FUNCTIONAL ANALYSIS AND ADDITIVE ARITHMETIC FUNCTIONS

P. D. T. A. ELLIOTT

In memory of Professor M. Kac

Geometry Prince has familiar,
Twin;
Upon looking out,
Sees himself looking in.

1. A function is *arithmetic* if it is defined on the positive integers. Those arithmetic functions which assume real values and satisfy $f(ab) = f(a) + f(b)$ for mutually prime integers a, b are classically called *additive*. The following examples illustrate the interest of these functions, both for themselves and for their applications.

An additive function is defined by its values on the prime powers. I shall denote a typical prime power by q, and the prime of which it is a power by q_0. A well-known additive function is $\omega(n)$ which, with $\omega(q) = 1$, counts the number of distinct prime divisors of n. Let $\nu_x(n; \ldots)$ denote the frequency amongst the positive integers not exceeding x of those for which property $\ldots$ holds. Then as $x \to \infty$

$$\nu_x\left(n; \omega(n) - \log\log x \leqslant z(\log\log x)^{1/2}\right) \Rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2}\,dt.$$

More generally, for an additive function $f(n)$ define

$$A(x) = \sum_{q \leqslant x} \frac{f(q)}{q}\left(1 - \frac{1}{q_0}\right), \qquad B(x) = \left(\sum_{q \leqslant x} \frac{|f(q)|^2}{q}\left(1 - \frac{1}{q_0}\right)\right)^{1/2} \geqslant 0.$$

Then if $|f(q)| \leqslant 1$ for all q and $B(x)$ is unbounded with x, the celebrated theorem of Erdős-Kac [18, 19] asserts (essentially) that, as $x \to \infty$,

$$\nu_x\left(n; f(n) - A(x) \leqslant zB(x)\right) \Rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2}\,dt.$$

It is natural to ask for the weakest conditions on f under which such a result can hold.

The ideas applied by Erdős and Kac were realized in a wider conceptual setting by Kubilius [25], who generalized and elaborated their theorem. Unfortunately the archetypal result to which their method aspires is that, as $x \to \infty$,

(1) $\nu_x(n; f(n) - \alpha(x) \leqslant z\beta(x)) \Rightarrow F(z)$

for some $\alpha(x), \beta(x) > 0$ if and only if

(2) $P\left(\sum_{q \leqslant x} Y_q - \alpha(x) \leqslant z\beta(x) \right) \Rightarrow F(z),$

where the independent random variables Y_q respectively assume the values $f(q)$ with probability $q^{-1}(1 - q_0^{-1})$, and 0 with probability $1 - q^{-1}(1 - q_0^{-1})$; and this result is false. Not only do the mechanics of the underlying sieve method fail (see, for example, Elliott [11, Chapter 3, pp. 145–146]), but such an 'isomorphism' would force all the limiting distributions of the frequencies (1) to be infinitely divisible, and it is known that they are not.

In the work of Erdős-Kac and Kubilius the philosophy is maintained by specializing $\alpha(x)$ and $\beta(x)$ to $A(x)$ and $B(x)$, and requiring (although it is not formulated in this way) that, for each fixed $y > 0$, $B(x^y)/B(x)$ approach 1 as $x \to \infty$. One steps from (1) to (2) with the probabilistic device of truncation, and the 'isomorphism' indeed holds. The number-theoretic problem of deciding which frequencies (1) converge is reduced to a problem in the theory of probability proper, with a readily available solution. This result has been bought at a high price, however. The difficulties of handling an occasional large value of $|f(q)|$ have been legislated away, and as a consequence many interesting distributions $F(z)$, including all the stable laws beyond the normal law, cannot be considered.

In fact a more exact 'isomorphism' between the behavior of the distributions in (1) and (2) can be obtained, but the random variables Y_q must be defined differently. If we restrict ourselves to the consideration of measurable functions $\beta(x)$ which as $x \to \infty$ satisfy $\beta(x) \to \infty$ and $\beta(x^y)/\beta(x) \to 1$ for each fixed $y > 0$, then

(3) $\nu_x(n; f(n) - \alpha(x) \leqslant z\beta(x)) \Rightarrow F(z)$

for some $\alpha(x)$, if and only if

(4) $P\left(\sum_{q \leqslant x} Z_q - \alpha_1(x) \leqslant z\beta(x) \right) \Rightarrow F(z),$

where for some constant D the independent random variables Z_q assume values $f(q) - D \log q$ with probability $q^{-1}(1 - q_0^{-1})$, and 0 with probability $1 - q^{-1}(1 - q_0^{-1})$. Moreover, $\alpha_1(x) = \alpha(x) - D \log x$. The deduction of (3) from (4) may be found in Levin and Timofeev [31] (see also Elliott [6]); that of (4) from (3) is in Elliott [8]. Here there is no restriction upon the additive function f. The method of proof, quite different from that of Erdős-Kac and Kubilius, involves extensive appeal to Fourier analysis, in particular to the theory of Dirichlet series in the manner of Halász [20]. I do not pursue here the

question of whether the growth condition upon $\beta(x)$ is appropriate, noting only that, in general, necessary and sufficient conditions for the weak convergence of the arithmetic frequencies (3) are still not known. Instead, I would like to draw attention to the term $-D \log q$ in the definition of the variable Z_q. Where does this logarithm come from?

In another direction let $a_1, a_2, \ldots,$ be a sequence of positive rationals. Let Q^* be the group of positive rationals with multiplication as group law, Γ its subgroup generated by the a_i, and G the quotient group Q^*/Γ. A homomorphism of G into the additive reals may be viewed as an additive arithmetic function f, extended to the rationals by $f(r/s) = f(r) - f(s)$, which satisfies $f(a_i) = 0$ for all i. Conditions involving coprimality may be suppressed, so that, in the standard terminology, f is *completely additive*.

Considering additive functions to be characters on G, and allowing them (by abuse of meaning) to assume values in differing groups, one sees that the determination of G is equivalent to the characterization of various additive functions which vanish on the a_i. Moreover, this procedure may be given a quantitative aspect. Thus, in order that some integral power of the integer n be representable in the form $a_1^{d_1} \cdots a_k^{d_k}$ with integers d_j, it is necessary and sufficient that every real completely additive f with $f(a_i) = 0$, $i = 1, \ldots, k$, also satisfy $f(n) = 0$. An account of the appropriate mathematics, which generalizes and extends earlier results of Dress and Volkmann [3], Wolke [43], Meyer [32]—see also Ruzsa [35]—may be found in Elliott [13, Chapter 15]. In the present paper I shall confine myself to real or complex additive functions.

As an example of this procedure we set $a_n = (6n + 5)/(12n + 11)$, and employ the inequality

(5)
$$\sum_{\substack{q \leqslant x \\ (q,6)=1}} \frac{1}{q} |F(q) - M \log q|^2 \ll \sup_{w \leqslant x^c} \frac{1}{w} \sum_{x < n \leqslant w} |f(6n + 5) - f(12n + 11)|^2$$

with a constant c, and function M of x, which is contained in Theorem (10.1) of Elliott [13]. If now m is an integer which is not divisible by 2 or 3, then it has a representation

$$m^v = \prod_{i=1}^{k} \left(\frac{6n_i + 5}{12n_i + 11} \right)^{d_i}$$

with integers v and d_i, each n_i lying in the interval $m < n_i \leqslant (23m)^c$. For it follows from (5) that if $f(6n + 5) - f(12n + 11) = 0$ for $m < n \leqslant (23m)^c$, then in particular

$$M \log \frac{6m + 5}{12m + 11} = f\left(\frac{6m + 5}{12m + 11} \right) = 0,$$

so that $M = 0$. In this case the infinite generators of G are provided by the images of the primes 2 and 3 under the canonical map $Q^* \to G$. Another argument shows that G is finitely generated, so that it has its finite torsion group as a direct summand.

It is essential that in the inequality (5) no restriction should be placed upon f beyond that it be additive; and it is clearly desirable to obtain for c a value which is as close to 1 as possible. The general direction in classical probabilistic number theory is to derive the value distribution of an additive function f on the integers from its given distribution on the prime powers. Here, as in the step from (3) to (4), we are going in the opposite direction. We are concerned with a function whose values on a specific sequence of integers are known, and whose values on the prime powers we wish to determine. This is the passage, from Dirichlet series to Euler product, which is taken in the proof of the prime number theorem, save that we do not now have any regularity conditions satisfied by f which would guarantee an analytic continuation of an appropriate Dirichlet series. Once again a logarithm intervenes.

In the present paper I do not consider in detail the probabilistic theory of numbers, nor the multiplicative representation of integers. An extensive account of the former can be found in my two-volume work [11], and of the latter in its companion volume [13]. Instead, I shall give a functional analytic overview, and indicate a source for the logarithms. This view enables a clearer presentation of many of the problems, and introduces a new precision into certain results. As an example, I prove that with an appropriately chosen function M of x, the inequality

$$(6) \qquad \sum_{q \leqslant x} \frac{1}{q} |f(q) - M \log q|^2 \ll \max_{w \leqslant x^c} \frac{1}{w} \sum_{n \leqslant w} |f(n+1) - f(n)|^2$$

certainly holds with any fixed $c > 8$. Moreover, by varying the spaces on which we view various operators, generalizations are suggested which would otherwise be less obvious.

2. We begin with the Turán-Kubilius inequality

$$(7) \qquad \sum_{n \leqslant x} |f(n) - A(x)|^2 \leqslant c_1 x B(x)^2.$$

With $f(n) = \omega(n)$ and $A(x)$, $B(x)$ replaced by $\log \log x$, this inequality was first established by Turán [39]. Note that for $x \geqslant 2$

$$\sum_{q \leqslant x} \frac{1}{q} = \log \log x + O(1).$$

An important application which he had in view was the simpler proof of a result of Hardy and Ramanujan [21] that for each positive δ most integers n have $|\omega(n) - \log \log n| \leqslant (\log \log n)^{1/2 + \delta}$. Turán was well aware of the flexibility of his argument. As he himself wrote to me in 1976: "...from H. R.'s paper in 1917 seventeen years passed until 1934, without the slightest sign of realising that such general theorems exist at all." With hindsight we may see in (7) an analogue of Tchebyshev's inequality in the theory of probability, but Turán knew little of probability in 1934. Indeed, it was this very paper of Turán which promoted the interest of Kac in these matters, leading to the joint work with Erdős mentioned earlier. Lacking any other aesthetic, Turán [39] generalized his inequality to cover the cases when $f(p)$ is bounded and $B(x)$ is

replaced by $A(x)$. His argument was to expand the inner square, invert the order of summation, and show that in the resulting estimations the main terms cancel.

Whilst Erdős employed the method of Turán in many special circumstances, Kubilius gave it the outcome (7). He followed the original argument of Turán, together with judicious applications of the Cauchy-Schwarz inequality, employing elementary results on the distribution of prime numbers. By this time the influence of probability was paramount, and in his monograph [26] of 1962 Kubilius views the inequaltiy (7) as a species of the 'Law of Large Numbers.'

In the form given above the Turán-Kubilius inequality has another advantage (see Elliott [5]): the quantities on both sides may be viewed as the squares of norms. Let $\mathbf{C}^s$ be the complex space with one coordinate for each of the s prime powers q not exceeding x. Let $\mathbf{C}^{[x]}$ be the complex space with one coordinate for each of the $[x]$ positive integers n not exceeding x. If on the space $\mathbf{C}^s$ we employ the L^2 norm

$$\|y\| = \left(\sum_{q \leqslant x} |y_q|^2 \right)^{1/2},$$

and on $\mathbf{C}^{[x]}$ the corresponding norm

$$\|z\| = \left(\sum_{n \leqslant x} |z_n|^2 \right)^{1/2},$$

then the Turán-Kubilius inequality asserts that the map $\Lambda: \mathbf{C}^s \to \mathbf{C}^{[x]}$ given by

$$(8) \quad \Lambda(f)(n) = \sum_{q \| n} f(q) \left(q\left(1 - \frac{1}{q_0}\right) \right)^{-1/2} - \sum_{q \leqslant x} f(q) \left(\frac{1}{q}\left(1 - \frac{1}{q_0}\right) \right)^{1/2}$$

has norm bounded uniformly for all $x \geqslant 2$. Here $q \| n$ denotes that q divides n but qq_0 does not, and I have avoided discussion of the slightly differing formulations of the Turán-Kubilius inequality.

