BULLETIN OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 84, Number 4, July 1978

THE ANALYTIC PRINCIPLE OF THE LARGE SIEVE

BY HUGH L. MONTGOMERY

E. Bombieri [12] has written at length concerning applications of the large sieve to number theory. Our intent here is to complement his exposition by devoting our attention to the analytic principle of the large sieve; we describe only briefly how applications to number theory are made. The large sieve was studied intensively during the decade 1965–1975, with the result that the subject has lost its mystery: We now possess a variety of simple ideas which provide very precise results and a host of variants. While the large sieve can no longer be considered deep, it nevertheless gives powerful estimates in many different settings.

1. Historical background. The large sieve originates in a short paper of Ju. V. Linnik [51]. Linnik [52] made a simple application to the distribution of quadratic nonresidues, but it was A. Rényi [72]–[81] who systematically studied the large sieve, and who first made an important application to number theory: Using the large sieve, Rényi [72], [73] was the first to show that every large even number $2n$ can be expressed in the form $2n = p + P_k$, where p is prime and P_k has at most k prime factors. (Rényi did not determine a value for k, but M. B. Barban [2], [3] showed that one can take $k = 4$. The mean value theorem of Bombieri enables one to take $k = 3$, and Chen [17], [18] (see also [36], [84]) has obtained $k = 2$. In all of these arguments the large sieve is a major tool.) The large sieve remained the province of a few specialists, until the appearance in 1965 of a fundamental paper of K. F. Roth [86], followed immediately by a major contribution of Bombieri [7]. As we consider it here, the large sieve was first reduced to its basic analytic principle by H. Davenport and H. Halberstam [21].

2. The nature of the large sieve. For $M + 1 \leqslant n \leqslant M + N$ we let a_n be arbitrary complex numbers, and we form the trigonometric polynomial

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

here $e(\theta) = e^{2\pi i\theta}$, so that $S(\alpha)$ has period 1. Let $\alpha_1, \dots, \alpha_R$ be points which are well spaced (mod 1) in the sense that

This is an expanded account of an invited address entitled *The large sieve for the mathematician in the street*, presented at the 733rd meeting of the American Mathematical Society, University of Illinois, Urbana, March 20, 1976; received by the editors October 20, 1977.

AMS (MOS) subject classifications (1970). Primary 10H30.

Key words and phrases. Large sieve, Bessel's inequality, Hilbert's inequality.

$$\|\alpha_r - \alpha_s\| \geqslant \delta \tag{1}$$

for $r \neq s$; here $\|\theta\|$ denotes the distance to the nearest integer,

$$\|\theta\| = \min_n |\theta - n|.$$

The large sieve is an inequality of the form

$$\sum_{r=1}^{R} |S(\alpha_r)|^2 \leqslant \Delta \sum_{n=M+1}^{M+N} |a_n|^2, \tag{2}$$

where $\Delta = \Delta(N, \delta)$. The parameter M is irrelevant, for if $T(\alpha) = \sum_{M+1+K}^{M+N+K} a_{n-K} e(n\alpha)$ then $T(\alpha) = e(K\alpha)S(\alpha)$, and hence $|T(\alpha)| = |S(\alpha)|$. We are interested in determining how Δ depends on N and δ. We find, for example, that we can take $\Delta = N + \delta^{-1}$, but before considering admissible Δ we observe that Δ can not be too small.

Suppose that $a_n = e(-n\alpha_1)$. Then

$$|S(\alpha_1)|^2 = N^2 = N \sum |a_n|^2;$$

thus $\Delta \geqslant N$. In fact we can take $\Delta = N$ when $R = 1$, for by Cauchy's inequality

$$|S(\alpha)|^2 \leqslant N \sum |a_n|^2. \tag{3}$$

If the α_r are equally spaced then we may regard $R^{-1} \sum |S(\alpha_r)|^2$ as a Riemann sum approximating to $\int_0^1 |S(\alpha)|^2 \, d\alpha$. This prompts us to note that

$$\int_0^1 \sum_r |S(\alpha_r + \alpha)|^2 \, d\alpha = R \int_0^1 |S(\alpha)|^2 \, d\alpha = R \sum |a_n|^2.$$

Hence for some value of α,

$$\sum |S(\alpha_r + \alpha)|^2 \geqslant R \sum |a_n|^2,$$

and thus $\Delta \geqslant [\delta^{-1}] \geqslant \delta^{-1} - 1$.

The power of the large sieve may be attributed to the fact that we need not take Δ to be much larger than is necessitated by the elementary considerations above. In taking $\Delta = N + \delta^{-1}$ we see that Δ does not depend very heavily on δ as long as $N\delta$ is large.

3. An elementary inequality of the Sobolev type. P. X. Gallagher [29] has given a very simple derivation of the large sieve, based on the idea that $|S(\alpha)|^2$ is approximately $\delta^{-1} \int_{\alpha-\delta/2}^{\alpha+\delta/2} |S(\beta)|^2 \, d\beta$, to within an amount depending on the size of $S(\beta)$ and $S'(\beta)$. For our purposes the following simple inequality is sufficient.

Lemma 1. *Suppose* $f \in C^1[0, 1]$. *Then for* $0 \leqslant x \leqslant 1$,

$$|f(x)| \leqslant \int_0^1 (|f| + |f'|),$$

and

$$|f(\tfrac{1}{2})| \leqslant \int_0^1 (|f| + \tfrac{1}{2}|f'|).$$

PROOF. We easily verify that

$$f(x) = \int_0^1 f(u)\,du + \int_0^x uf'(u)\,du + \int_x^1 (u-1)f'(u)\,du.$$

The multiplier of $f'(u)$ has modulus not exceeding 1, and if $x = \frac{1}{2}$ then it has modulus $\leqslant \frac{1}{2}$. Thus the lemma is immediate.

In Lemma 1 we find a simple representative of a large class of inequalities, studied by Sobolev (see [1], [92]), in which a norm of f is bounded in terms of other norms of f and f'.

Using Lemma 1 we shall obtain

THEOREM 1 (GALLAGHER). *The inequality* (2) *holds with* $\Delta = \delta^{-1} + \pi N$.

This bound is asymptotically correct when $N\delta$ is small, but the secondary term is larger than it need be. Bombieri and Davenport [15] have shown that when $N\delta \leqslant 1$, the optimal Δ satisfies the bounds

$$\delta^{-1}\left(1 + \tfrac{1}{12}(N\delta)^3\right) \leqslant \Delta(N, \delta) \leqslant \delta^{-1}\left(1 + 270(N\delta)^3\right).$$

PROOF. We change variables in Lemma 1 in order to treat the interval $[\alpha_r - \frac{1}{2}\delta, \alpha_r + \frac{1}{2}\delta]$; we find that

$$|f(\alpha_r)| \leqslant \delta^{-1}\int_{\alpha_r-\delta/2}^{\alpha_r+\delta/2} |f(\beta)|\,d\beta + \frac{1}{2}\int_{\alpha_r-\delta/2}^{\alpha_r+\delta/2} |f'(\beta)|\,d\beta.$$

Taking $f(\alpha) = S(\alpha)^2$, we deduce that

$$|S(\alpha_r)|^2 \leqslant \delta^{-1}\int_{\alpha_r-\delta/2}^{\alpha_r+\delta/2} |S(\beta)|^2\,d\beta + \int_{\alpha_r-\delta/2}^{\alpha_r+\delta/2} |S(\beta)S'(\beta)|\,d\beta.$$

The intervals $(\alpha_r - \frac{1}{2}\delta, \alpha_r + \frac{1}{2}\delta)$ are nonoverlapping (mod 1). Hence

$$\sum_{r=1}^{R} |S(\alpha_r)|^2 \leqslant \delta^{-1}\int_0^1 |S(\beta)|^2\,d\beta + \int_0^1 |S(\beta)S'(\beta)|\,d\beta.$$

Here by Parseval's identity the first term on the right is $= \delta^{-1}\Sigma|a_n|^2$. By Cauchy's inequality and Parseval's identity the second term on the right is

$$\leqslant \left(\int_0^1 |S|^2\right)^{1/2}\left(\int_0^1 |S'|^2\right)^{1/2} = \left(\sum |a_n|^2\right)^{1/2}\left(\sum |2\pi n a_n|^2\right)^{1/2}$$

$$\leqslant 2\pi\left(\max_{M+1\leqslant n\leqslant M+N} |n|\right)\sum |a_n|^2.$$

We have already observed that $\Delta(N, \delta)$ is independent of M; thus we may assume that $M = -[\frac{1}{2}(N+1)]$. Then $\max|n| \leqslant \frac{1}{2}N$, and we obtain the desired bound,

$$\sum_{r=1}^{R} |S(\alpha_r)|^2 \leqslant (\delta^{-1} + \pi N)\sum |a_n|^2.$$

4. Duality. If X and Y are two Banach spaces and the linear operator A maps X to Y, then we define the norm of A to be

$$\|A\| = \sup_{x\in X} \|Ax\|/\|x\|.$$

If X^* and Y^* denote the dual spaces of X and Y, then the adjoint operator adj A maps Y^* to X^*, and as a general principle

$$\|\text{adj}\, A\| = \|A\|. \tag{4}$$

If

$$1/p + 1/q = 1, \qquad p \ge 1, q \ge 1, \tag{5}$$

then l_p is dual to l_q. Hence if A is a matrix, $A\colon l_{p_1} \to l_{p_2}$ then adj $A\colon l_{q_2} \to l_{q_1}$; we now derive (4) in this special setting. Our interest is confined to the special case $p_1 = p_2 = 2$; we prove more than we require in order to emphasize the generality of this duality principle.

