lemma 2.

Our proof of Theorem 9 depends on Abel's Theorem that any finite abelian group is isomorphic to the direct product of cyclic groups, but we can prove Corollary 10 without appealing to this result, as follows. By the Chinese Remainder Theorem we see that

$$\big(\mathbb{Z}/q\mathbb{Z}\big)^\times \cong \bigotimes_{p^\alpha \| q} \big(\mathbb{Z}/p^\alpha\mathbb{Z}\big)^\times.$$

If p is odd then the reduced residue classes (mod p^α) form a cyclic group; in classical language we say there is a primitive root g. Thus if $(n,p)=1$ then there is a unique ν (mod $\varphi(p^\alpha)$) such that $g^\nu \equiv n \pmod{p^\alpha}$. The number ν is called the index of n, and is denoted $\nu = \operatorname{ind}_g n$. From Lemma 2 it follows that the characters (mod p^α), $p>2$, are given by

$$\chi k(n) = e\Big(\frac{k \operatorname{ind}_g n}{\varphi(p^\alpha)}\Big) \tag{16}$$

for $(n,p)=1$. We obtain $\varphi(p^\alpha)$ different characters by allowing k to assume integral values in the range $1 \le k \le \varphi(p^\alpha)$. By Lemma 8 it follows that if q is odd then the general character (mod q) is given by

$$\chi(n) = e\Big(\sum_{p^\alpha \| q} \frac{k \operatorname{ind}_g n}{\varphi(p^\alpha)}\Big) \tag{17}$$

for $(n,q)=1$, where it is understood that $k = k(p^\alpha)$ is determined (mod $\varphi(p^\alpha)$) and that $g = g(p^\alpha)$ is a primitive root (mod p^α).

The multiplicative structure of the reduced residues (mod 2^α) is more complicated. For $\alpha = 1$ or $\alpha = 2$ the group is cyclic (of order 1 or 2, respectively), and (16) holds as before. For $\alpha \ge 3$ the group is not cyclic, but if n is odd then there exist unique μ (mod 2) and ν (mod $2^{\alpha-2}$) such that $n \equiv (-1)^\mu 5^\nu \pmod{2^\alpha}$. In group-theoretic terms this means that

$$\big(\mathbb{Z}/2^\alpha\mathbb{Z}\big)^\times \cong C_2 \otimes C_{2^{\alpha-2}}$$

when $\alpha \ge 3$. By Lemma 8 the characters in this case take the form

$$\chi(n) = e\Big(\frac{j\mu}{2} + \frac{k\nu}{2^{\alpha-2}}\Big) \tag{18}$$

for odd n where $j = 0$ or 1 and $1 \le k \le 2^{\alpha-2}$. Thus (17) holds if $8 \nmid q$, but if $8|q$ then the general character takes the form

$$\chi(n) = e\Big(\frac{j\mu}{2} + \frac{k\nu}{2^{\alpha-2}} + \sum_{\substack{p^\alpha \| q \\ p>2}} \frac{\ell \operatorname{ind}_g n}{\varphi(p^\alpha)}\Big) \tag{19}$$

when $(n,q)=1$.

By definition, if $f(n)$ is totally multiplicative, $f(n)=0$ whenever $(n,q)>1$, and $f(n)$ has period q, then f is a Dirichlet character (mod q). It is useful to note that the first condition can be relaxed.

Theorem 12. *If f is multiplicative, $f(n)=0$ whenever $(n,q)>1$, and f has period q, then f is a Dirichlet character modulo q.*

Proof. It suffices to show that f is totally multiplicative. If $(mn,q)>1$ then $f(mn) = f(m)f(n)$ since $0=0$. Suppose that $(mn,q)=1$. Hence in particular $(m,q)=1$, so that the map $k \mapsto n+kq \pmod m$ permutes the residue classes (mod m). Thus there is a k for which $n+kq \equiv 1 \pmod m$, and consequently $(m, n+kq)=1$. Then

$$\begin{aligned} f(mn) &= f(m(n+kq)) && \text{(by periodicity)} \\ &= f(m)f(n+kq) && \text{(by multiplicativity)} \\ &= f(m)f(n) && \text{(by periodicity),} \end{aligned}$$

and the proof is complete.

Primitive characters

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 13. *Let χ be a character* (mod q)*. We say that d is a* quasiperiod *of χ if $\chi(m) = \chi(n)$ whenever $m \equiv n$* (mod 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 14. *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 15. *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 $\pmod{q/d}$, the numbers $n = ac + kcd$ run through all residues $\pmod{q}$ 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.

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 16. *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 (mod q/d), the numbers $n = hc + kcd$ run through all residues (mod 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

Theorem 17, The Pólya [1918]–I. M. Vinogradov [1918] inequality. *Suppose that χ is a non-principal character modulo q. Then*

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

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

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 16 we have

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

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

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

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.