In this form the Turán-Kubilius inequality represents something intrinsic about the positive integers. This may be illustrated in another way by considering the dual map $\Lambda': (\mathbf{C}^{[x]})' \to (\mathbf{C}^s)'$, which has the same norm as Λ. Otherwise expressed, this yields that the inequality

$$(9) \quad \sum_{q \leqslant x} q\left(1 - \frac{1}{q_0}\right) \left| \sum_{\substack{n \leqslant x \\ n \cong 0 \,(\mathrm{mod}\, q)}} a_n - \frac{1}{q}\left(1 - \frac{1}{q_0}\right) \sum_{n \leqslant x} a_n \right|^2 \ll x \sum_{n \leqslant x} |a_n|^2$$

($n \cong b \pmod q$ denoting that $q \|(n - b)$) holds for all complex numbers a_n, for all $x \geqslant 2$. In some average sense, random sequences a_n are well distributed in residue classes to prime powers q all the way up to x. This inequality looks like others, that arose in the theory of the Large Sieve. I shall return to this topic again.

Since $\|\Lambda\|$ is finite, how big is it? An elementary argument, as for example in Elliott [9], shows that, as $x \to \infty$, $\gamma = \limsup \|\Lambda\|^2$ satisfies $\gamma \leqslant 2$, and the consideration of special cases shows that $\gamma \geqslant 1$. The asymptotic value γ has

only recently been obtained (Kubilius [28, 29]; Hildebrand [23]). I first consider the treatment in [28].

For convenience of exposition consider the ratio

$$\tau = (nB^2)^{-1} \sum_{m=1}^{n} \left| f(m) - \sum_{p \leqslant n} f(p)p^{-1} \right|^2,$$

with $B^2 = \sum_{p \leqslant n} f(p)^2 p^{-1}$, which is appropriate to a *strongly-additive* arithmetic function, an additive function which also satisfies $f(p^k) = f(p)$ for all positive k. Following the argument of Turán we expand the square and invert the order of summation, to obtain for τ the representation $Q_1 + Q_2 - Q$, where

$$Q_1 = n^{-1} \sum_{p \leqslant n} \left[\frac{n}{p} \right] px_p^2 - \sum_{p \leqslant n^{1/2}} p^{-1}x_p^2,$$

$$Q_2 = \sum_{p \leqslant n} \sum_{l \leqslant n} \left(\alpha(p)p^{1/2}l^{-1/2} + \alpha(l)l^{1/2}p^{-1/2} \right)x_p x_l$$

$$- \sum_{\substack{pl \leqslant n \\ p \neq l}} \alpha(pl)p^{1/2}l^{1/2}x_p x_l,$$

with $\alpha(v) = v^{-1} - n^{-1}[n/v]$,

$$Q = \sum_{\substack{p \leqslant n \ l \leqslant n \\ pl > n}} p^{-1/2}l^{-1/2}x_p x_l,$$

and for each prime p (or l) $x_p = f(p)p^{-1/2}B^{-1}$, so that

$$\sum_{p \leqslant n} x_p^2 = 1.$$

A straightforward application of the Cauchy-Schwarz inequality gives $Q_2 \ll (\log n)^{-1/2}$, and it is clear that the bound $Q_1 \leqslant 1$ is essentially best possible. The main interest lies in Q.

The coefficient matrix G of Q is symmetric, and represents a self-adjoint operator on the L^2 space $\mathbf{C}^t$, where t denotes the number of primes p not exceeding n. We need an estimate for its smallest eigenvalue. In fact almost its whole spectrum can be obtained. We shall apply the notion that an approximate eigenvector gives rise to an approximate eigenvalue.

LEMMA 1. *Let T be a self-adjoint operator on a complex Hilbert space. For any vector x and complex number α, there is a λ in the closure of the spectrum of T so that*

$$|\lambda - \alpha| \, \|x\| \leqslant \|Tx - \alpha x\|.$$

PROOF. This result is equivalent to the assertion that for a self-adjoint operator T the norm of the resolvent operator $R(\alpha; T)$ does not exceed the reciprocal of the distance of α from the spectrum of T. In his paper Kubilius makes the assertion of the lemma only for a symmetric matrix, giving no reference. For that case a simple proof may be given.

Let I denote the identity map on the same space as T, and let d_j, $j = 1, \ldots, k$, be an orthonormal basis for the space made up of eigenvectors of T. Then

$$(T - \alpha I)x = \sum_{j=1}^{k} (\lambda_j - \alpha) c_j d_j,$$

where λ_j denotes the eigenvalue corresponding to d_j, and c_j is the inner product of x against d_j. Hence

$$\sum_{j=1}^{k} |(\lambda_j - \alpha) c_j|^2 = \|Tx - \alpha x\|^2,$$

and the desired result follows by choosing for λ an eigenvalue which minimizes $|\lambda_j - \alpha|$.

An inductive proof of the existence of the basis d_j (the spectral decomposition theorem for finite spaces) may be found in Lang [30].

There is a companion to this lemma which I shall employ later. It asserts that for isolated eigenvalues an approximate eigenfunction must be near to a genuine eigenfunction.

LEMMA 2. *Let λ be an eigenvalue of the self-adjoint operator T which is at a positive distance δ from the rest of the spectrum. Then for any vector x there is a further vector z, belonging to the eigenspace of T corresponding to λ, so that $\delta\|x - z\| \leqslant \|Tx - \lambda x\|$.*

PROOF. This also follows from the spectral decomposition of the underlying space with respect to T. For finite spaces, and in the notation of the proof of the previous lemma, we have

$$\|(T - \lambda I)x\|^2 = \sum_{j=1}^{k} |\lambda_j - \lambda|^2 |c_j|^2.$$

If now z is the sum of the $c_j d_j$, where j runs through the values for which $\lambda_j = \lambda$, then the right-hand side of this equation is at least as large as $\delta^2\|x - z\|^2$.

A result similar to Lemma 2 I proved as Lemma (5.7) of my book [13], save that I made rather a meal of it—all the more curious since I employed the above simple argument in a typical situation on pages 433 and 434 of that same volume. In this paper I shall employ Lemmas 1 and 2 for compact operators. A straightforward modification of the argument for finite spaces will suffice in this case. A detailed discussion of the general theory of self-adjoint operators is given in Yosida [44].

To obtain suitable candidates for the eigenvectors of G, Kubilius considers the problem of estimating

$$V(\psi) = \frac{\iint \psi(u)\psi(v) u^{-1} v^{-1} \, du \, dv}{\int_0^1 \psi(u)^2 u^{-1} \, du},$$

where the double integral is taken over the triangle $0 \leqslant u \leqslant 1$, $0 \leqslant v \leqslant 1$, $u + v > 1$. This can be viewed as an extremal problem involving functions of the class L^2 with respect to the measure $u^{-1} du$ on $0 \leqslant u \leqslant 1$. That it is appropriate is not immediately clear, and I shall discuss this presently.

By considering $V(\psi + yh)$ for a temporarily fixed h, and varying the real numbers y, Kubilius shows that any function ψ which gives an extremal value λ of $V(\psi)$ must satisfy

$$\int_{1-u}^{1} \psi(v) v^{-1} \, dv = \lambda \psi(u)$$

on $0 \leqslant u \leqslant 1$. It must be an eigenvalue of the operator

$$(10) \qquad \psi(u) \mapsto \int_{1-u}^{1} \psi(v) v^{-1} \, dv.$$

Let us call this operator T; it will appear many times in what follows.

The eigenfunctions of T will satisfy the relation

$$\lambda(1-u)\psi'(u) = \psi(1-u)$$

where $'$ denotes differentiation with respect to u, and hence the second-order differential equation

$$(11) \qquad u(1-u)\phi''(u) - u\phi'(u) + \lambda^{-2}\phi(u) = 0.$$

This is an equation of hypergeometric type, and solutions

$$(12) \qquad \phi_r(u) = u \sum_{k=0}^{r-1} (-1)^{r-1-k} \binom{r}{k} \binom{r-1}{k} u^{r-k-1} (1-u)^k$$

with $\lambda_r = (-1)^{r-1} r^{-1}$ are exhibited, one for each positive integer r. In terms of the standard notation for Jacobi polynomials,

$$\phi_r(u) = u P_{r-1}^{(1,0)}(1 - 2u).$$

See, for example, Rainville [34, Chapter 6]. In particular

$$\int_0^1 \phi_r(v)\phi_s(v) v^{-1} \, dv = \begin{cases} 0 & \text{if } r \neq s, \\ 1/2r & \text{if } r = s. \end{cases}$$

It will be convenient to use $\psi_r(u)$ to denote the (renormalized) function $(2r)^{1/2}\phi_r(u)$.

Let z_r be the vector in $\mathbf{C}^t$ with p-component $p^{-1/2}\psi_r(\log p/\log n)$. A direct computation shows that

$$|Gz_r - \lambda_r z_r| \ll (\log n)^{-1/2},$$

which by Lemma 1 guarantees the existence of an eigenvalue ν_r of G in the range $|\nu_r - \lambda_r| \ll (\log n)^{-1/2}$. For $r \leqslant c_1 (\log n)^{1/4}$ and a suitably chosen constant c_1, the eigenvalues for differing r will be distinct. We do not as yet know whether any occur multiply for G.

Since G is symmetric,

$$\sum_{r=1}^{t} \nu_r^2 = \text{trace}(G^2),$$

and this trace Kubilius estimates, in a straightforward manner, to be $\zeta(2) + O(\log\log n/\log n)$ in terms of the Riemann zeta function $\zeta(k) = \sum_{r=1}^{\infty} 1/r^k$. Since demonstrably

$$\sum_{r=1}^{\infty} \lambda_r^2 = \zeta(2),$$

we see that all eigenvalues of G with $r > c_1(\log n)^{1/4}$ are $O((\log n)^{-1/4})$ in size. Their effect in G is small.

The highest eigenvalue of $-G$ is thus $-\nu_2$, giving

$$\tau \leqslant 3/2 + O\big((\log n)^{-1/2}\big).$$

Moreover, with $x = z_2$ we can obtain for τ a similar estimate in the other direction. This gives for the maximal value of τ the estimate $3/2 + O((\log n)^{-1/2})$, which is the main result of Kubilius' paper.

A second method is sketched, as follows. Direct computation shows that for each positive integer k

$$\operatorname{tr}(G^k) = \sum (p_1 \cdots p_k)^{-1},$$

where the summation is taken over those k-tuples of primes that satisfy $p_j \leqslant n$ for every j, $p_i p_{i+1} > n$ for $1 \leqslant i \leqslant k-1$, $p_k p_1 > n$. This sum in turn is closely approximated by

$$(13) \qquad \int (u_1 \cdots u_k)^{-1} du_1 \cdots du_k$$

taken over the region $0 \leqslant u_j \leqslant 1$ for every j, $u_i + u_{i+1} > 1$ for $1 \leqslant i \leqslant k-1$, and $u_k + u_1 > 1$.

Defining

$$\theta(u_k, u_1) = \begin{cases} 1 & \text{if } u_k > 1 - u_1, \\ 0 & \text{otherwise,} \end{cases}$$

this last integral may be written in the form

$$\int_0^1 u_1^{-1} du_1 \int_{1-u_1}^1 u_2^{-1} du_2 \cdots \int_{1-u_{k-1}}^1 \theta(u_k, u_1) u_k^{-1} du_k.$$

Expanding the function θ in terms of the eigenfunctions ϕ_r, say

$$\theta(u, u_1) \sim \sum_{r=1}^{\infty} a_r \phi_r(u),$$

we see that

$$a_r = 2r \int_{1-u_1}^1 \phi_r(u) u^{-1} du = 2r\lambda_r \phi_r(u_1),$$

and readily obtain the estimate

$$(14) \qquad \operatorname{tr}(G^k) = \sum_{r=1}^{\infty} (-1)^{(r-1)k} r^{-k} + o(1),$$

valid as $n \to \infty$.