Lemma 2. *For $i = 1, 2$ let p_i and q_i be fixed numbers satisfying* (5), *and let $[c_{nr}]$ be a fixed $N \times R$ matrix. Then the following three assertions concerning the constant D are equivalent*:

(i) *For any x_n,*

$$\left(\sum_r \left|\sum_n c_{nr}x_n\right|^{p_2}\right)^{1/p_2} \le D\left(\sum_n |x_n|^{p_1}\right)^{1/p_1}.$$

(ii) *For any x_n, y_r,*

$$\left|\sum_{n,r} c_{nr}x_ny_r\right| \le D\left(\sum_n |x_n|^{p_1}\right)^{1/p_1}\left(\sum_r |y_r|^{q_2}\right)^{1/q_2}.$$

(iii) *For any y_r,*

$$\left(\sum_n \left|\sum_r c_{nr}y_r\right|^{q_1}\right)^{1/q_1} \le D\left(\sum_r |y_r|^{q_2}\right)^{1/q_2}.$$

Proof. We may assume that $1 < p_i < \infty$, as the limiting cases follow by continuity. (i) implies (ii). By Hölder's inequality,

$$\begin{aligned}\left|\sum_{n,r} c_{nr}x_ny_r\right| &= \left|\sum_r y_r \sum_n c_{nr}x_n\right| \\ &\le \left(\sum_r |y_r|^{q_2}\right)^{1/q_2}\left(\sum_r \left|\sum_n c_{nr}x_n\right|^{p_2}\right)^{1/p_2}.\end{aligned}$$

We use (i) to bound the second factor on the right to obtain (ii).

(ii) implies (i). Put $L_r = \sum_n c_{nr}x_n$, and take $y_r = |L_r|^{p_2-2}\overline{L_r}$. By (ii),

$$\begin{aligned}\sum_r |L_r|^{p_2} = \sum_r L_ry_r &= \sum_{n,r} c_{nr}x_ny_r \\ &\le D\left(\sum |x_n|^{p_1}\right)^{1/p_1}\left(\sum |y_r|^{q_2}\right)^{1/q_2}.\end{aligned}$$

But $|y_r|^{q_2} = |L_r|^{p_2}$, so the above gives

$$\left(\sum_r |L_r|^{p_2}\right)^{1-1/q_2} \le D\left(\sum |x_n|^{p_1}\right)^{1/p_1},$$

which is (i).

The equivalence of (ii) and (iii) is the same.

The most direct approach to bounding $\sum_r |S(\alpha_r)|^2$ would involve multiplying out the square and taking the sum over r inside:

$$\sum_r |S(\alpha_r)|^2 = \sum_{m,n} a_m \overline{a_m} \sum_r e((m-n)\alpha_r).$$

The problem with this is that our weak information concerning the α_r does not permit us to evaluate or estimate the inner sum. However, by duality the inequality (2) is equivalent to having

$$\sum_n \left| \sum_r y_r e(n\alpha_r) \right|^2 \le \Delta \sum |y_r|^2$$

for all y_r. Here the left-hand side is

$$= \sum_{r,s} y_r \overline{y_s} \sum_{M+1}^{M+N} e(n(\alpha_r - \alpha_s))$$

$$= N \sum_{r=1}^{R} |y_r|^2 + \sum_{r \ne s} y_r e(T\alpha_r) \overline{y_s}\, e(-T\alpha_s) \frac{\sin \pi N(\alpha_r - \alpha_s)}{\sin \pi(\alpha_r - \alpha_s)},$$

where $T = M + N + \frac{1}{2}$. Taking $u_r = y_r e(T\alpha_r)$, we see that we may take $\Delta = N + C$, provided that

$$\left| \sum_{r \ne s} u_r \overline{u_s} \frac{\sin \pi N(\alpha_r - \alpha_s)}{\sin \pi(\alpha_r - \alpha_s)} \right| \le C \sum |u_r|^2. \tag{6}$$

In §6 we shall show that this holds with $C = \delta^{-1}$, but first we note that our problem can be made easier by introducing weights.

Suppose that $b_n \ge 0$ for all n, that $b_n > 0$ for $M + 1 \le n \le M + N$, and consider the inequality

$$\sum_r \left| \sum_{M+1}^{M+N} a_n e(n\alpha_r) \right|^2 \le B \sum_{M+1}^{M+N} |a_n|^2 b_n^{-1}. \tag{7}$$

By changing variables we see that this is equivalent to the inequality

$$\sum_r \left| \sum_{M+1}^{M+N} a_n b_n^{1/2} e(n\alpha_r) \right|^2 \le B \sum_{M+1}^{M+N} |a_n|^2,$$

which by duality is equivalent to

$$\sum_{M+1}^{M+N} b_n \left| \sum_r y_r e(n\alpha_r) \right|^2 \le B \sum_r |y_r|^2.$$

This in turn is implied by the inequality

$$\sum_{-\infty}^{+\infty} b_n \left| \sum_r y_r e(n\alpha_r) \right|^2 \le B \sum_r |y_r|^2.$$

By squaring out on the left and taking the sum over n inside, we see that this may be written

$$\sum_{r,s} y_r \overline{y_s}\, B(\alpha_r - \alpha_s) \le B \sum |y_r|^2, \tag{8}$$

where

$$B(\alpha)=\sum_{-\infty}^{+\infty} b_n e(n\alpha).$$

If $b_n = 1$ for $M+1 \leqslant n \leqslant M+N$, $b_n = 0$ otherwise then we arrive again at (6). If $b_n \geqslant 1$ for $M+1 \leqslant n \leqslant M+N$ then (7) gives (2) with $\Delta = B$; in §7 we discuss the problem of finding such b_n for which we can establish (8) with a reasonably small value of B.

Matthews [54], [55], [56], Kobayashi [48], [49] and Elliott [24] were the first to consider the large sieve via duality. Earlier, Rényi had based his arguments on generalizations of Bessel's inequality. We now formulate such an inequality, which provides a second method of reducing the large sieve to (6) or (8). However, we see ultimately that the generalized Bessel inequality is equivalent to the case $p_1 = p_2 = 2$ of Lemma 2.

5. Bessel's inequality. Let $\xi, \varphi_1, \ldots, \varphi_R$ be vectors in an inner product space. If the φ_r are orthonormal then Bessel's inequality asserts that

$$\sum_{r=1}^{R} |(\xi, \varphi_r)|^2 \leqslant \|\xi\|^2.$$

In order to have an inequality of this sort it is not necessary to assume that the φ_r are orthonormal; we have

LEMMA 3 (BOAS [6]). *Let $\varphi_1, \ldots, \varphi_R$ be vectors in an inner product space. Then the following two assertions concerning the number B are equivalent*:

(iv) *For any ξ,*

$$\sum_{r=1}^{R} |(\xi, \varphi_r)|^2 \leqslant B\|\xi\|^2.$$

(v) *For any y_r,*

$$\sum_{r,s} y_r \overline{y_s}\,(\varphi_r, \varphi_s) = \Big\|\sum_r y_r \varphi_r\Big\|^2 \leqslant B \sum_r |y_r|^2.$$

Note that if the φ_r are orthonormal then we clearly have $B = 1$ in (v), so that (iv) gives Bessel's inequality.

PROOF. (iv) implies (v). Take $\xi = \sum y_r \varphi_r$. Then by Cauchy's inequality and (iv),

$$\|\xi\|^2 = \sum_r \overline{y_r}\,(\xi, \varphi_r) \leqslant \Big(\sum_r |y_r|^2\Big)^{1/2}\Big(\sum_r |(\xi, \varphi_r)|^2\Big)^{1/2}$$
$$\leqslant B^{1/2}\|\xi\|\Big(\sum_r |y_r|^2\Big)^{1/2}.$$

This gives (v).