Since G is symmetric,

$$\sum_{r=1}^{t} \nu_r^k = \text{tr}(G^k) = \zeta(k) + o(1)$$

for even integers k. Assuming that $|\nu_r|$ does not increase with r, we deduce the estimate

$$\limsup_{n \to \infty} |\nu_1| \leqslant \liminf_{k \to \infty} \zeta(k)^{1/k} = 1.$$

Indeed, for $x = z_1$, Q has the value $1 + O((\log n)^{-1/2})$, so that $\nu_1 = 1 + o(1)$.

Removing ν_1^k from the trace gives

$$\nu_2^k \leqslant \zeta(k) - 1 + o(1)$$

for every even k, and so

$$\limsup_{n \to \infty} |\nu_2| \leqslant \liminf_{k \to \infty} (\zeta(k) - 1)^{1/k} = 1/2.$$

Since for $x = z_2$, Q has the value $-1/2 + O((\log n)^{-1/2})$, we get $\nu_2 = -1/2 + o(1)$, and so on.

This method leads to the value $\gamma = 3/2$ but does not seem as well able to give a quantitative estimate for τ in terms of n.

It seems that these two arguments were devised in the opposite order. Improving upon his earlier work [27], Kubilius, in a lecture given at Oberwolfach, Germany, in November 1980, reduced the problem of estimating τ to the evaluation of the multiple integral (13).

At that meeting H. Montgomery showed that the polynomials ϕ_r were eigenfunctions of the operator T, with eigenvalues λ_r. Once we decide to look for polynomial eigenfunctions of the operator T at (10) it is clear that those of degree r would have the corresponding eigenvalue λ_r. The ψ_r may then be found by direct substitution. However, for the asymptotic estimation of trace(G^k) it is sufficient that such polynomials, of arbitrary degree, should exist. This, as was pointed out by Wirsing in a letter to Montgomery in early January 1981, may be obtained by noting that the linear operator $T - \lambda_r I$ takes the finite-dimensional space of polynomials of degree at most r and vanishing at $u = 0$ into the similar space where the degree is at most $r - 1$, a space of smaller dimension. In the same letter the relation $\sum_{r=1}^{\infty} \lambda_r^2 = \zeta(2)$ is already employed.

There was an exchange of letters between Montgomery and Kubilius, and at some stage the alternative method, presented preferentially in Kubilius' paper, was developed. This gives more directly estimates for the eigenvalues of the operator under consideration, so that they are sharper. For this argument to succeed when applied to a discrete operator it is very helpful that the eigenvalues decrease steadily in size, because no analogue of Weierstrass' approximation of a continuous function by polynomials, with which to gain completeness, is available.

In a later section I shall show that when not working in a Hilbert space it is advantageous to combine the two methods. The use of iterations to determine

the spectral radius of an operator is, of course, well known. To put this into practice is another matter. Thus in connection with the Large Sieve I suggested in [4], 1971, that one might iterate the operator. However, even though one could formulate an appropriate problem in terms of gamma functions, which have convenient convolutions, I balked at the complications; and later developments have not used this method at all.

The following background remarks may be helpful. In treating sums involving primes it is convenient to apply the prime number theorem in the form

$$\sum_{p \leqslant x} 1 = \int_2^x \frac{dt}{\log t} + O\left(x \exp\left(-c\sqrt{\log x}\right)\right)$$

with $c > 0$ (see, for example, Davenport [2], Prachar [33]). Then by partial summation

$$\sum_{p \leqslant x} \frac{1}{p} = \log \log x + c_2 + O\left(\exp\left(-c_3\sqrt{\log x}\right)\right)$$

with constants $c_2, c_3 > 0$, uniformly for $x \geqslant 2$. This enables sums over ranges $x/l < p \leqslant x$ to be well approximated by continuous functions provided l is not too near to x. When l is near to x, upper bounds are usually sufficient. Similar results hold when the summations are over prime powers, rather than primes.

Let W be the space of complex-valued functions $h(u) = ug(u)$ where $g(u)$ is continuous on the interval $0 \leqslant u \leqslant 1$, and define

$$\|h\| = \sup_{0 \leqslant u \leqslant 1} |g(u)|.$$

This space is complete. It is mapped into itself by the transformation T which takes h to the function

$$u \mapsto \int_{1-u}^1 \frac{h(v)}{v}\, dv, \qquad 0 \leqslant u \leqslant 1.$$

This transformation is bounded on W, of norm 1.

Let H be the Hilbert space of complex-valued Lebesgue measurable functions $k(u)$, defined on the unit interval, with

$$\int_0^1 \frac{|k(u)|^2}{u}\, du < \infty.$$

We make the usual identification of functions differing on a set of measure zero. The inner product on H is given by

$$[k, t] = \int_0^1 \frac{k(u)\,\overline{t(u)}}{u}\, du.$$

The points of W are everywhere dense in H (in the topology of H), and an application of Fubini's theorem shows that T is a self-adjoint operator in H. An application of the Cauchy-Schwarz inequality shows that it is bounded.

In view of the well-known theorem of Weierstrass, we see that the polynomials $\psi_r(u)$ span W, and therefore H. In particular, for every nonzero complex

number ρ, and member x of H,

$$\|(T - \rho I)x\| \geqslant \|x\| \min_r |\lambda_r - \rho|.$$

This shows that the spectrum of T contains only the eigenvalues λ_r, and possibly the real number zero.

If $Tg = 0$ in H, then by differentiation of Tg, g is seen to be zero also. Thus zero is not an eigenvalue of T. However, the range of T is not the whole of H, since there are functions in H which are not absolutely continuous. Thus zero does belong to the spectrum of H. Of course T is a compact operator, of Hilbert-Schmidt type.

A feeling for the size of the polynomials $\psi_j(u)$ may be obtained by evaluating the integral

$$\int_{1-v}^1 \theta(u, v) \, \frac{du}{u},$$

once employing the spectral decomposition of $\theta(u, v)$ by T, and once directly. This leads to the representation

$$-\log(1 - v) = \sum_{j=1}^\infty \frac{\psi_j(v)^2}{j^2},$$

valid in the interval $0 \leqslant v < 1$.

Why are continuous analogues on the (product of) interval(s) $(0, 1)$, such as $V(\psi)$ and (10), reasonable?

The logarithmic function appears first in probabilistic number theory in Erdős' characterization of *finitely distributed* additive functions [17], 1946. There it is introduced to make the proof work, and not for any background philosophical reason. He applies his characterization to show that if a real-valued additive function f is monotone, or satisfies $f(n + 1) - f(n) \to 0$ as $n \to \infty$, then it must be a constant multiple of $\log n$. Erdős's arguments were rather special, and at the end of his paper he conjectured that this last conclusion could be drawn from the hypothesis that

$$\lim_{x \to \infty} x^{-1} \sum_{n \leqslant x} |f(n + 1) - f(n)| = 0.$$

Moreover, any additive function f for which $f(n + 1) - f(n)$ was bounded above should differ by a bounded function from a multiple of $\log n$.

These conjectures were proved valid by Kátai [24] and Wirsing [42], and by Wirsing [41], respectively, and largely regarded as giving characterizations of the function $\log n$. Although ingenious, the proofs were again of a special nature. Besides this, the application of Halász's method in the theory of Dirichlet series [20] made possible a better understanding of the limiting behavior of translations of additive functions (e.g., Elliott and Ryavec [16], Levin and Timofeev [31]), and so of finitely distributed functions.

Altogether one gained the diffuse notion that any real additive function which is sufficiently regular must in some sense be near to a multiple of a logarithm. To make this precise in any given situation one needed a source for

the logarithm, which in the above works was obtained indirectly. More recent experience in probabilistic number theory helps (Elliott [7]):

Let measurable functions $\alpha(x), \beta(x) > 0$ exist so that as $x \to \infty$

$$(15) \qquad \nu_x(n; f(n) - \alpha(x) \leqslant z\beta(x)) \to \begin{cases} 1 & \text{if } z > 0, \\ 0 & \text{if } z < 0; \end{cases}$$

the frequencies converge weakly to the improper law with jump at the origin. Suppose further that $\beta(x) \to \infty$ as $x \to \infty$, and that $\limsup \beta(x^y)/\beta(x)$ is finite for each positive y. Then the function $\alpha(x)$ has a decomposition $\alpha(x) = \alpha_1(x) + \alpha_2(x)$ with respect to the group of substitutions $x \mapsto x^y$, so that $\alpha_1(x^y) = y\alpha_1(x) + o(\beta(x))$, and $\alpha_2(x^y) = \alpha_2(x) + o(\beta(x))$, as $x \to \infty$. Moreover, if $\limsup \beta(x^2)/\beta(x) < 2$, this ensures that $\alpha_1(x)$ has the form $A \log x + o(\beta(x))$ for some constant A. In this sense, and in a more exact sense in that same paper, f is near to a logarithm.

During the considerations of that paper the function $\Lambda(a, b) = \alpha(a) - \alpha(a/b) - \alpha(b)$ was studied on the direct sum of four copies of $\mathbb{R}^*$, the multiplicative group of the positive real numbers. It proved most natural to investigate the size of $\Lambda(a, b) - \Lambda(a_1, b_1)$ in terms of the renormalized Haar measure μ_1 on $\mathbb{R}^*$ induced by taking the Lebesgue measure of the image under the map $a \mapsto (\log a)/\log x$. In particular this transforms the range $1 \leqslant n \leqslant x$ into the interval $0 \leqslant u \leqslant 1$, a situation which is implicitly reproduced in Kubilius' paper.

Since the argument in my paper [7] was implicitly concerned with bounding a resolvent operator in measure, cannot the determination of the spectrum of the discrete analogue of T say something about the implications for $f(q)$ on prime powers q of a known value distribution of an additive function $f(n)$ on the integers n, at least in an L^2 sense? Indeed it can, as I presently show.

Another proof that $\gamma = 3/2$ was given by Hildebrand [23]. In his paper he gives a kind of asymptotic formula for the analogue $S(\mathbf{x})$ of τ,

$$(16) \qquad S(\mathbf{x}) = n^{-1} \sum_{m=1}^{n} \left| f(m) - n^{-1} \sum_{m=1}^{n} f(m) \right|^2,$$

where $x_q = f(q)$, $q \leqslant n$, dealing at once with a general additive function. He introduces the polynomials $p_k(u)$, which are orthogonal over the interval $0 \leqslant u \leqslant 1$ with weight u. Thus $up_k(u) = \psi_k(u)$. He then estimates $S(\mathbf{x})$ by reducing the problem to the minimization of an appropriate quadratic form, implicitly constructing a spectral decomposition of $f(q)$. As in the integral method above, the estimations do not lend themselves to the derivation of a sharp local bound for $S(\mathbf{x})$. I give a view of his result in [13, pp. 430–432]. Here I shall employ only the following lemma.

LEMMA 3. *Let* $\mathbf{y}$ *be the vector with q-coordinate* $2^{1/2} \log q / \log n$, $q \leqslant n$. *Then for any* μ

$$S(\mathbf{x} + \mu\mathbf{y}) = S(\mathbf{x}) + \frac{\mu^2(1 + o(1))}{(\log n)^2} + o\left(\frac{|\mu| B(n)}{\log n} \right)$$

as $n \to \infty$.

PROOF. Following Hildebrand we can obtain this directly, employing the good distribution into residue classes (mod q) of the logarithmic function.

Since we are essentially evaluating an inner product, the cross-terms typified by

$$U = 2^{1/2}\mu(n\log n)^{-1}\sum_{m=1}^{n}\left(f(m) - n^{-1}\sum_{m=1}^{n}f(m)\right)\left(\log m - n^{-1}\sum_{m=1}^{n}\log m\right)$$

are the most interesting. Here $\log m$ is translated by its mean, so that the innermost average of f may be omitted. Representing $f(m)$ as the sum of the $f(q)$ with $q\|m$, and inverting the order of summation over m and q, gives U the alternative form

$$2^{1/2}\mu(n\log n)^{-1}\sum_{q\leqslant n}f(q)\eta_q$$

with

$$\eta_q = \sum_{\substack{m=1 \\ m\equiv 0\,(\mathrm{mod}\,q)}}^{n}\left(\log m - n^{-1}\sum_{m=1}^{n}\log m\right).$$

The coefficients η_q may be estimated individually, either using several integrations by parts as in Hildebrand [23, pp. 163–165], or by applications of Stirling's approximation for $n!$. Altogether they satisfy

$$\sum_{q\leqslant n}q\eta_q^2 \ll \frac{(n\log\log n)^2}{\log n}.$$

An appropriate application of the Cauchy-Schwarz inequality now yields $U = o(|\mu|B(n)/\log n)$. The proof of Lemma 3 is readily completed.

3. Reappraisal. Before reappraising these results it is convenient to make a few more remarks of a functional analytic nature.

Let X be a complex Hilbert space, with inner product $[\,,\,]$. Then we can identify X with its dual space X' by the map J which takes the vector x to the linear map

$$u\mapsto [u,x], \qquad u\in X.$$

This map preserves norms but is not quite linear since, for a complex number α, $J(\alpha x) = \bar{\alpha}J(x)$.

Suppose further that Y is a Hilbert space, and let K be its identification with its dual space Y'. The *adjoint* of a linear map $L\colon X\to Y$ is then defined to be the map $L^*\colon Y\to X$ given by

$$L^* = J^{-1}L'K,$$

where L' is the map $Y'\to X'$ dual to L. This ensures that

$$[Lx, y] = [x, L^*y]$$

for all x in X, y in Y, where the inner products are defined on X and Y, respectively. Then $(L^*)^* = L$.

It is readily checked that for each x in X, $\|Lx\|^2 = [x, L^*Lx]$, so that

$$\|L\|^2 = \sup_{\|x\|=1} |[x, L^*Lx]|.$$

Similarly, for each y in Y, $\|L^*y\|^2 = [y, LL^*y]$, from which

$$\|L'\|^2 = \sup_{\|y\|=1} |[y, LL^*y]|.$$

The maps L^*L and LL^* are self-adjoint on the spaces X, Y respectively, and for a bounded operator L their spectra are contained in a bounded interval of nonnegative numbers.

Let x denote a positive real number, $x \geqslant 2$.

We allow the values of an additive function $f(m) = \Sigma_{q\|m} f(q)$ to be complex, and regard the $f(q)$, $q \leqslant x$, as determining a vector in the complex space $\mathbb{C}^s$, where s denotes the number of prime powers q not exceeding x. It is convenient to topologize this space with the norm

$$\|z\| = \left(\sum_{q \leqslant x} \frac{1}{q}\left(1 - \frac{1}{q_0}\right)|z_q|^2 \right)^{1/2}$$

and to regard it as the space of functions on a set of s points. These functions are viewed as being of class L^2 with respect to the measure $\widehat{du}$ that assigns a weight $q^{-1}(1 - q_0^{-1})$ to the point corresponding to the qth coordinate.

Likewise, $\mathbb{C}^{[x]}$ will be topologized with the norm

$$\|y\| = \left(\frac{1}{[x]} \sum_{n \leqslant x} |y_n|^2 \right)^{1/2},$$

and viewed as an L^2 space on a set of $[x]$ points, with the frequency measure assigning a weight $[x]^{-1}$ to each point corresponding to an integer n not exceeding x.

In this notation the Turán-Kubilius inequality (7) asserts that the map A_2: $L^2(\mathbb{C}^s) \to L^2(\mathbb{C}^{[x]})$ given by

(17) $$(A_2 f)(n) = f(n) - M(f),$$

where

(18) $$M(f) = \sum_{l \leqslant x} \frac{f(l)}{l}\left(1 - \frac{1}{l_0}\right)$$

and l runs through prime powers, has a norm bounded independently of x. The peculiar notation A_2 is for future purposes. Determining the best constant in the Turán-Kubilius inequality is equivalent to finding the spectral radius of $A_2^* A_2$.

The dual operator A_2' is given by

$$(A_2' \mathbf{a})(q) = \frac{q}{(1 - 1/q_0)[x]}\left(\sum_{\substack{n \leqslant x \\ n \equiv 0 \,(\text{mod } q)}} a_n - \frac{1}{q}\left(1 - \frac{1}{q_0}\right)\sum_{n \leqslant x} a_n \right).$$

A direct computation shows that

$$A_2^* A_2 = I - T + N_1 + N_2,$$

where T is given by

(19) $$(Tf)(q) = \sum_{x/q < l \leqslant x} \frac{f(l)}{l}\left(1 - \frac{1}{l_0}\right),$$

l continuing to denote a prime power,

$$(N_1 f)(q) = -\sum_{\substack{l \leqslant x/q \\ (l,q) > 1}} \frac{f(l)}{l} + \left(\frac{qq_0}{(q_0 - 1)[x]}\left(\left[\frac{x}{q}\right] - \left[\frac{x}{qq_0}\right]\right) - 1\right) f(q),$$

and the q-coordinate of $N_2 f$ is

$$\ll \frac{q}{x} \sum_{\substack{l \leqslant x/q \\ (l,q) = 1}} |f(l)| + \frac{1}{x} \sum_{l \leqslant x} |f(l)|.$$

In particular, after a number of applications of the Cauchy-Schwarz inequality employing the upper bound $s \ll x/\log x$, $\|N_2\| \ll (\log\log x/\log x)^{1/2}$. When aiming for the value of $\|A_2^* A_2\|$ it is better to deal with $A_2^* A_2$ directly, thus avoiding the extra factor $(\log\log x)^{1/2}$. However, the form given here demonstrates clearly the similarity between the discrete operator T and the continuous operator T on the Hilbert space $H = L^2(0, 1)$ of functions measurable with respect to $u^{-1} du$, given by

$$h \mapsto \int_{1-u}^{1} \frac{h(v)}{v}\, dv.$$

I shall continue to use the same notation for these operators, distinguishing between them by the adjectives discrete and continuous, or by T_d, T_c respectively.

In practice the operator N_1 may often be removed by considering $\check{A}_2$, a variant of A_2 which is most simply defined through its adjoint. The operator $\check{A}_2^* : \mathbf{C}^{[x]} \to \mathbf{C}^s$ given by

(20) $$\left(\check{A}_2^* \mathbf{a}\right)(q) = \frac{1}{[x/q] - [x/qq_0]}\left(\sum_{\substack{n \leqslant x \\ n \equiv 0 \,(\mathrm{mod}\, q)}} a_n - \frac{1}{q} \sum_{\substack{n \leqslant x \\ (n,q) = 1}} a_n\right)$$

has bounded norm, since it largely differs from A_2^* by an operator the norm of whose effect may be directly estimated, using the Cauchy-Schwarz inequality, not to exceed

$$\left(\sum_{q \leqslant x} \frac{1}{q}\left(1 - \frac{1}{q_0}\right) \frac{1}{[x]^2} \sum_{\substack{n \leqslant x \\ (n,q) > 1}} |a_n|^2 \sum_{\substack{n \leqslant x \\ (n,q) > 1}} 1\right)^{1/2} \ll \|\mathbf{a}\|\left(\sum_{q \leqslant x} \frac{1}{qq_0}\right)^{1/2} \ll \|\mathbf{a}\|,$$

for each $\mathbf{a}$ in $\mathbf{C}^{[x]}$. Then the composition $\check{A}_2^* A_2$ has a representation $I - T + N$ where $\|N\| \ll (\log\log x/\log x)^{1/2}$.

It is possible to give an explicit treatment of the notion that as $x \to \infty$ the discrete T approaches the continuous version of T. Let d_j, $j = 1, \ldots, s$, be a basis for $\mathbf{C}^s$ comprised of unit eigenvectors of T, with corresponding eigenvalues ν_j so that $|\nu_j|$ does not increase with j. By using the approximate eigenfunctions $\psi_j(\log q / \log x)$ and Lemma 1 one obtains an estimate $|\nu_j - \lambda_j| \ll c^j(\log x)^{-1/2}$ with some constant $c > 1$, uniformly in j, x. Once again

$$\sum_{j=1}^{s} \nu_j^2 = \operatorname{trace} T^2 = \zeta(2) + O(\log\log x / \log x).$$

From this one readily deduces the estimate $|\nu_j - \lambda_j| \ll (\log\log 4x)^{-1/2}$, uniformly for $1 \leqslant j \leqslant s$, $x \geqslant 2$.

We inject $\mathbf{C}^s$ into the Hilbert space $L^2(0,1)$ by

$$D: \sum_{j=1}^{s} c_j d_j \mapsto \sum_{j=1}^{s} c_j \psi_j(u).$$

This map depends upon the choice of the eigenvectors d_j when $j \gg (\log x)^{1/4}$, but since the corresponding eigenvalues are small, this has only a small effect upon T. If P_s is the projection of H onto the space spanned by the first s eigenfunctions $\psi_j(u)$, then for y in H we have

$$\left\| T_c P_s y - D T_d D^{-1} y \right\|^2 = \sum_{j=1}^{s} (\lambda_j - \nu_j)^2 c_j^2 \ll \|y\|^2 (\log\log 4x)^{-1}.$$

Moreover, $\|T_c - T_c P_s\| \ll s^{-1} \ll x^{-1} \log x$, and we see that

$$\left\| D T_d D^{-1} - T_c \right\| \ll (\log\log 4x)^{-1}.$$

In this sense we gain a quantitative representation of the convergence of T_d to T_c. In practice it is sometimes sharper to mimic in $L^2(\mathbf{C}^s)$ a proof that is valid in H, rather than invoke directly this last result.

As an example in the application of the foregoing analysis I establish the following result, first proved by Ruzsa [37] using an altogether different method.

THEOREM 4.

$$V(f) \asymp \min_{\lambda}\left(|\lambda|^2 + \sum_{q \leqslant x} \frac{1}{q} |f(q) - \lambda \log q|^2 \right)$$

uniformly in f, $x \geqslant 2$.

Here $V(f)$ is the same as $S(x)$ defined at (16) but with n replaced by $[x]$, and $u \asymp v$ denotes that there are positive constants c_1, c_2 so that $c_1 u \leqslant v \leqslant c_2 u$.

PROOF. The upper bound follows readily from the Turán-Kubilius inequality, if we note that

$$\frac{1}{[x]} \sum_{m \leqslant x} f(m) = \sum_{q \leqslant x} \frac{f(q)}{q}\left(1 - \frac{1}{q_0}\right) + O\left(\left(\frac{1}{\log x} \sum_{q \leqslant x} \frac{|f(q)|^2}{q} \right)^{1/2} \right),$$

and that

$$\frac{1}{x} \sum_{n \leqslant x} \left(\log \frac{n}{x} \right)^2$$

is bounded uniformly for $x \geqslant 2$.

The lower bound for $V(f)$ requires more effort. Adopting the philosophy: *if operator ↔ sufficiency, then dual of operator ↔ necessity*, we consider the composition $A_2^* A_2$. In fact

$$\| \check{A}_2^* A_2 f \| \ll V(f) + (\log x)^{-1/2} \| f \|.$$

In view of our considerations in the reappraisal we have

$$\| (I - T) f \| \ll V(f) + (\log \log x / \log x)^{1/2} \| f \|,$$

with the norms in $\mathbb{C}^s$.

It follows from Lemma 2 that with $F = (f, d_1)$,

$$\| f - F d_1 \| \ll V(f) + (\log \log x / \log x)^{1/2} \| f \|,$$

where (as earlier) d_1 is the unit eigenvector of the discrete T corresponding to the eigenvalue ν_1.