(v) implies (iv). For any y_r,

$$0 \leqslant \Big\|\xi - \sum_r y_r \varphi_r\Big\|^2$$
$$= \|\xi\|^2 - 2\operatorname{re} \sum_r \overline{y_r}\,(\xi, \varphi_r) + \Big\|\sum_r y_r \varphi_r\Big\|^2.$$

Then by (v),

$$\left(2\mathrm{re}\sum_r \overline{y_r}\,(\xi, \varphi_r) \leq \|\xi\|^2 + B\sum |y_r|^2\right).$$

Taking $y_r = B^{-1}(\xi, \varphi_r)$, we see that this is (iv).

The bilinear form in (v) is positive definite and Hermitian, so the best constant B is precisely the largest eigenvalue of the inner product matrix $[(\varphi_r, \varphi_s)]$. If we take $\xi = \{x_n\}$, $\varphi_r = \{c_{nr}\}$ then the equivalence of (iv) and (v) becomes that of (i) and (iii) in the case $p_1 = p_2 = 2$. Conversely, by taking an orthonormal basis for span $(\varphi_1, \ldots, \varphi_R)$ then we find that we can derive Lemma 3 from this case of Lemma 2. If we take $\xi = \{a_n b_n^{-1/2}\}$, $\varphi_r = \{b_n^{1/2} e(-n(\alpha_r))\}$, then (iv) becomes (7) and (v) becomes (8).

Having shown that the best constant in (iv) is the same as that in (v), it would be helpful to have a way of bounding the bilinear form in (v). This generally a tricky business, but as a first step in this direction we have

LEMMA 4. *Let $C = [c_{rs}]$ be an arbitrary Hermitian matrix, and let $\kappa_1, \ldots, \kappa_R$ be positive numbers. If*

$$\sum_s \kappa_s |c_{rs}| \leq B\kappa_r \qquad (1 \leq r \leq R) \tag{9}$$

then for arbitrary y_r,

$$\left|\sum_{r,s} c_{rs} y_r \overline{y_s}\right| \leq B\sum |y_r|^2. \tag{10}$$

In particular, we can take $B = \max_r \sum_s |c_{rs}|$.

PROOF. From the inequality $|\alpha\beta| \leq \frac{1}{2}|\alpha|^2 + \frac{1}{2}|\beta|^2$ with $\alpha = y_r\kappa_r^{-1}$, $\beta = \overline{y_s}\,\kappa_s^{-1}$, we find that

$$\left|\sum c_{rs} y_r \overline{y_s}\right| \leq \sum_{r,s} |c_{rs}|\kappa_r\kappa_s\left(\frac{1}{2}\left|\frac{y_r}{\kappa_r}\right|^2 + \frac{1}{2}\left|\frac{y_s}{\kappa_s}\right|^2\right)$$

$$= \sum_r |y_r|^2 \kappa_r^{-1} \sum_s \kappa_s |c_{rs}|.$$

An appeal to (9) now gives the desired bound. For the last assertion we take $\kappa_r = 1$ for all r.

Let $\rho(C)$ denote the maximum modulus of the eigenvalues of a square matrix C. Perron proved that $\rho(C) \leq B$ for arbitrary C, if B satisfies (9). But $\|C\| = \rho(C)$ for Hermitian C, so the above is a consequence of Perron's theorem. Our criterion (9) depends only on $|c_{rs}|$, and as such it is best possible: If the c_{rs} are nonnegative then $\rho(C)$ is an eigenvalue whose eigenvector has positive coordinates. Taking the κ_r to be these coordinates, we see that we may take $B = \rho(C)$ in (9). Unfortunately, if the c_{rs} vary greatly in sign or argument then the bound provided by Lemma 4 is usually rather weak.

Combining Lemmas 3 and 4 we obtain an inequality of Bombieri **[10]**:

$$\sum_r |(\xi, \varphi_r)|^2 \leq \|\xi\|^2 \max_r \sum_s |(\varphi_r, \varphi_s)|. \tag{11}$$

Similarly we can derive a bound of Selberg (see Bombieri [10]),

$$\sum_r |(\xi, \varphi_r)|^2 \Big(\sum_s |(\varphi_r, \varphi_s)|\Big)^{-1} \leqslant \|\xi\|^2.$$

Boas [6] and Bellman [5] have given another generalization of Bessel's inequality, which also follows by combining Lemma 3 with a simple upper bound for the bilinear form in (v).

6. Hilbert's inequality. Hilbert showed that

$$\left|\sum_{r \neq s} \frac{w_r \overline{w_s}}{r - s}\right| \leqslant c \sum |w_r|^2 \tag{12}$$

with $c = 2\pi$. Schur [90] was the first to obtain the best constant $c = \pi$, and Toeplitz gave the following elegant proof: Let $W(\alpha) = \sum w_r e(r\alpha)$, $K(\alpha) = \sum_{k \neq 0} k^{-1} e(k\alpha)$. Then

$$\left|\sum_{r \neq s} \frac{w_r \overline{w_s}}{r - s}\right| = \left|\int_0^1 |W(\alpha)|^2 K(\alpha)\, d\alpha\right|$$

$$\leqslant (\text{ess sup}|K|) \int_0^1 |W(\alpha)|^2\, d\alpha.$$

But $K(0) = 0$ and $K(\alpha) = \pi - 2\pi\alpha$ for $0 < \alpha < 1$, so we have (12) with $c = \text{ess sup}|K| = \pi$. In fact (12) holds with strict inequality.

With the aim of establishing (6), we generalize Hilbert's inequality as follows.

THEOREM 2 (MONTGOMERY AND VAUGHAN [63]). *Suppose that* $\lambda_1 < \lambda_2 < \cdots < \lambda_R$, *and that* $\lambda_{r+1} - \lambda_r \geqslant \delta$ *for* $1 \leqslant r < R$. *Then for any* w_r,

$$\left|\sum_{r \neq s} \frac{w_r \overline{w_s}}{\lambda_r - \lambda_s}\right| \leqslant \pi\delta^{-1} \sum |w_r|^2.$$

PROOF. By Cauchy's inequality the above is

$$= \left|\sum_r w_r \sum_{\substack{s \\ s \neq r}} \frac{\overline{w_s}}{\lambda_r - \lambda_s}\right|$$

$$\leqslant \Big(\sum_r |w_r|^2\Big)^{1/2} \left[\sum_r \left|\sum_{\substack{s \\ s \neq r}} \frac{\overline{w_s}}{\lambda_r - \lambda_s}\right|^2\right]^{1/2}.$$

Thus it suffices to show that

$$\sum_r \left|\sum_{\substack{s \\ s \neq r}} \frac{\overline{w_s}}{\lambda_r - \lambda_s}\right|^2 \leqslant \pi^2\delta^{-2} \sum |w_r|^2, \tag{13}$$

as is also evident by Lemma 2. We multiply out the square on the left and take the sum over r inside to see that the left-hand side is

$$= \sum_{s,t} \overline{w_s} w_t \sum_{\substack{r \\ r \neq s \\ r \neq t}} (\lambda_r - \lambda_s)^{-1}(\lambda_r - \lambda_t)^{-1}.$$

Writing the diagonal terms separately we see that this is

$$= \sum_{s} |w_s|^2 \sum_{\substack{r \\ r \neq s}} (\lambda_r - \lambda_s)^{-2}$$

$$+ \sum_{\substack{s,t \\ s \neq t}} \frac{\overline{w_s} w_t}{\lambda_s - \lambda_t} \sum_{\substack{r \\ r \neq s \\ r \neq t}} (\lambda_r - \lambda_s)^{-1} - (\lambda_r - \lambda_t)^{-1}.$$

Call the first term Σ_1. The second term we write as the difference of two terms in which the inner summands are $(\lambda_r - \lambda_s)^{-1}$ and $(\lambda_r - \lambda_t)^{-1}$, respectively. In the first of these we introduce the new term for $r = t$, and similarly for the second. Thus the second term above is

$$= \sum_{s \neq t} \frac{\overline{w_s} w_t}{\lambda_s - \lambda_t} \sum_{\substack{r \\ r \neq s}} (\lambda_r - \lambda_s)^{-1}$$

$$- \sum_{s \neq t} \frac{\overline{w_s} w_t}{\lambda_s - \lambda_t} \sum_{\substack{r \\ r \neq t}} (\lambda_r - \lambda_t)^{-1} + 2 \sum_{s \neq t} \frac{\overline{w_s} w_t}{(\lambda_s - \lambda_t)^2}$$

$$= \Sigma_2 - \Sigma_3 + 2\Sigma_4,$$

say. Thus far we have followed Schur's proof of (12); at this point Schur uses his assumption that $\lambda_r = r$ to show that the inner sums in Σ_2 and Σ_3 telescope and hence $\Sigma_2 = \Sigma_3 = 0$. Since this does not succeed for us, we introduce a new idea: We may assume that the w_r are extremal. Since our coefficient matrix is skew-Hermitian, the extremal w_r are the coordinates of an eigenvector. Hence there is a real μ such that

$$\sum_{\substack{r \\ r \neq s}} w_r(\lambda_r - \lambda_s)^{-1} = i\mu w_s \qquad (1 < s < R).$$

Taking the sum over t inside in Σ_2, and using the above, we find that

$$\Sigma_2 = -i\mu \sum_{s} |w_s|^2 \sum_{\substack{r \\ r \neq s}} (\lambda_r - \lambda_s)^{-1}.$$

Making the same simplification in Σ_3 we find that $\Sigma_2 = \Sigma_3$ for those extremal w_r. Thus we have

$$\sum_{r} \left| \sum_{\substack{s \\ s \neq r}} \overline{w_s}(\lambda_r - \lambda_s)^{-1} \right|^2 < \Sigma_1 + 2\Sigma_4.$$

We use the inequality $|\overline{w_s} w_t| < \frac{1}{2}|w_s|^2 + \frac{1}{2}|w_t|^2$ in Σ_4, as in the proof of Lemma 4, and then see that $\Sigma_4 < \Sigma_1$. Thus to prove (13) it suffices to show that

$$\sum_{\substack{r \\ r \neq s}} (\lambda_r - \lambda_s)^{-2} < \tfrac{1}{3}\pi^2\delta^{-2}.$$

But $|\lambda_r - \lambda_s| \geqslant \delta|r - s|$, so the above sum is

$$\leqslant \sum_{\substack{r \\ r \neq s}} \delta^{-2}(r-s)^{-2} < 2\delta^{-2}\sum_{k=1}^{\infty} k^{-2} = \frac{1}{3}\pi^2\delta^{-2}.$$

We now deduce

COROLLARY 1. *For any* w_r,

$$\left|\sum_{r \neq s} \frac{w_r\overline{w_s}}{\sin \pi(\alpha_r - \alpha_s)}\right| \leqslant \delta^{-1}\sum |w_r|^2. \tag{14}$$

We could prove this by the method employed above, but the following simple argument now suffices. We appeal to Theorem 2 with a doubly-indexed set of RK variables w_{rm}, $1 \leqslant r \leqslant R$, $1 \leqslant m \leqslant K$, and well-spaced constants λ_{rm}. Then

$$\left|\sum_{\substack{r,s,m,n \\ (r,m)\neq(s,n)}} w_{rm}\overline{w_{sn}}\,(\lambda_{rm} - \lambda_{sn})^{-1}\right| \leqslant \pi\delta^{-1}\sum_{r,m} |w_{rm}|^2.$$

Now put $w_{rm} = w_r(-1)^m$, $\lambda_{rm} = \alpha_r + m$. Then

$$\left|\sum_{(r,m)\neq(s,n)} (-1)^{m-n}w_r\overline{w_s}(\lambda_r - \lambda_s + m - n)^{-1}\right| \leqslant K\pi\delta^{-1}\sum |w_r|^2.$$

As $\Sigma_r|w_r|^2\,\Sigma_{m\neq n}(-1)^{m-n}(m-n)^{-1} = 0$, we may replace the condition $(r, m) \neq (s, n)$ by the simpler condition $r \neq s$. We also put $k = m - n$ and divide by K to see that

$$\left|\sum_{r\neq s} w_r\overline{w_s}\sum_{-K}^{K}(1 - |k|/K)(-1)^k(\lambda_r - \lambda_s + k)^{-1}\right| \leqslant \pi\delta^{-1}\sum |w_r|^2. \tag{15}$$

But for $\alpha \notin \mathbf{Z}$,

$$(\sin \pi\alpha)^{-1} = \pi^{-1}\sum_k (-1)^k(\alpha - k)^{-1},$$

and hence

$$\lim_{K\to\infty}\sum_{-K}^{K}(1 - |k|/K)(-1)^k(\alpha + k)^{-1} = \frac{\pi}{\sin \pi\alpha}.$$

Thus on allowing K to tend to infinity we see that (15) gives (13).

COROLLARY 2. *For arbitrary real t, and any* u_r,

$$\left|\sum_{r\neq s} u_r\overline{u_s}\,\frac{\sin t(\alpha_r - \alpha_s)}{\sin \pi(\alpha_r - \alpha_s)}\right| \leqslant \delta^{-1}\sum |u_r|^2;$$

in the large sieve we may take $\Delta = N + \delta^{-1}$.

The second assertion follows from the first by (6). To see the first, we note that

$$u_r \overline{u_s} \sin t(\alpha_r - \alpha_s) = \frac{1}{2i} u_r \overline{u_s} e^{it(\alpha_r - \alpha_s)} - \frac{1}{2i} u_r \overline{u_s} e^{-it(\alpha_r - \alpha_s)},$$

and apply Corollary 1 twice with $w_r = u_r e^{\pm it\alpha_r}$.

Arguing similarly from Theorem 2, we have

COROLLARY 3. *For arbitrary a_r,*

$$\int_0^T \left| \sum_{r=1}^R a_r e^{i\lambda_r t} \right|^2 dt = (T + 2\pi\theta\delta^{-1}) \sum |a_r|^2$$

for some θ, $|\theta| \leqslant 1$.

Weaker results of this sort have been obtained by Wiener [97], Paley and Wiener [70], Marcinkievic and Zygmund [53], Ingham [46], and Titchmarsh [94].

The advantage of this approach to the large sieve is that it can be used to obtain more delicate results. Suppose that $|\lambda_r - \lambda_s| \geqslant \delta_r$ for any s distinct from r. Montogomery and Vaughan [63] have shown that

$$\left| \sum_{r \neq s} \frac{w_r \overline{w_s}}{\lambda_r - \lambda_s} \right| \leqslant \frac{3}{2} \pi \sum_r |w_r|^2 \delta_r^{-1}.$$

Selberg (unpublished) has shown that $3\pi/2$ can be replaced by 3.2, but it is not known whether the above holds with the constant π. One can derive analogues of the corollaries from the above; as a variant of the large sieve we obtain

$$\sum_r |S(\alpha_r)|^2 \left(N + \frac{3}{2} \delta_r^{-1} \right)^{-1} \leqslant \sum_{M+1}^{M+N} |a_n|^2. \tag{16}$$

Useful arithmetic applications of this are found in Montgomery and Vaughan [62]. Corresponding to Corollary 3 we have

$$\int_0^T \left| \sum a_r e^{i\lambda_r t} \right|^2 dt = \sum |a_r|^2 (T + 3\pi\theta\delta_r^{-1})$$

with $|\theta| \leqslant 1$. Hence in particular,

$$\int_0^T \left| \sum_{n=1}^{\infty} a_n n^{-it} \right|^2 dt = \sum_{n=1}^{\infty} |a_n|^2 (T + O(n)).$$

7. An extremal problem. We now consider the problem of deriving the large sieve from (8). From (8) and Lemma 4 we see that if $b_n \geqslant \chi_{[M+1,M+N]}(n)$ then we can take

$$\Delta = \max_r \sum_s |B(\alpha_r - \alpha_s)|$$

in the large sieve. Here $\chi_{\mathbb{S}}$ denotes the characteristic function of $\mathbb{S}$. If we take $b_n = \chi_{[M+1,M+N]}(n)$ then $|B(\alpha)| = |\sin \pi N\alpha / \sin \pi\alpha|$, and the above is $= N + O(\delta^{-1} \log \delta^{-1})$. We do better to take smoother b_n, for example $b_n = 0$ for $n \leqslant M + 1 - A$, $b_n = 1 - (M + 1 - n)/A$ for $M + 1 - A < n \leqslant M + 1$, $b_n = 1$ for $M + 1 < n \leqslant M + N$, $b_n = 1 - (n - M - N)/A$ for $M + N <$

$n \leqslant M + N + A$, $b_n = 0$ for $n > M + N + A$. For integral A we then have

$$|B(\alpha)| = |(\sin \pi A\alpha)(\sin \pi(N + A)\alpha)(\sin \pi\alpha)^{-2}|,$$

and on taking a suitable $A \approx \delta^{-1}$ we find that $\Delta \leqslant N + 2\delta^{-1}$; see Bombieri [10].

If $b_n \geqslant \chi_{[M+1,M+N]}(n)$ and $B(\alpha) = 0$ for $\|\alpha\| \geqslant \delta$ then from (8) we see that

$$\sum_r |S(\alpha_r)|^2 \leqslant B(0) \sum |a_n|^2. \tag{17}$$

One way to construct such b_n is to construct a function $b(x)$ such that $b(x) \geqslant \chi_{[M+1,M+N]}(x)$, $b \in L^1(\mathbf{R})$, $\hat{b}(t) = 0$ for $|t| \geqslant \delta$. Then we put $b_n = b(n)$, and see by the Poisson summation formula that

$$B(\alpha) = \sum_n b(n)e(n\alpha) = \sum_k \hat{b}(k - \alpha).$$

Hence $B(\alpha) = 0$ for $\|\alpha\| \geqslant \delta$, and $B(0) = \hat{b}(0) = \int_{-\infty}^{+\infty} b(x)\, dx$.