Let $\hat{\psi}_j$ denote the function $\psi_j (\log q / \log x)$ rescaled to have the norm 1. Since

$$\left\| (T - \nu_1 I) \psi_1 \left(\frac{\log q}{\log x} \right) \right\| \ll (\log x)^{-1/2},$$

we see from another application of Lemma 2 that

$$\left\| \psi_1 \left(\frac{\log q}{\log x} \right) - (\psi_1, d_1) d_1 \right\| \ll (\log x)^{-1/2}.$$

Moreover, $\| d_1 \| = 1 = \| \hat{\psi}_1 \|$, so that in view of the general norm inequality $| \|a\| - \|b\| | \leqslant \|a - b\|$, we similarly derive $\| \hat{\psi}_1 - d_1 \| \ll (\log x)^{-1/2}$. Thus with $F_0 = (f, \hat{\psi}_1)$ we have

$$\| f - F_0 \hat{\psi}_1 \| \ll V(f) + (\log \log x)^{-1/2} \| f \|.$$

If we replace f in this inequality by $f - F_0 \hat{\psi}_1$, then F_0 is replaced by zero, and we obtain

$$\| f - F_0 \hat{\psi}_1 \| \ll V(f - F_0 \hat{\psi}_1).$$

This puts us into a position to apply Lemma 3, and so deduce that

$$\lambda^2 + \sum_{q \leqslant x} \frac{1}{q} | f(q) - \lambda \log q |^2 \ll V(f)$$

with $\lambda = \sqrt{2} (f, \hat{\psi}_1)(\log x)^{-1}$.

For this argument to succeed it would be enough to know that the eigenvalue ν_1, which is near to 1, is bounded away from the rest of the spectrum of T, uniformly for all large values of x. For later reference we note that a bound $\| \hat{\psi}_j - d_j \| \ll (\log x)^{-1/2}$ holds for each fixed j.

I should add that I first applied to the study of additive (and multiplicative) arithmetic functions the above philosophy concerning duality and necessity in [5], 1972, and in connection with the Turán-Kubilius inequality. Duality between point and line is a central feature in projective geometry. There the philosophy is exemplified (possibly I repeat myself) by Desargues' theorem, whose dual is already its converse. At that time I consciously sought an alternative to the application of Kac's idea that being divisible by differing primes represented independent events, since I knew it to be not generally valid. Seen from this distance it is clear that a particular outcome of the above philosophy is that the notion of independence has been replaced by that of self-adjointness of an operator.

Whilst this approach gives a unifying view of probabilistic number theory, many of the important theorems of that theory are susceptible to treatments derived from different motivations. This is not the case in the study of the differences of additive functions, where the use of functional analysis within the above philosophy leads to results which, at present, are the most precise and the widest in application.

4. Differences of additive functions. An analogue of the Turán-Kubilius inequality for differences is readily obtained by interpolating $M(f)$:

$$\sum_{n+1 \leqslant x} |f(n+1) - f(n)|^2$$

$$\leqslant 2 \sum_{n+1 \leqslant x} |f(n+1) - M(f)|^2 + 2 \sum_{n+1 \leqslant x} |f(n) - M(f)|^2$$

$$\ll xB(x)^2.$$

Likewise we obtain the analogue of the upper bound in Theorem 4:

$$x^{-1} \sum_{n+1 \leqslant x} |f(n+1) - f(n)|^2 \ll x^{-1}|\lambda|^2 + \sum_{q \leqslant x} q^{-1}|f(q) - \lambda \log q|^2,$$

valid for each λ. What about the lower bound? This is of particular interest when considering applications of arithmetic functions to the representation of rationals by products. For simplicity of exposition, and to show more clearly the dependence of localization upon the size of various parameters, I shall consider only the differences $f(n+1) - f(n)$.

Let $B: \mathbb{C}^s \to \mathbb{C}^{[x]}$ be the operator given by

$$(Bf)(n) = f(n-1) - M(f),$$

with the convention that $f(0) = 0$. Thus B is bounded uniformly for $x \geqslant 2$. According to the philosophy employed in the previous section we consider the composition $\check{A}_2^*(A_2 - B)$. Then

$$\|\check{A}_2^*(A_2 - B)f\| \ll \|(A_2 - B)f\|,$$

where

$$\|(A_2 - B)f\| = \left([x]^{-1} \sum_{n \leqslant x} |f(n) - f(n-1)|^2\right)^{1/2}.$$

If we could somehow remove the operator $\check{A}_2^* B$ we should be in a position analogous to that of the previous section, reduced to a study of a discrete version of T. It is required that in some sense $\check{A}_2^*$ (or A_2^*) be orthogonal to B. Note that $\|\check{A}_2^* Bf\|^2$ is equal to

$$
\sum_{q \leqslant x} q\left(1 - \frac{1}{q_0}\right)^{-1} \left| \frac{1}{[x]} \sum_{\substack{n \leqslant x-1 \\ n \equiv 1 \,(\mathrm{mod}\, q)}} f(n) - \frac{1}{q}\left(1 - \frac{1}{q_0}\right)\frac{1}{[x]} \sum_{\substack{n \leqslant x-1 \\ (n-1, q)=1}} f(n) \right|^2.
$$

Defining

$$
E(y, D, r) = \sum_{\substack{n \leqslant y \\ n \equiv r \,(\mathrm{mod}\, D)}} f(n) - \frac{1}{\phi(D)} \sum_{\substack{n \leqslant y \\ (n, D)=1}} f(n)
$$

for integers $r, D > 0$, and real y, I can prove

THEOREM 5. *For each fixed $\varepsilon > 0$*

$$
\sum_{q \leqslant x^{1/2 - \varepsilon}} \phi(q) \max_{(r, q)=1} \max_{y \leqslant x} |E(y, q, r)|^2 \ll \frac{x^2 (\log \log x)^4}{\log x} \|f\|^2,
$$

where

$$
\|f\| = \left(\sum_{q \leqslant x} \frac{1}{q}\left(1 - \frac{1}{q_0}\right)|f(q)|^2 \right)^{1/2}.
$$

Moreover, a similar result holds if $n \equiv r \pmod q$ is replaced by $n \cong r \pmod q$, and $\phi(q)^{-1}$ is replaced by q^{-1} in the definition of $E(y, q, r)$. If for integers d and $c > 1$ the outer summation is restricted to those q which are prime to c, then $n \equiv d \pmod c$ may be required in the sums defining $E(y, q, r)$.

Perhaps a result of this kind holds with $x^{1/2 - \varepsilon}$ replaced by $x^{1 - \varepsilon}$, or better. With $q = 3$, applications of the Cauchy-Schwarz and then the Turán-Kubilius inequalities yield

$$
|E(x, q, r)|^2 \ll x \sum_{\substack{n \leqslant x \\ n \equiv r \,(\mathrm{mod}\, 3)}} |f(n) - M(f)|^2 + x \sum_{\substack{n \leqslant x \\ (n-1, 3)=1}} |M(f) - f(n)|^2
$$

$$
\ll x^2 \|f\|^2,
$$

so that for one modulus the inequality saves about $(\log x)^{-1}$.

If we define

$$
F(y, D, r) = \sum_{\substack{p \leqslant y \\ p \equiv r \,(\mathrm{mod}\, D)}} 1 - \frac{1}{\phi(D)} \sum_{p \leqslant y} 1,
$$

then the Bombieri-Vinogradov theorem [1] asserts that for each fixed $c_1 > 0$ and appropriate $c_2 > 0$

$$
\sum_{D \leqslant x^{1/2}(\log x)^{-c_2}} \max_{(r, D)=1} \max_{y \leqslant x} |F(y, D, r)| \ll x(\log x)^{-c_1}.
$$

Employing the Brun-Titchmarsh bound $F(y, D, r) \ll y(\phi(D)\log y)^{-1}$ which,

given $c_0 > 0$, is valid for $(r, D) = 1$, $2 \leqslant D \leqslant y^{1-c_0}$, we may deduce from the Bombieri-Vinogradov theorem that

$$(21) \qquad \sum_{D \leqslant x^{1/2}(\log x)^{-c_2}} \phi(D) \max_{(r,D)=1} \max_{y \leqslant x} |F(y, D, r)|^2 \ll x^2 (\log x)^{-c_1-1}.$$

In Theorem 5 the primes p are replaced by an arbitrary additive function $f(n)$, and this demands a qualitative change in the argument. The traditional applications of the Large Sieve to the study of primes in arithmetic progressions employ at some stage an L^1 argument. Here we must work in an L^2 sense from the outset, and so be careful with our estimates of the spectra of the underlying operators.

Let

$$E_1(y, k, r) = \sum_{\substack{n \leqslant y \\ n \equiv d \,(\mathrm{mod}\ c) \\ n \equiv r \,(\mathrm{mod}\ k)}} \frac{f(n)}{n^\sigma}\left(1 - \frac{n}{y}\right) - \frac{1}{\phi(k)} \sum_{\substack{n \leqslant y \\ n \equiv d \,(\mathrm{mod}\ c) \\ (n,k)=1}} \frac{f(n)}{n^\sigma}\left(1 - \frac{n}{y}\right)$$

where c, k are positive integers, d, r are integers, and σ, y are positive real numbers. In Chapter 7 of [13] I proved, as Theorem (7.1), that for each σ in the interval $1/2 < \sigma < 1$

$$(22) \qquad \sum_{\substack{\log x < p \leqslant Q \\ (p,c)=1}} (p-1) \max_{y \leqslant x} |E_1(y, p, r)|^2 \ll \left(\frac{x^{1-\sigma}}{\log x} + Q^\mu\right) \sum_{q \leqslant x} \frac{|f(q)|^2}{q^\sigma},$$

uniformly for all additive functions $f(n)$, for all $x \geqslant 2$ and $Q \geqslant 1$, with μ a certain function of σ. The presence of the conditions involving c, d is unimportant here. For Q a small enough power of x, the first factor in this upper bound is $\ll x^{1-\sigma}(\log x)^{-1}$. In comparison with the inequality (21) derived from the Bombieri-Vinogradov theorem we have saved only $\log x$ rather than an essentially arbitrary power $(\log x)^{c_1+1}$. In this generality, however, (22) is best possible. I proved in the same volume that, for the above range of values of Q, nontrivial additive functions exist, defined on the interval $[1, x]$, so that the inequality at (22) goes in the other direction.

My argument employed an integral representation for $E_1(y, q, r)$, by Dirichlet series, to split the sum involving E_1 into two pieces. One piece was estimated somewhat directly, and the other piece through its dual. An elaboration of this procedure may be applied to obtain the above theorem. It proves convenient to replace σ by a complex variable s, and to work with $\mathrm{Re}(s)$ near to $1/2$.

We see that we have control over only the first $x^{1/2-\varepsilon}$ of the coordinates of $\hat{A}_2^* B f$; in terms of the measure $\widehat{du}$ roughly half of them. Thus we obtain

$$(23)$$

$$\left(\sum_{q \leqslant x^\delta} \frac{1}{q}\left(1 - \frac{1}{q_0}\right)|f(q) - (Tf)(q)|^2\right)^{1/2} \ll \|(A_2 - B)f\| + \frac{\log\log x}{(\log x)^{1/2}}\|f\|,$$

for each fixed $\delta < 1/2$.

Several variants of this inequality occur in Chapter 8 of [13], under the title 'The Loop'. If $f(n + 1) - f(n)$ is small, then a (weighted) estimate for the size of $f(q)^2$ over the range $q \leqslant x$ gives a better (implicit) estimate for $f(q)^2$ over the range $q \leqslant x^{1/2-\varepsilon}$. In that volume the functional inequality is solved by considering the L^1 form of it, giving it a continuous form (not surprisingly involving the measure $u^{-1}\,du$), and obtaining an approximate differential equation. The solution of this differential equation gives the form of the first eigenfunction of T.