As Selberg (unpublished) observed, the folllowing lemma provides a means of making a good choice of $b(x)$.

Lemma 5. *Let*

$$F(z) = \left(\frac{\sin \pi z}{\pi}\right)^2 \left(\sum_{n=0}^{\infty} (z - n)^{-2} + \sum_{n=1}^{\infty} (z + n)^{-2} + 2z^{-1}\right).$$

Then F is entire, $F(z) = O(e^{2\pi|\operatorname{Im} z|})$, $F(x) \geqslant \operatorname{sgn} x$ for real x, and

$$\int_{-\infty}^{+\infty} F(x) - \operatorname{sgn} x\, dx = 1. \tag{18}$$

Since $F \notin L^1(\mathbf{R})$ we can not define $\hat{F}$ in the usual way, but the estimate for $|F|$ can be interpreted to mean that $\hat{F}(t) = 0$ for $|t| \geqslant 1$. Beurling has shown that among such functions which majorize sgn x, the one considered here is the unique function which minimizes the quantity in (18).

Proof. The first two assertions are clear. For the third we recall that

$$\left(\frac{\sin \pi z}{\pi}\right)^2 \left(\sum_{-\infty}^{+\infty} (z - n)^{-2}\right) = 1,$$

and note that for $x > 0$

$$\sum_{n=1}^{\infty} (x + n)^{-2} < \sum_{n=1} \int_{x+n-1}^{x+n} u^{-2}\, du = x^{-1}$$

$$= \sum_{n=0}^{\infty} \int_{x+n}^{x+n+1} u^{-2}\, du < \sum_{n=0}^{\infty} (x + n)^{-2}.$$

From these relations we see that $F(x) \geqslant \operatorname{sgn} x$ for all x. Finally

$$\int_{-\infty}^{+\infty} F(x) - \operatorname{sgn} x\, dx$$

$$= \int_0^{+\infty} F(x) + F(-x)\, dx = 2\int_0^{+\infty} \left(\frac{\sin \pi x}{\pi x}\right)^2 dx = 1.$$

For $d > 0$ we now put $G(x) = \frac{1}{2}F(x) + \frac{1}{2}F(d - x)$. Then $G(x) \geqslant \chi_{[0,d]}(x)$, $G \in L^1(\mathbf{R})$, $\hat{G}(t) = 0$ for $|t| \geqslant 1$, and

$$G(x) = \chi_{[0,d]}(x) + \tfrac{1}{2}(F(x) - \operatorname{sgn} x) + \tfrac{1}{2}(F(d - x) - \operatorname{sgn}(d - x)),$$

so that $\hat{G}(0) = \int_{-\infty}^{+\infty} G(x)\,dx = d + 1$. We can now take $d = \delta(N - 1)$, and put $b(x) = G(\delta x)$. Then $b(x) \geqslant \chi_{[0,N-1]}(x)$, $b \in L^1(\mathbf{R})$, and $\hat{b}(t) = \delta^{-1}\hat{G}(t/\delta)$. Thus $\hat{b}(t) = 0$ for $|t| \geqslant \delta$, $\hat{b}(0) = \delta^{-1}(d + 1) = N - 1 + \delta^{-1}$. Hence from (17) we have

THEOREM 3 (SELBERG). *The large sieve is valid with* $\Delta = N - 1 + \delta^{-1}$.

NOTE ADDED IN PROOF. Paul Cohen has observed that Theorem 3 can be derived from Corollary 2 by means of the following trick: Let $T(\alpha) = S(K\alpha)$, so that

$$K\sum_{r=1}^{R} |S(\alpha_r)|^2 = \sum_{k=1}^{K} \sum_{r=1}^{R} |S(\alpha_r + k)|^2 = \sum_{k=1}^{K} \sum_{r=1}^{R} |T((\alpha_r + k)/K)|^2.$$

The points $(\alpha_r + k)/K$ are spaced by δ/K, so by Corollary 2 applied to $T(\alpha)$, the above is $\leqslant (K(N - 1) + 1 + K/\delta)\Sigma|a_n|^2$. We divide by K and let K tend to infinity to obtain Theorem 3.

Among those $G \in L^1(\mathbf{R})$ for which $G(x) \geqslant \chi_{[0,d]}(x)$, $\hat{G}(t) = 0$ for $|t| \geqslant 1$, it can be shown that there is one for which $\hat{G}(0)$ is minimal. Since $\hat{G}(0) = \int G \geqslant \int \chi_{[0,d]} = d$, we write $\hat{G}(0) = d + \theta(d)$ for this minimal G; then $\theta(d) \geqslant 0$. The construction above shows that $\theta(d) \leqslant 1$ for all d. Our method gives

$$\sum_r |S(\alpha_r)|^2 \leqslant (N - 1 + \delta^{-1}\theta(\delta(N - 1))) \sum_{M+1}^{M+N} |a_n|^2. \tag{19}$$

If $\alpha_r = r/R + \alpha_0$, $\delta^{-1} = R^{-1}$, $R|(N - 1)$, $a_n = ae(-n\alpha_0)$ for $R|n$, $a_n = 0$ otherwise, then

$$\sum_{r=1}^{R} \left|\sum_{n=0}^{N-1} a_n e(n\alpha_r)\right|^2 = R|a|^2\left(\frac{N-1}{R} R + 1\right)^2$$

$$= (N - 1 + \delta^{-1}) \sum_{0}^{N-1} |a_n|^2,$$

so we have equality in Theorem 3 and in (19). Selberg has shown that this is the only situation in which equality occurs in (19); hence the same may be said for Theorem 3.

8. Applications to number theory. Take the points α_r to be the numbers a/q with $(a, q) = 1$, $q \leqslant Q$. If $a/q \neq a'/q'$ then

$$\left\| \frac{a}{q} - \frac{a'}{q'} \right\| = \left\| \frac{aq' - a'q}{qq'} \right\| \geqslant (qq')^{-1} \geqslant Q^{-2};$$

thus we may take $\delta = Q^{-2}$, and we have

$$\sum_{q \leqslant Q} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} |S(a/q)|^2 \leqslant (N + Q^2) \sum_{M+1}^{M+N} |a_n|^2. \tag{20}$$

Almost all applications of the large sieve to number theory proceed from (20), although the applications are of two different sorts. Less frequently the dual of (20) is useful:

$$\sum_{M+1}^{M+N} \left| \sum_{q<Q} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \lambda_{q,a} e(na/q) \right|^2 \leqslant (N+Q^2) \sum_{q<Q} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} |\lambda_{q,a}|^2.$$

Let $\mathfrak{N}$ be a set of Z integers in $[M+1, M+N]$, put $S(\alpha) = \sum_{n\in\mathfrak{N}} e(n\alpha)$, and let $Z(q, h)$ denote the number of members of $\mathfrak{N}$ which are congruent to $h \pmod q$. Then

$$S(a/q) = \sum_{n\in\mathfrak{N}} e(an/q) = \sum_{h=1}^{q} e(ah/q) Z(q,h),$$

and hence

$$\sum_{a=1}^{q} \left| S\left(\frac{a}{q}\right) \right|^2 = \sum_{a} \left| \sum_{h} e(ah/q) Z(q,h) \right|^2$$

$$= \sum_{h=1}^{q} \sum_{k=1}^{q} Z(q,h) Z(q,k) \sum_{a=1}^{q} e\left(\frac{a(h-k)}{q} \right).$$

The innermost sum vanishes unless $h \equiv k \pmod q$, in which case it is equal to q. Thus we find that

$$\sum_{a=1}^{q} \left| S\left(\frac{a}{q}\right) \right|^2 = q \sum_{h=1}^{q} Z(q,h)^2. \tag{21}$$

The average of $Z(q, h)$ is Z/q, since

$$\sum_{h=1}^{q} Z(q,h) = Z. \tag{22}$$

Noting that $Z = S(1)$, we see from (21) and (22) that

$$q \sum_{h=1}^{q} (Z(q,h) - Z/q)^2 = q \sum_{h=1}^{q} Z(q,h)^2 - Z^2 = \sum_{a=1}^{q-1} \left| S\left(\frac{a}{q}\right) \right|^2.$$

Here a runs over all nonzero residue classes, whereas in (20) we must restrict a to reduced residue classes in (20). However, if q is prime then these conditions coincide; hence by (20) we have

$$\sum_{p<Q} p \sum_{h=1}^{p} (Z(p,h) - Z/p)^2 \leqslant (N+Q^2) Z. \tag{23}$$

This is a powerful estimate, since among all sets $\mathfrak{N}$ with Z elements in $[M+1, M+N]$, for most of them the left-hand side is $\approx ZQ^2/\log Q$; see Erdős and Rényi [27].

Now let $\omega(p)$ be the number of $h \pmod p$ for which $Z(p, h) = 0$. Then

$$\sum_{h=1}^{p} (Z(p,h) - Z/p)^2 \geqslant \omega(p)(Z/p)^2,$$

and hence by (23),

$$Z \leqslant (N + Q^2)\Big(\sum_{p<Q} \omega(p)/p\Big)^{-1}.$$

At last we see the large sieve formulated as a sieve! If in particular $\mathcal{P}$ is a set of primes $p \leqslant N^{1/2}$ for which $\omega(p) \geqslant \tau p$, then $|\mathcal{P}| \leqslant 2N(\tau Z)^{-1}$. Here the emphasis is on primes for which $\omega(p)$ is large; hence the term *large* sieve. If we apply the above when $\omega(p)$ is small then we obtain weak results. This may be traced to the fact that we are using (20) only for prime values of q. Montgomery [58] derived form (20) a sharper bound,

$$Z \leqslant (N + Q^2)\Big(\sum_{q<Q} \mu^2(q) \prod_{p|q} \frac{\omega(p)}{p - \omega(p)}\Big)^{-1};$$

a simple proof of this has been given by Johnsen [47] and Gallagher [32] (see also Bombieri [13, p. 20]). The above bound is comparable to that obtained by Selberg's method. Indeed the above can be derived from the dual of the large sieve, making clear the connection with Selberg's sieve; see Halberstam and Richert [36, pp. 125–126], Huxley [44], Kobayashi [49], Matthews [56], Motohashi [69]. Letting $\pi(x; q, a)$ denote the number of primes $p \leqslant x$ for which $p \equiv a \pmod q$, we can easily derive the bound

$$\pi(x + y; q, a) - \pi(x; q, a) \leqslant \frac{2y}{\varphi(q)\log y/q}\left(1 + O\left(\frac{\log\log 3y/q}{\log 2y/q}\right)\right)$$

for $y \geqslant q$. By arguing more carefully from (16), Montgomery and Vaughan [62] have shown that the above is valid without the error term on the right.

A second application of the large sieve to number theory arises in bounding averages of a character sum

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

where χ is a Dirichlet character (mod q). The $\varphi(q)$ characters (mod q) are orthogonal, so that

$$\sum_{\chi} |T(\chi)|^2 = \varphi(q) \sum_{\substack{h=1\\(h,q)=1}}^{q} \Big|\sum_{n\equiv n(q)} a_n\Big|^2 \leqslant (N + q) \sum_{M+1}^{M+N} |a_n|^2;$$

here χ runs over all characters (mod q). If we wish to sum over $q \leqslant Q$ as well, then we use the large sieve. Gallagher [29] has given a simple proof that

$$\sum_{\chi}{}^{*} |T(\chi)|^2 \leqslant \frac{\varphi(q)}{q} \sum_{\substack{a=1\\(a,q)=1}}^{q} \Big|S\Big(\frac{a}{q}\Big)\Big|^2,$$

where the sum over χ is restricted to primitive characters χ (mod q). Hence by (20),

$$\sum_{q<Q} \frac{q}{\varphi(q)} \sum_{\chi}{}^{*} |T(\chi)|^2 \leqslant (N + Q^2) \sum_{M+1}^{M+N} |a_n|^2.$$

This inequality is an essential tool in studying L-functions and the distribution of prime numbers. It is central to the proof of Bombieri's mean value theorem, which asserts that

$$\sum_{q<Q} \max_{y<x} \max_{\substack{a \\ (a,q)=1}} \left|\pi(y; q, a) - \frac{li(y)}{\varphi(q)}\right| = O_A\left(x(\log x)^{-A}\right)$$

provided $Q \leqslant x^{1/2}(\log x)^{-B}$; $B = B(A)$. Vaughan (to appear) has given a very elegant proof of this estimate. Other important applications of the large sieve have been made by Bombieri and Davenport [13], Davenport and Halberstam [22], Chen [17], [18], Montgomery [60], Gallagher [31], Montgomery and Vaughan [64], and Hooley [40]. Barban [4] and Bombieri [12] have surveyed the applications of the large sieve to number theory.

9. Variants. We consider first variants in which we drop the condition (1) of well-spacing, and substitute other conditions on the α_r; we have already formulated one such variant in (16). Suppose that $N_\delta(\alpha)$ is the number of r for which $\|\alpha - \alpha_r\| < \delta$. Then

$$\sum_r N_\delta(\alpha_r)^{-1}|S(\alpha_r)|^2 \leqslant (\pi N + \delta^{-1})\sum |a_n|^2. \tag{24}$$

This implies Theorem 1; it is easily derived (see Montgomery [61, Theorem 2.1]) by combining the technique of §3 with the observation that

$$\sum_{\substack{r \\ \|\alpha-\alpha_r\|<\delta/2}} N_\delta(\alpha_r)^{-1} \leqslant 1$$

for all α. Following an unpublished observation of Bombieri, we may use the ideas of §§4, 7 to show that if μ is a measure on $\mathbf{T}$ then for any $\delta > 0$

$$\int_0^1 |S(\alpha)|^2\, d\mu(\alpha) \leqslant (N + 2\delta^{-1})\left(\max_\alpha \int_\alpha^{\alpha+\delta} d\mu\right)\sum_{M+1}^{M+N} |a_n|^2.$$

The method of §3 is very flexible, and can be applied to other classes of functions, such as ordinary polynomials or Dirichlet polynomials; see Davenport [20] and Montgomery [59]. We can also consider exponents other than 2; see Davenport and Halberstam [21] and Forti and Viola [28]. For example,

$$\sum_{r=1}^R |S(\alpha_r)|^q = O_q\left((N + \delta^{-1})\sum |a_n|^q(|n| + 2)^{q-2}\right)$$

for real $q \geqslant 2$. However, this tells us little more than we know already if the $|a_n|$ are generally of the same size. To see this, suppose that $|a_n| \leqslant 1$ for all n. Then the upper bound above is $O_q((N + \delta^{-1})N^{q-1})$. But $|S(\alpha)| \leqslant N$, so that $|S(\alpha)|^q \leqslant N^{q-2}|S(\alpha)|^2$, and hence by the large sieve,

$$\sum_{r=1}^R |S(\alpha_r)|^q \leqslant (N + \delta^{-1})N^{q-1}$$

if $q \geqslant 2$ and $|a_n| \leqslant 1$ for all n.

Let

$$(Mf)(\alpha) = \sup_{h>0} \frac{1}{2h} \int_{\alpha-h}^{\alpha+h} |f(\beta)|\, d\beta$$

be the maximal function of f. Montgomery (to appear) has shown that

$$\sum_{r=1}^{R} |(MS)(\alpha_r)|^2 \leqslant C(N+\delta^{-1}) \sum_{M+1}^{M+N} |a_n|^2.$$

Also, by Hunt's theorem on the strong [2, 2] boundedness of the maximal partial sum operator,

$$\sum_{r=1}^{R} \max_{k} \left| \sum_{M+1 \leqslant n \leqslant k} a_n e(n\alpha_r) \right|^2 \leqslant C(N+\delta^{-1}) \sum_{M+1}^{M+N} |a_n|^2;$$

this improves on an elementary estimate of Uchiyama [95].

In the dual of the large sieve we can derive a corresponding lower bound with the factor $N + 1 - \delta^{-1}$, but for the large sieve itself we generally have no such lower bound; see Boas [6]. Wolke [103] has discussed lower bound counterparts of (20). Lower bounds for the irregularity of the distribution of a set $\mathfrak{N}$ of integers into arithmetic progressions have been given by Roth [85], Montgomery [61, Chapter 5], Huxley [45], and Montgomery and Sárközy (to appear).