5. Approximate functional equations. For motivation we return to the example at (15), and assume that in the sense described there

$$\beta(x)^{-1}(f(n) - \alpha(x))$$

obeys the weak law of large numbers. It is a consequence of this assumption that for every positive v

(24) $$\sideset{}{'}\sum_{p \leqslant x} \frac{1}{p} \to 0, \qquad x \to \infty,$$

where the summation is confined to those primes p for which

$$|f(p) - \alpha(x) + \alpha(x/p)| > v\beta(x).$$

Moreover, the function $\alpha(x)$ can be shown not to vary violently as x increases.

Using only these facts together with weak information concerning the distribution of prime numbers, we can prove the existence of functions $\omega(x)$, $\varepsilon = \varepsilon(x)$, so that as $x \to \infty$ we have $\varepsilon(x) \to 0$, whilst the approximate functional equation

(25) $$\alpha(ab) = \alpha(a) + \alpha(b) + o(\beta(x))$$

holds uniformly for all pairs $x^\varepsilon \leqslant a \leqslant b \leqslant x^{1/\varepsilon}$, with the possible exception of a set E of pairs for which $\mu_2 E \leqslant \varepsilon(x)$. Here μ_2 is the product measure on the direct sum of two copies of $\mathbb{R}^*$, induced by the Haar measure μ_1. Detailed proofs of these assertions may be found in my original 1975 paper [7], or Chapters 13 and 14 of the book [11].

With respect to the measure $\widetilde{du}$ from §3, we may regard (24) as an in-measure version of (23). Of course, the summation in (23) extends only over $q \leqslant x^{1/2-\varepsilon}$ rather than $q \leqslant x$, but this is compensated by the fact that the inequality is in mean-square form, and $\alpha(x)$ is given explicitly:

$$\alpha(x) = \sum_{q \leqslant x} \frac{1}{q}\left(1 - \frac{1}{q_0}\right) f(q).$$

In particular, this gives us some control over the growth rate of $\alpha(x)$. We hope for an L^2 approximate functional equation, whose solution we can essentially determine. In view of the form of (25), $\alpha(x)$ should in some sense be near to a logarithm.

In the event it is convenient to work in an L^1 sense.

THEOREM 6. *For $0 < \gamma < \delta \leqslant 1$ and $\alpha(x)$ as defined above, let*

(26) $$\theta(x) = \sum_{x^\gamma < q \leqslant x^\delta} \frac{1}{q} |f(q) - \alpha(x) + \alpha(x/q)|.$$

Then for each pair of positive numbers ε, B, there is a positive constant c so that

$$\alpha(t) = G(x) \log t - \eta(x) + O(Y(x)),$$

with

$$Y(x) = \sup_{x^c \leqslant w \leqslant x} \theta(w) + (\log x)^{-B} \sum_{q \leqslant x} \frac{|f(q)|}{q} + \max_{x^c \leqslant q \leqslant x} q^{-1}|f(q)|,$$

holds uniformly for $x^\varepsilon \leqslant t \leqslant x$, and all x sufficiently large in terms of γ, δ, ε, B.

The functions G, η are Lebesgue measurable, and the constant c depends at most upon ε, γ, and δ. It is possible to reduce the explicit dependence of the error term upon the values $f(q)$.

LEMMA 7. *In the notation of Theorem 7, there is a constant d so that*

$$\sum_{q \leqslant x} \frac{|f(q)|}{q} \ll \left(\max_{q \leqslant d} |f(q)| + \sup_{w \leqslant x^{1/\delta}} \theta(w) \right) (\log x)^d$$

for all $x \geqslant 2$.

I give here only a lightning sketch of these results, which I first exposed in the spring of 1980. Full details of the necessary arguments may be found in Chapter 9 of my book [13].

After subtraction, (26) implies that for any fixed $k \geqslant 1$

(27) $$\sum_{x^\gamma < q \leqslant x^\delta} \frac{1}{q} |\alpha(z) - \alpha(z/q) - \alpha(w) + \alpha(w/q)| \ll Y(x^k)$$

uniformly for $x \leqslant w \leqslant z \leqslant x^k$. In particular, as was the case in the above treatment of the law of large numbers, the individual $f(q)$ disappear. The operator T with $(Tf)(q) = \alpha(x) - \alpha(x/q)$ was initially defined on the space $\mathbb{C}^s$, where $s = s(x)$ denotes the number of prime powers q not exceeding x. With the step to (27) we widen the study of T to a patch of neighboring spaces $\mathbb{C}^{s(w)}$, $x \leqslant w \leqslant x^k$. Note that in the usual metric on the line, x and x^k are far apart for large x. In the metric on $\mathbb{R}^*$ given by

$$\rho(u, v) = \left| \frac{\log u}{\log x} - \frac{\log v}{\log x} \right|$$

they are at a distance $k - 1$, bounded independently of x ($\geqslant 2$), and so close.

With $\Lambda(x, y)$ defined to be $\alpha(x) - \alpha(x/y) - \alpha(y)$, as in the paragraph following (15), we give (27) a continuous form:

$$\int_{x^\gamma}^{x^\delta} |\Lambda(z, u) - \Lambda(u, w)| \frac{du}{u} \ll LY(x^k)$$

with $L = \log x$, uniformly for $x \leqslant w \leqslant z \leqslant x^k$.

Let $a = 1 - \delta + \delta^2$, $b = \delta^2/(2a)$. Suppose that $l \geqslant ab$, and $0 < \tau < 1$. Then

$$\iint\limits_{D} |\Lambda(w, u) - \Lambda(w, v)| \, \frac{dw \, dv}{wv} \ll L^2 Y(x^a)$$

holds uniformly for $x^{ab} \leqslant u^\delta \leqslant x^l$. Here the domain D of integration is given by $x^{a(1-\tau)} \leqslant v \leqslant x^a$, $v^{b(1-\tau)} \leqslant uv/w \leqslant v^b$, a form which betrays the integral to be on $\mathbb{R}^* \oplus \mathbb{R}^*$, rather than the two-dimensional real plane.

We continue to measure the rate at which $\Lambda(w, u)$ varies with w, u, by proving that

$$\iiint |\Lambda(z, u) - \Lambda(w, v)| \, \frac{du \, dv \, dw}{uvw} \ll L^3 Y(x^k)$$

uniformly for $x \leqslant z \leqslant x^k$, the integral being over the region

$$x^{\delta(1-\tau)} \leqslant u \leqslant x^\delta, \quad x^{a(1-\tau)} \leqslant v \leqslant x^a, \quad v^{b(1-\tau)} \leqslant uv/w \leqslant v^b.$$

For τ sufficiently small in terms of γ, δ, this region is nonempty. It may be more readily visualized by subjecting each of the variables to the transformation $y \mapsto \log y / \log x$.

Employing this triple integral, a function $\lambda(x)$ may be defined so that

$$(28) \qquad \int_{x^{\delta(1-\delta\tau/9)}}^{x^\delta} |\Lambda(z, u) - \lambda(x)| \, \frac{du}{u} \ll LY(x^k)$$

uniformly for $x \leqslant z \leqslant x^k$.

These results are obtained by splitting the integrals into a few pieces, and making appropriate changes of variable. It is not difficult to deduce from (28) that for each z_1 ($\geqslant 1$ and) not exceeding a certain power of x, the function $\alpha(z) - \alpha(z/z_1)$ is essentially constant over the interval $x \leqslant z \leqslant x^k$. This is an analogue of the approximate functional equation (25).

Moreover, it follows from (26) that

$$\sum_{x^\gamma < q \leqslant t} \frac{1}{q}\left(1 - \frac{1}{q_0}\right)(f(q) - \Lambda(x, q) + \alpha(q)) \ll \theta(x)$$

uniformly for $x^\gamma \leqslant t \leqslant x^\delta$. This equation, too, may be put into a continuous form so that, together with (28) in the case $z = x$, it gives

$$\alpha(t) - \int_{x^\kappa}^{t} \frac{\alpha(u)}{u \log u} \, du = F(x) + \lambda(x) \log \log t + O(Y(x))$$

with $\kappa = \delta(1 - \delta\tau/9)$. Here the function $F(x)$ can be given explicitly.

Denoting the integral by J we note that the left-hand side of this approximate differential equation has the form

$$t(\log t)^2 \frac{d}{dt}\left(\frac{J}{\log t}\right),$$

demonstrating an integrating factor. We can essentially solve the equation.

This leads to the desired representation for $\alpha(t)$ over the range $x^\kappa \leqslant t \leqslant x^\delta$, and its validity over the longer range $x^\varepsilon \leqslant t \leqslant x$ may be obtained from (28).

Suppose now that for $0 < \gamma < \delta < 1$

$$\psi(x) = \sum_{x^\gamma < q \leqslant x^\delta} \frac{1}{q} |f(q) - (Tf)(q)|^2,$$

where T is defined on $\mathbb{C}^s$, as earlier. An application of the Cauchy-Schwarz inequality, together with an estimate from elementary number theory, shows that the function $\theta(x)$ defined in Theorem 6 is $\ll \psi(x)^{1/2}$. With the corresponding representation of $\alpha(t)$ to hand we readily derive

THEOREM 8. *For suitable $G(x)$, c, and any fixed B*

$$\sum_{x^\gamma < q \leqslant x^\delta} \frac{1}{q} |f(q) - G(x) \log q|^2 \ll \sup_{x^c \leqslant w \leqslant x} \psi(w) + (\log x)^{-B} \sum_{q \leqslant x} \frac{|f(q)|^2}{q}$$

uniformly for all f, and $x > 2$.

This may be compared with the inequality

$$\sum_{q \leqslant x} \frac{1}{q} |f(q) - H(x) \log q|^2 \ll \|f - Tf\|^2 + (\log x)^{-1/2} \sum_{q \leqslant x} \frac{|f(q)|^2}{q}$$

obtained by employing the spectral decomposition of $\mathbb{C}^s$ with respect to T. The method of approximate functional equations involves some loss of precision, but as a general method is very flexible. Not only does it immediately apply to norms other than the mean-square norm considered here; at a deeper level, the interpolating points q may be replaced by others, provided they are reasonably well distributed. This may, of course, change the measure such as $u^{-1} du$ in the resulting continuous version of the approximate functional equation. Progress may then be hoped for if there is a corresponding exact functional equation whose solutions can be usefully classified.

In our present circumstances the role of the exact functional equation is played by $\alpha(ab) = \alpha(a) + \alpha(b)$ on the positive reals, which is Cauchy's equation on $\mathbb{R}^*$.

Studies of versions of Cauchy's equation and of the equation $g(x + y) = g(x)g(y)$, which are to hold on the line, but only approximately with respect to some measure, may be found in Elliott [14, 15]. Each case has an associated approximate differential equation.

Define

$$\omega(x) = \sup_{x^c \leqslant y \leqslant x} \left(\frac{1}{y} \sum_{n \leqslant y} |f(n) - f(n-1)|^2 \right)^{1/2}.$$

It follows from our half-norm inequality (23) and Theorem 8 that for any fixed $0 < \varepsilon < \delta < 1/2$, and suitable c,

(29)

$$\sum_{x^\varepsilon < q \leqslant x} \frac{1}{q} |f(q) - K \log q|^2 \ll \omega(x^{1/\delta})^2 + L^{-1}(\log L)^2 \sum_{q \leqslant x^{1/\delta}} \frac{|f(q)|^2}{q}.$$

Here K denotes $G(x^{1/\delta})$. If we introduce the inner product

$$[f, g] = \sum_{x^\varepsilon < q \leqslant x} \frac{f(q)\,\overline{g(q)}}{q}$$

on the space of complex tuples, one coordinate for each prime power q in the interval $x^\varepsilon < q \leqslant x$, and the corresponding norm $|f|_1 = [f, f]^{1/2} \geqslant 0$, then, with $K_0|\log|_1^2 = [f, \log]$, we have

$$|f - K\log|_1^2 = |f - K_0\log|^2 + |K - K_0|^2|\log|_1^2.$$

We shall apply this device of translating so that a certain inner product vanishes, several times in the remainder of this section. As an application here, we see that the inequality (29) is valid with K replaced by K_0, so that without loss of generality

$$|K|^2 \ll L^{-2} \sum_{x^\varepsilon < q \leqslant x} \frac{|f(q)|^2}{q}$$

may be assumed.