The large sieve is a bound for the norm of a matrix whose coefficients are $e(n\alpha_r)$. Here the α_r may be irregularly spaced, but the n lie in arithmetic progression. Selberg (unpublished) has observed that we can obtain similar results for the more general coefficients $e(\nu_k\alpha_r)$: Suppose that

$$T_0 = \alpha_0 < \alpha_1 < \cdots < \alpha_R = T_0 + T$$

and that $\alpha_r - \alpha_{r-1} \geqslant \delta$ for $1 \leqslant r \leqslant R$. Suppose also that

$$M_0 = \nu_0 < \nu_1 < \cdots < \nu_K = M_0 + N$$

and that $\nu_k - \nu_{k-1} \geqslant \Delta$ for $1 \leqslant k \leqslant K$. Then for any a_k,

$$\sum_{r=0}^{R} \left| \sum_{k=0}^{K} a_k e(\nu_k\alpha_r) \right|^2 \leqslant \left(T + \delta + \frac{1}{\Delta}\right)\left(\frac{1}{\delta} + \pi N\right) \sum_{k=0}^{K} |a_k|^2.$$

To establish this we proceed as in §3, and then appeal to Corollary 3. If we dualize first then we obtain the above with the factor on the right replaced by $(\pi T + 1/\Delta)(1/\delta + \Delta + N)$. Considering the symmetry of the situation, it would be desirable to have a symmetric upper bound, such as $(T + 1/\Delta)(N + 1/\delta)$.

Let $\mathfrak{Q}$ be a set of X integers $q \leqslant Q$. Then we may ask for a factor $\Delta(N, Q, X)$ such that

$$\sum_{q\in\mathfrak{Q}} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} |S(a/q)|^2 \leqslant \Delta(N, Q, X) \sum_{M+1}^{M+N} |a_n|^2.$$

By (20) we have $\Delta(N, Q, X) \leqslant N + Q^2$. Alternatively, by the large sieve we see that

$$\sum_{\substack{a=1\\(a,q)=1}}^{q} |S(a/q)|^2 \leqslant \sum_{a=1}^{q} |S(a/q)|^2 \leqslant (N+q) \sum_{M+1}^{M+N} |a_n|^2;$$

this can also be seen by a direct elementary argument. Summing over $q \in \mathfrak{Q}$, we find that $\Delta(N, Q, X) \leqslant X(N+Q)$. Hence

$$\Delta(N, Q, X) \leqslant \min(N + Q^2, XN + XQ). \tag{25}$$

Moreover, examples can be constructed to show that the above is never more than a constant factor from the truth. This is disappointing, since one might have hoped to be able to take $\Delta(N, Q, X) = C(N + QX)$. For certain sets $\mathfrak{Q}$ one can do a little better than (25); see Wolke **[99]**, **[101]**, Sokolovskiĭ **[93]**. Burgess **[16]** has shown that

$$\sum_{q\in\mathfrak{Q}} \sum_{\substack{a=1\\(a,q)=1}}^{q} |S(a/q)| \leqslant C\big(QX(N+QX)\sum |a_n|^2\big)^{1/2}.$$

This may be derived by Cauchy's inequality from (24) with $\delta = (QX)^{-1}$, since it may be shown that

$$\sum_{q\in\mathfrak{Q}} \sum_{\substack{a=1\\(a,q)=1}}^{q} N_\delta(a/q) = O(QX).$$

Multidimensional versions of the large sieve have been established for use in algebraic number fields; see Rieger **[82]**, **[83]**, Samandarov **[88]**, Huxley **[41]**–**[43]**, Wilson **[98]**, Schaal **[89]**, Hlawka **[38]**, **[39]**.

BIBLIOGRAPHY

1. R. A. Adams, *Sovolev spaces*, Academic Press, New York, 1975.

2. M. B. Barban, *The density of zeros of Dirichlet L-series and the problem of the addition of primes and almost primes*, Dokl. Akad. Nauk UzSSR **1** (1963), 9–10. (Russian)

3. ______, *The "density" of the zeros of Dirichlet L-series and the problem of the sum of primes and "near primes,"* Mat. Sb. **61** **(103)** (1963), 418–425.

4. ______, *The "large sieve" method and its applications in the theory of numbers*, Uspehi Mat. Nauk **21** (1966), 51–102 = Russian Math. Surveys **21** (1966), 49–103.

5. R. Bellman, *Almost orthogonal series*, Bull. Amer. Math. Soc. **50** (1944), 517–519.

6. R. P. Boas, *A general moment problem*, Amer. J. Math. **63** (1941), 361–370.

7. E. Bombieri, *On the large sieve*, Mathematika **12** (1965), 201–225.

8. ______, *Nuovi metodi e nuovi risultati nella teoria dei numeri*, Bol. Un. Mat. Ital. (4) **1** (1968), 96–106.

9. ______, *On a theorem of van Lint and Richert*, Symposia Mathematica. IV (INDAM, Rome, 1968/69), Academic Press, London, 1970, pp. 175–180.

10. ______, *A note on the large sieve*, Acta Arith. **18** (1971), 401–404.

11. ______, *On the large sieve inequalities and their applications*, Proc. Internat. Conf. Number Theory (Moscow, 1971), Trudy Mat. Inst. Steklov **132** (1973), 251–256, 266.

12. ______, *Le grand crible dans la théorie analytique des nombres*, astérisque 18, Société Math. France, 1974.

13. E. Bombieri and H. Davenport, *Small differences between prime numbers*, Proc. Roy. Soc. Ser. A **293** (1966), 1–18.

14. ______, *On the large sieve method*, Number Theory and Analysis (Papers in honor of Edmund Landau), Plenum, New York, 1969, pp. 9–22.

15. ______, *Some inequalities involving trigonometrical polynomials*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. **23** (1969), 223–241.

16. D. A. Burgess, *The average of the least primitive root modulo p^2*, Acta Arith. **18** (1971), 263–271.

17. J. Chen, *On the representation of a large even integer as the sum of a prime and the product of at most two primes*, Kexue Tongbao **17** (1966), 385–386.

18. ______, *On the representation of a large even integer as a sum of a prime and the product of at most two primes*, Sci. Sinica **16** (1973), 157–176.

19. H. Davenport, *Multiplicative number theory*, Markham, Chicago, 1967. (Delete Theorems 4, 4A in §23.)

20. ______, *The zeros of trigonometrical polynomials*, Mathematika **19** (1972), 88–90.

21. H. Davenport and H. Halberstam, *The values of a trigonometric polynomial at well spaced points*, Mathematika **13** (1966), 91–96. *Corrigendum and addendum*, Mathematika **14** (1967), 229–232.

22. ______, *Primes in arithmetic progressions*, Michigan Math. J. **13** (1966), 485–489. *Corrigendum*, Michigan Math. J. **15** (1968), 505.

23. P. D. T. A. Elliott, *The Turán-Kubilius inequality, and a limitation theorem for the large sieve*, Amer. J. Math. **92** (1970), 293–300.

24. ______, *On inequalities of large sieve type*, Acta Arith. **18** (1971), 405–422.

25. ______, *On connections between the Turán-Kubilius inequality and the large sieve: some applications*, Proc. Sympos. Pure Math., vol. 24, Amer. Math. Soc., Providence, R.I., 1973, pp. 77–82.

26. P. Erdős, *Remarks on number theory*. V, Mat. Lapok **17** (1966), 135–155.

27. P. Erdős and A. Rényi, *Some remarks on the large sieve of Yu. V. Linnik*, Ann. Univ. Sci. Budapest. Eötvös Sect. Math. **11** (1968), 3–13.

28. M. Forti and C. Viola, *On the large sieve type estimates for the Dirichlet series operator*, Proc. Sympos. Pure Math., vol. 24, Amer. Math. Soc., Providence, R.I., 1973, pp. 31–49.

29. P. X. Gallagher, *The large sieve*, Mathematika **14** (1967), 14–20.

30. ______, *Bombieri's mean value theorem*, Mathematika **15** (1968), 1–6.

31. ______, *A large sieve density estimate near $\sigma = 1$*, Invent. Math. **11** (1970), 329–339.

32. ______, *Sieving by prime powers*, Acta Arith **24** (1973), 491–497.

33. ______, *The large sieve and probabilistic Galois theory*, Proc. Sympos. Pure Math., vol. 24, Amer. Math. Soc., Providence, R.I., 1973, pp. 91–101.

34. S. W. Graham, *Applications of sieve methods*, Ph.D. Dissertation, Univ. of Michigan, Ann Arbor, 1977.

35. H. Halberstam, *The large sieve*, Number Theory (Colloq. János Bolyai Math. Soc., Debrecen, 1968), North-Holland, Amsterdam, 1970, pp. 123–131.

36. H. Halberstam and H.-E. Richert, *Sieve Methods*, Academic Press, London, 1974.

37. H. Halberstam and K. F. Roth, *Sequences*, Oxford Univ. Press, London and New York, 1966.

38. E. Hlawka, *Bemerkungen zum grossen Sieb von Linnik*, Österreich. Akad. Wiss. Lit. Mainz Abh. Math.-Natur. Kl. S.-B. II **178** (1970), 13–18.

39. ______, *Zum grossen Sieb von Linnik*, Acta Arith. **27** (1975), 89–100.