To remove the condition $q > x^\varepsilon$ in (29), let g_2 coincide with $f - K\log$ on the prime powers q in $x^\varepsilon < q \leqslant x$, and be zero otherwise. Define $g_2 = f - K\log - g_1$. Then by the Turán-Kubilius inequality

$$\|(A_2 - B)g_2\|^2 \ll \|g_2\|^2 \ll \sigma$$

with σ denoting the upper bound at (29). Since $\|(A_2 - B)\log\| \ll x^{-1}$ we have

$$\|(A_2 - B)(f - K\log)\|^2 \ll \|(A_2 - B)f\|^2 + x^{-2}|K|^2,$$

and in view of our bound for the size of K

$$\|(A_2 - B)g_1\|^2 \ll \sigma.$$

The inequality (23) is now applied to g_1 rather than f:

$$\sum_{q \leqslant x^\delta} \frac{1}{q}\left(1 - \frac{1}{q_0}\right)|g_1(q) - (Tg_1)(q)|^2 \ll \sigma + L^{-1}(\log L)^2\|g_1\|^2.$$

Since $\varepsilon < 1 - \delta$, $(Tg_1)(q) = 0$ for each $q \leqslant x^\delta$. The sum on the left-hand side of this inequality is $\|g_1\|^2$, and we obtain the analogue of (29) for the range $q \leqslant x^\varepsilon$.

Define

$$\rho(x) = \left(\sum_{q \leqslant x} \frac{1}{q}\left(1 - \frac{1}{q_0}\right)|f(q) - \eta(x)\log q|^2\right)^{1/2}$$

with $\eta(x)\|\log\|^2 = (f, \log)$ on $\mathbb{C}^s$. In the inequality (29) we omit the condition $q > x^\varepsilon$, and replace f by $f - \eta(x^{1/\delta})\log$:

$$\|f - (K + \eta(x^{1/\delta}))\log\|^2$$
$$\ll \omega(x^{1/\delta})^2 + L^{-1}(\log L)^2\rho(x^{1/\delta})^2 + x^{-2c}|\eta(x^{1/\delta})|^2.$$

The inner product argument allows us to replace $K + \eta(x^{1/\delta})$ by $\eta(x)$, to derive

$$(30) \qquad \rho(x) \ll \omega(x^{1/\delta}) + L^{-1/2} \log L \rho(x^{1/\delta}) + x^{-c}|\eta(x^{1/\delta})|.$$

This gives an iterative scheme for the improvement of bounds on $\|f\|$, but such a scheme needs an initial bound. In the present circumstances this is provided by the following result (from Elliott [10]).

LEMMA 9.

$$\sum_{n \leqslant x} |f(n)|^2 \ll x \log x \sum_{n \leqslant x} n^{-1}|f(n) - f(n-1)|^2.$$

A short inductive proof of this result appears in the above reference. A consideration of the general difference $f(an + b) - f(cn + d)$, which is more complicated and involves the application of Kloosterman sums, is made in Chapter 3 of my book [13].

Since the function $M_0(f) = [x]^{-1}\sum_{n \leqslant x} f(n)$ satisfies

$$|M_0(f)|^2 \ll x^{-1} \sum_{n \leqslant x} |f(n)|^2,$$

we can assert that

$$x^{-1} \sum_{n \leqslant x} |f(n) - M_0(f)|^2 \ll \log x \sum_{n \leqslant x} n^{-1}|f(n) - f(n-1)|^2 \ll \beta(x) \log x,$$

say. Together with Ruzsa's theorem (Theorem 4 of §3) this guarantees a λ such that

$$|\lambda|^2 + \|f - \lambda \log\|^2 \ll \beta(x) \log x.$$

In particular

$$(31) \qquad \rho(x)^2 \ll \beta(x) \log x, \qquad \|f\|^2 \ll \beta(x)(\log x)^3.$$

Applied to the inequality (30) these last give

$$\rho(x) \ll \omega(x^{1/\delta}) + \beta(x^{1/\delta})^{1/2} \log \log x.$$

It is apparent that the factor $\log \log x$ could be expensive, since it gives an increasing weight to the small differences $f(n) - f(n-1)$. To overcome this, three iterations of (30) are helpful:

$$\rho(x) \ll \sum_{j=1}^{3} \left(\frac{\log \log x}{\sqrt{\log x}} \right)^{j-1} \theta(x^{1/\delta^j}) + \left(\frac{\log \log x}{\sqrt{\log x}} \right)^3 \rho(x^{1/\delta^3})$$

where $\theta(y) = \omega(y) + y^{-c\delta}|\eta(y)|$. Bearing in mind the bounds (31) we obtain, in particular,

$$\rho(x) \ll \sup_{x^{1/\delta} \leqslant y \leqslant x^{1/\delta^3}} \omega(y) + \frac{(\log \log x)^3}{\log x} \left(\sum_{n \leqslant x^{1/\delta^3}} \frac{|f(n) - f(n-1)|^2}{n} \right)^{1/2}.$$

With $\eta = \eta(x)$ we have

$$x^{-1}|\eta|^2 \ll \|(A_2 - B)\eta \log\|^2 \ll \|(A_2 - B)(f - \eta \log)\|^2 + \omega(x)^2$$
$$\ll \rho(x)^2 + \omega(x)^2,$$

having applied the Turán-Kubilius inequality for the last time in this section.

Altogether we have established the following new result:

THEOREM 10. *For any fixed $c > 8$ the inequality*

$$x^{-1}|\eta|^2 + \sum_{q \leqslant x} \frac{1}{q}\left(1 - \frac{1}{q_0}\right)|f(q) - \eta \log q|^2$$

$$\ll \sup_{2 \leqslant y \leqslant x^c} y^{-1} \sum_{n \leqslant y} |f(n) - f(n - 1)|^2,$$

with $\eta\|\log\|^2 = (f, \log)$, holds for all additive functions, and real $x \geqslant 2$.

With more care the supremum in the statement of Theorem 10 may be taken over a range $x^d \leqslant y \leqslant x^c$ for a positive d, independent of f. The first step is to obtain a constant d, possibly larger than c, so that the inequality is valid with the supremum over $x^{1/d} \leqslant y \leqslant x^d$.

With a squeeze perhaps only two iterations of (30) might be needed, thus reducing the condition $c > 8$ in the theorem to $c > 4$.

The more general differences of additive functions mentioned earlier may be similarly treated, save that at present a larger value of c is needed.

6. Other norms. There are analogues of the Turán-Kubilius and related inequalities which involve norms other than L^2. As a background example consider the inequality

$$(32) \qquad \sum_{q \leqslant x} \frac{1}{q}\left(1 - \frac{1}{q_0}\right)\left|f(q) - \sum_{x/q < l \leqslant x} \frac{f(l)}{l}\left(1 - \frac{1}{l_0}\right)\right|^\alpha \leqslant \varepsilon^\alpha,$$

where $\alpha > 1$, and l denotes a prime power. In order to study a continuous analogue we view the operator T as being defined on the space L^α with respect to the measure $u^{-1}\,du$ on $0 < u \leqslant 1$. A straightforward application of Hölder's inequality shows, that, into L^α, T is bounded, indeed compact.

The map T is one-to-one, but its image in L^α contains only continuous functions, and so doesn't cover the whole of L^α. Thus 0 is the sole member of its continuous spectrum.

The dual space of L^α may be identified with $L^{\alpha'}$, $\alpha^{-1} + (\alpha')^{-1} = 1$, and the dual operator T' is formally the same as T. Moreover, since T is compact, T and T' have the same spectra, even down to the finite multiplicity of each eigenvalue. Thus, without loss of generality we may confine ourselves to the cases $\alpha \geqslant 2$. This is convenient, since we can then drop down to L^2.

In fact $u^{-1}\,du$ does not give a finite measure to the whole interval $(0, 1)$. It is advantageous to consider the functions f of $L^\alpha(0, 1)$ restricted to the interval $\delta \leqslant u \leqslant 1 - \delta$, and so belonging to $L^2(\delta, 1 - \delta)$, $0 < \delta < 1/2$. I shall denote a typical restricted function by f_1.

On $L^2(\delta, 1 - \delta)$ the operator

$$G: h \mapsto \int_{1-u}^{1-\delta} h(v) \frac{du}{v}$$

is self-adjoint, and we may employ the means used earlier when $\delta = 0$. The eigenfunctions of G once again satisfy the differential equations (11), but they can no longer be a set of polynomials orthogonal with respect to the measure $u^{-1} du$, and a simple explicit representation for them does not seem forthcoming. Instead we think of δ as small, and of $L^2(\delta, 1 - \delta)$ as an approximation to $L^2(0, 1)$.

Let g be a unit eigenfunction of T on $L^\alpha(0, 1)$, with associated eigenvalue λ. Then on the interval $\delta \leqslant u \leqslant 1$ we have

$$Tg = Tg_1 + \int_{1-\delta}^{1} g(v) \frac{dv}{v}.$$

We see that in $L^2(0, 1)$

$$\|Tg_1 - \lambda g_1\|^2 = \int_{\delta}^{1-\delta} \left| \int_{1-\delta}^{1} g(v) \frac{dv}{v} \right|^2 \frac{du}{u} + \int_{1-\delta}^{1} |Tg_1|^2 \frac{du}{u}.$$

Bearing in mind that on $L^\alpha(0, 1)$, $\|g_1\| \leqslant 1$, we may majorize $|Tg_1|$ by means of Hölder's inequality, and obtain for the last of these integrals the upper bound $2\delta(\log 1/\delta)^{2/\alpha'}$. Similarly

$$\left| \int_{1-\delta}^{1} g(v) \frac{dv}{v} \right| \leqslant (-\log(1 - \delta))^{2/\alpha'},$$

and the penultimate integral is not more than $(2\delta)^{2/\alpha'} \log 1/\delta$. Altogether

$$\|Tg_1 - \lambda g_1\| \ll \delta^c$$

for some positive constant c. Moreover, since g is not almost surely zero, $\|g_1\| \geqslant c_0 > 0$ holds in $L^2(0, 1)$ for all sufficiently small δ.

It now follows from the spectral decomposition of $L^2(0, 1)$ with respect to T, that for some eigenvalue $\lambda_j = (-1)^{j-1} j^{-1}$ we have $|\lambda - \lambda_j| \ll \delta^{c/2}$, the implied constant depending at most upon α. Letting $\delta \to 0^+$ we see that λ belongs to the closure of the spectrum of T on $L^2(0, 1)$, and so to that spectrum itself. Moreover, the polynomials $\psi_j(u)$ belong to $L^\alpha(0, 1)$ for every $\alpha > 0$, and are still eigenfunctions of T. Thus, disregarding multiplicities, the spectrum of T on L^α, $\alpha > 1$, is the same as that of T on L^2.

In fact every eigenvalue of T is still simple. For example, when $\alpha > 1$ we can define the map

$$T_1: f \mapsto Tf - (f, \psi_1)\psi_1,$$

where the inner product (f, ψ_1) is taken in $L^2(0, 1)$. Since ψ_1 belongs to every $L^\beta(0, 1)$, T_1 is well defined. T_1 is (again) compact and formally self-dual. The above argument shows that its spectrum is that of T with the eigenvalue 1 omitted, so that its spectral radius is $1/2$. Suppose now that g is an eigenvalue of T with corresponding eigenvalue 1, and fix an integer k so that $\|T_1^k\| < 3/4$. Then a simple check shows that $T_1^k g = g - (g, \psi_1)\psi_1$, from which

$$g = (I - T_1^k)^{-1}(g, \psi_1)\psi_1 = (g, \psi_1)\psi_1.$$

Apart from its normalization, g is uniquely determined. In particular the eigenvalue 1 is simple.