40. C. Hooley, *On the Barban-Davenport-Halberstam theorem*. I, J. Reine Angew. Math. **274/275** (1975), 206–223.

41. M. N. Huxley, *The large sieve inequality for algebraic number fields*, Mathematika **15** (1968), 178–187.

42. ______, *The large sieve inequality for algebraic number fields*. II. *Means of moments of Hecke zeta-functions*, Proc. London Math. Soc. (3) **21** (1970), 108–128.

43. ______, *The large sieve inequality for algebraic number fields*. III. *Zero-density results*, J. London Math. Soc. (2) **3** (1971), 233–240.

44. ______, *The distribution of prime numbers*, Oxford Mathematical Monographs, Oxford Univ. Press, London and New York, 1972.

45. ______, *Irregularity in sifted sequences*, J. Number Theory **4** (1972), 437–454.

46. A. E. Ingham, *Some trigonometrical inequalities with applications to the theory of series*, Math. Z. **41** (1936), 367–379.

47. J. Johnsen, *On the large sieve method in $GF[q, x]$*, Mathematika **18** (1971), 172–184.

48. I. Kobayashi, *Remarks on the large sieve method*, Proc. United States-Japan Seminar on Number Theory, Tokyo, 1971.

49. ______, *A note on the Selberg sieve and the large sieve*, Proc. Japan Acad. **49** (1973), 1–5.

50. I. P. Kubilius, *Probabilistic methods in the theory of numbers*, Uspehi Mat. Nauk **11** (1956), no. 2 (68), 31–36.

51. Ju. V. Linnik, *The large sieve*, Dokl. Akad. Nauk SSSR **30** (1941), 292–294. (Russian)

52. ______, *A remark on the least quadratic non-residue*, Dokl. Akad. Nauk SSSR **36** (1941), 119–120. (Russian)

53. J. Marcinkievic and A. Zygmund, *Proof of a gap theorem*, Duke Math. J. **4** (1938), 469–472.

54. K. R. Matthews, *On a bilinear form associated with the large sieve*, J. London Math. Soc. **5** (1972), 567–570.

55. ______, *On an inequality of Davenport and Halberstam*, J. London Math. Soc. **4** (1972), 638–642.

56. ______, *Hermitian forms and the large and small sieves*, J. Number Theory **5** (1973), 16–23.

57. Ming-Chit Liu, *On a result of Davenport and Halberstam*, J. Number Theory **1** (1969), 385–389.

58. H. L. Montgomery, *A note on the large sieve*, J. London Math. Soc. **43** (1968), 93–98.

59. ______, *Mean and large values of Dirichlet polynomials*, Invent. Math. **8** (1969), 334–345.

60. ______, *Zeros of L-functions*, Invent. Math. **8** (1969), 346–354.

61. ______, *Topics in multiplicative number theory*, Lecture Notes in Math., vol. 227, Springer-Verlag, Berlin, 1971.

62. H. L. Montgomery and R. C. Vaughan, *The large sieve*, Mathematika **20** (1973), 119–134.

63. ______, *Hilbert's inequality*, J. London Math. Soc. (2) **8** (1974), 73–81.

64. ______, *The exceptional set in Goldbach's problem*, Acta Arith. **27** (1975), 353–370.

65. Y. Motohashi, *A note on the large sieve*, Proc. Japan Acad. **53** (1977), 17–19.

66. ______, *On Gallagher's prime number theorem*, Proc. Japan Acad. **53** (1977), 50–52.

67. ______, *Introduction to the theory of the distribution of prime numbers*, Sûgaku **26** (1974), no. 1, 1–12.

68. ______, *On the density theorem of Linnik*, Proc. Japan Acad. **51** (1975), suppl. 815–917.

69. ______, *A note on the large sieve*. II, Proc. Japan Acad. **53** (1977), 122–124.

70. R. E. A. C. Paley and N. Wiener, *Fourier transforms in the complex domain*, Amer. Math. Soc. Colloq. Publ. Vol. 19, Amer. Math. Soc., Providence, R.I., 1934.

71. P. A. B. Pleasants, *A sum related to the distribution modulo* 1 *of sets of real numbers*, Quart. J. Math. Oxford Ser. (2) **21** (1970), 321–336.

72. A. Rényi, *On the representation of an even number as the sum of a single prime and a single almost-prime number*, Dokl. Akad. Nauk SSSR **56** (1947), 455–458. (Russian)

73. ______, *On the representation of an even number as the sum of a single prime and a single almost-prime number*, Izv. Akad. Nauk SSSR Ser. Mat. **12** (1948), 57–78; English transl., Amer. Math. Soc. Transl. (2) **19** (1962), 299–321.

74. ______, *Un nouveau théorème concernant les fonctions indépendantes et ses applications à la théorie des nombres*, J. Math. Pures Appl. (9) **28** (1949), 137–149.

75. ______, *Probability methods in number theory*, Publ. Math. Collectae Budapest **1** (1949), no. 21, 1–9.

76. ______, *On a theorem of the theory of probability and its application in number theory*, Časopis Pešt. Mat. Fys. **74** (1949), 167–175.

77. ______, *Sur un théorème général de probabilité*, Ann. Inst. Fourier (Grenoble) **1** (1950), 43–52.

78. ______, *On the large sieve of Ju. V. Linnik*, Compositio Math. **8** (1950), 68–75.

79. ______, *On the probabilistic generalization of the large sieve of Linnik*, Magyar Tud. Akad. Mat. Kutató Int. Közl. **3** (1958), 199–206.

80. ______, *Probabilistic methods in number theory*, Proc. Internat. Congress Math., Cambridge Univ. Press, London and New York, 1958, pp. 529–539.

81. ______, *A new version of the probabilistic generalization of the large sieve*, Acta Math. Acad. Sci. Hungar **10** (1959), 217–226.

82. G. J. Rieger, *Zum Sieb von Linnik*, Arch. Math. **11** (1960), 14–22.

83. ______, *Das grosse Sieb von Linnik für algebraische Zahlen*, Arch. Math. **12** (1961), 184–187.

84. P. M. Ross, *On Chen's theorem that each large even number has the form $p_1 + p_2$ or $p_1 + p_2p_3$*, J. London Math. Soc. **10** (1975), 500–506.

85. K. F. Roth, *A remark on integer sequences*, Acta Arith. **9** (1964), 257–260.

86. ______, *On the large sieves of Linnik and Rényi*, Mathematika **12** (1965), 1–9.

87. ______, *The large sieve*. Inaugural Lecture, January 23, 1968, Imperial College of Science and Technology, London, 1968.

88. A. G. Samandarov, *On the large sieve in algebraic number fields*, Mat. Zametki **6** (1967), 673–680.

89. W. Schaal, *On the large sieve method in algebraic number fields*, J. Number Theory **2** (1970), 249–270.

90. I. Schur, *Bemerkungen zur Theorie der beschrankten Bilinearformen mit unendlich vielen Verändlichen*, J. Reine Angew. Math. **140** (1911), 1–28.

91. W. Schwarz, *Einführung in Siebmethoden der analytischen Zahlentheorie*, Bibliographisches Institut, Mannheim-Vienna-Zurich, 1974.

92. S. L. Sobolev, *Applications of functional analysis in mathematical physics*, Transl. Math. Monographs, vol. 7, 1963.

93. A. V. Sokolovskiĭ, *The large sieve*, Acta Arith. **25** (1973/74), 301–306.

94. E. C. Titchmarsh, *A class of trigonometrical series*, J. London Math. Soc. **3** (1928), 300–304.

95. S. Uchiyama, *The maximal large sieve*, Hokkaido Math. J. **1** (1972), 117–126.

96. A. I. Vinogradov, *On the density hypothesis for Dirichlet L-functions*, Izv. Akad. Nauk SSSR Ser. Mat. **29** (1965), 903–934. Correction: Izv. Akad. Nauk SSSR Ser. Mat. **30** (1966), 719–720.

97. N. Wiener, *A class of gap theorems*, Ann. Sculoa Norm. Sup. Pisa **3** (1934), 367–372.

98. R. J. Wilson, *The large sieve in algebraic number fields*, Mathematika **16** (1969), 189–204.

99. D. Wolke, *Farey-Bruche mit primen Nenner und das grosse Sieb*, Math. Z. **114** (1970), 145–158.

100. ______, *Einige Anwendungen des grossen Siebes auf zahlentheoretische Funktionen*, Habilitationsschrift, Phillipps-Universitat Marburg/Lahn, Marburg/Lahn, 1970.

101. ______, *On the large sieve with primes*, Acta Math. Acad. Sci. Hungar. **22** (1971/72), 239–247.

102. ______, *Über eine Ungleichung von A. I. Vinogradov*, Arch. Math. (Basel) **23** (1972), 625–629.

103. ______, *A lower bound for the large sieve inequality*, Bull. London Math. Soc. **6** (1974), 315–318.

Department of Mathematics, University of Michigan, Ann Arbor Michigan 48109