We are now in a positive to 'solve' the inequality (32). It may be viewed as $\|f - Tf\| \leqslant \varepsilon$, where T is given by the map (19), and the space is $\mathbb{C}^s$ with the norm

$$\left(\sum_{q \leqslant x} \frac{1}{q}\left(1 - \frac{1}{q_0}\right)|x_q|^\alpha \right)^{1/\alpha}.$$

Since T is bounded, $\|T^j f - T^{j+1} f\| \ll \varepsilon$ for each positive j, and after several applications of the triangle inequality, $\|f - T^k f\| \ll \varepsilon$ for each positive k.

It is readily proved by induction that

$$T_1^k f = T^k f - \nu_1\left(T^{k-1}f, d_1\right)d_1$$

where, as earlier, d_1 is the first eigenfunction of T in $L^2(\mathbb{C}^s)$ and T_1 is the operator

$$f \mapsto Tf - \nu_1(f, d_1)d_1.$$

Of course d_1 is still an eigenfunction of T, but its norm in L^α may possibly grow with x. Then we have

$$\left\|\left(I - T_1^k\right)f - \nu_1\left(T^{k-1}f, d_1\right)d_1\right\| \ll \varepsilon.$$

Denoting by R_1 the resolvent of $I - T_1^k$, and noting that $R_1(d_1) = d_1$, we obtain the solution

$$\left\|f - \nu_1\left(T^{k-1}f, d_1\right)d_1\right\| \ll \varepsilon,$$

with one caveat; we need a bound for the norm of the resolvent which does not depend upon x.

In the present circumstances we do not have a space on which to exhibit any approximation to the continuous T by its discrete analogues. Since L^α is not a Hilbert space, there is no spectral decomposition with respect to T available, and although the spectral radius of T_1 may be expected to approach $1/2$, the number of iterations k needed to ensure $\|T_1^k\| < 3/4$ may grow with x. I shall show that effectively it does not.

It is convenient to introduce the operator S_1: $f \mapsto Tf - (f, \hat\psi_1)\hat\psi_1$, to be used as an approximation to T_1. So far we only know this approximation to be true in $L^2(\mathbb{C}^s)$.

We write S_1 in the form

$$f \mapsto \int_2^x h(u, v)f(v)\,\widehat{dv},$$

where $h(u, v) = \delta(u, v) - s(u)s(v)$ with $\delta(u, v) = 1$ if $uv > x$, and $= 0$ otherwise; $s(q)$ is the q-coordinate of $\hat\psi_1$. Then S_1^k may be represented by a kernel $\rho(u, v)$ given by (reading right to left)

$$\int_2^x h(u, u_{k-1})\,\widehat{du_{k-1}} \cdots \int_2^x h(u_2, u_1)h(u_1, v)\,\widehat{du_1}.$$

An application of Hölder's inequality yields the bound

$$\left\| S_1^k \right\|^\alpha \leqslant \int_2^x \widehat{du} \left(\int_2^x |\rho(u,v)|^{\alpha'} \widehat{dv} \right)^{\alpha/\alpha'}.$$

For each fixed v, $\rho(u,v) = T_1^{k-1}(h(\cdot,v))$, and we now approximate to the continuous T_1. We note that after the substitution τ: $u \to (\log u)/\log x$ the unit square $[0,1]^2$ can be divided by straight-line segments into finitely many pieces (depending upon k) on each of which integrands such as $\rho(\tau u, \tau v)$ are continuous. By enclosing these lines in strips of width $(\log x)^{-1/2}$, and removing similar end-strips $1 - (\log x)^{-1/2} < \tau u \leqslant 1$, $1 - (\log x)^{-1/2} < \tau v \leqslant 1$, we can ensure that on the remainder $\rho(\tau u, \tau v)$ has continuous partial derivatives in τu, τv, uniformly bounded by a power of $\log x$. This enables the relevant parts of the various integrals to be replaced by continuous analogues. The contribution remaining which corresponds to the removed strips is anyway small, since the integrands themselves are bounded by a power of $\log \log x$. For example, $|h(u,v)| \leqslant 3$ uniformly in u, v, so that $\rho(u,v) \ll (\log \log x)^{k-1}$.

In this way we obtain (extending a previous notation)

$$(33) \qquad \left\| S_1^k \right\|^\alpha \leqslant \int_0^1 \frac{du}{u} \left(\int_0^1 \left| T_{1c}^{k-1}(t(\cdot,w)) \right|^{\alpha'} \frac{dw}{w} \right)^{\alpha/\alpha'} + o(1),$$

where $t(w,s) = 1 - 2sw$ if $s + w > 1$, and $= -2sw$ otherwise. This argument could not be carried out for the discrete operator T_1, since we have no information concerning the individual coordinates of the eigenfunction d_1.

If $g(z)$ denotes the function $T_{1c}^{k-2}(t(\cdot,w))$, then

$$T_{1c}^{k-1}(t(\cdot,w))(u) = \int_{1-u}^1 g(z) \frac{dz}{z} - 2u \int_0^1 g(z) \, dz.$$

Applications of the Cauchy-Schwarz inequality give

$$\|g\|_2 \left(\sqrt{\log(1-u)} + u\sqrt{2} \right)$$

as an upper bound for these expressions, where $\| \ \|_2$ denotes the norm on $L^2(0,1)$. Since T_{1c} has a spectral radius $1/2$ on that space,

$$\|g\|_2 \leqslant 2^{-k+2} \|t(\cdot,w)\|_2.$$

A direct calculation shows that this last norm is $\ll \sqrt{\log(1-w)} + w$, and we see that the innermost integral in the bound at (33) does not exceed a constant multiple of

$$2^{-k} \left(\sqrt{\log(1-u)} + u \right)^{\alpha'}.$$

This in turn leads to a bound

$$\left\| S_1^k \right\| \leqslant c_0 2^{-k}$$

for all $x \geqslant 2$. An important feature of this result is that the constant c_0 is absolute. Given $0 < \varepsilon < \frac{1}{2}$ we may choose k large enough that the bound does not exceed $(\frac{1}{2} + \varepsilon)^k$.

By considering the powers of S_1 k at a time, we can obtain a constant c so that $\|S_1^m\| \leqslant c(\frac{1}{2} + \varepsilon)^m$ holds for all $m \geqslant 1$. For $|\lambda| \geqslant \frac{1}{2} + 2\varepsilon$ the resolvent operators $(\lambda I - S_1)^{-1}$ are then bounded, with norms not exceeding c/ε.

Returning to our initial hypothesis (32), we write it in the form

$$\|(I - S_1)f - (f, \hat{\psi}_1)\hat{\psi}_1\| \leqslant \varepsilon,$$

and deduce that

$$(34) \qquad \left\| f - (f, \hat{\psi}_1)(I - S_1)^{-1}\hat{\psi}_1 \right\| \ll \varepsilon.$$

A direct computation shows that $\|S_1\hat{\psi}_1\| = \|T\hat{\psi}_1 - \hat{\psi}_1\| \ll (\log x)^{-1/\alpha}$, and since

$$(35) \qquad \lambda(\lambda I - S)^{-1} - I = \lambda(\lambda I - S)^{-1}S,$$

we also have

$$\left\| (I - S_1)^{-1}\hat{\psi}_1 - \hat{\psi}_1 \right\| \ll (\log x)^{-1/\alpha}.$$

However, we cannot simplify (34) further without some information concerning the inner product $(f, \hat{\psi}_1)$.

The (first) eigenvalue ν_1 of the discrete T is only known to be near to 1, but it will exceed $3/4$ for all sufficiently large x. Then

$$(\nu_1 I - S_1)d_1 = (\nu_1 I - T)d_1 + (d_1, \hat{\psi}_1)\hat{\psi}_1,$$

so that

$$d_1 = (d_1, \hat{\psi}_1)(\nu_1 I - S_1)^{-1}\hat{\psi}_1.$$

This not only shows (once again) that the eigenvalue ν_1 of T on $L^\alpha(\mathbb{C}^s)$ is simple, but also, since $(d_1, \hat{\psi}_1)$ may be computed in $L^2(\mathbb{C}^s)$, that d_1 is bounded in $L^\alpha(\mathbb{C}^s)$, independently of x. Another application of (35), and we obtain

$$(36) \qquad \|d_1 - \hat{\psi}_1\| \ll (\log x)^{-\delta}, \qquad \delta = \min(1/2, 1/\alpha),$$

of the same form as a result in §3, but with a different norm. This enables us to show that T_1 and S_1 are genuinely close.

In fact

$$(T_1 - S_1)f = (f, \hat{\psi}_1 - d_1)\hat{\psi}_1 + (f, d_1)(\hat{\psi}_1 - d_1).$$

By Hölder's inequality

$$|(f, \hat{\psi}_1 - d_1)| \leqslant \|f\|_\alpha \|\hat{\psi}_1 - d_1\|_{\alpha'},$$

where the second of these norms, considered on $L^{\alpha'}(\mathbb{C}^s)$, is $\ll (\log x)^{-\sigma}$, $\sigma = \min(1/2, 1/\alpha')$. In this manner we see that for a suitable δ

$$\|T_1 - S_1\| \ll (\log x)^{-\delta}.$$

The resolvents $(I - T_1^j)^{-1}$ are now seen to be bounded for each positive j. In particular, we may write the hypothesis (32) in the form

$$\|f - T_1 f - \nu_1(f, d_1)d_1\| \leqslant \varepsilon,$$

and deduce the aesthetically more pleasing

(37) $$\|f - \nu_1(f, d_1)d_1\| \ll \varepsilon.$$

We may continue with this analysis, introducing the operator

$$S_j: f \mapsto Tf - \sum_{i=1}^{j} \lambda_i(f, \hat{\psi}_i)\hat{\psi}_i.$$

As for S_1, and with a similar meaning, we prove that asymptotically the spectral radius of S_j is not more than $1/(j + 1)$. For $|\lambda| \geqslant \beta > 1/(j + 1)$ the resolvents $(\lambda I - S_j)^{-1}$ have norms bounded in terms of β and j, uniformly for x sufficiently large.

Further direct calculation shows that if $i \neq j$, $(\hat{\psi}_i, \hat{\psi}_j) \ll (\log x)^{-\delta}$ for some positive constant δ not depending upon i, j or x ($\geqslant 2$). This enables us to show that $\|S_j \hat{\psi}_i\| \ll (\log x)^{-\delta}$ for $1 \leqslant i \leqslant j$, and in view of (35) that

$$\left\| \lambda(\lambda I - S_j)^{-1}\hat{\psi}_i - \hat{\psi}_i \right\| \ll (\log x)^{-\delta}$$

uniformly for $1 \leqslant i \leqslant j$, $|\lambda| \geqslant \beta$, $x \geqslant 2$. Hence we obtain a representation

$$d_i = \sum_{r=1}^{j} \lambda_r(d_i, \hat{\psi}_r)(\nu_i I - S_j)^{-1}\hat{\psi}_r$$

valid for $1 \leqslant i \leqslant j$; that the $\|d_i\|$ are bounded uniformly in x; and that

$$\left\| \lambda_i d_i - \sum_{r=1}^{j} \lambda_r(d_i, \hat{\psi}_r)\hat{\psi}_r \right\| \ll (\log x)^{-\delta}.$$

Here the approximation $\nu_i - \lambda_i \ll (\log x)^{-1/2}$ has been applied.

Since the operator T is bounded we may apply it $k - 1$ times to the function inside this norm, and so obtain

$$\left\| \lambda_i^k d_i - \sum_{r=1}^{j} \lambda_r^k(d_i, \hat{\psi}_r)\hat{\psi}_r \right\| \ll (\log x)^{-\delta},$$

for $k = 1, 2, \ldots, j$. These j inequalities may be treated in the manner of linear equations, using Cramer's rule. The coefficient matrix (λ_i^k), $1 \leqslant i, k \leqslant j$, has a van der Monde determinant, not zero because the λ_i are distinct and nonzero. Thus we obtain

(38) $$\|d_i - \hat{\psi}_i\| \ll (\log x)^{-\delta}, \qquad (d_i, \hat{\psi}_r) \ll (\log x)^{-\delta}$$

for each fixed $i \neq r$.

Defining the operator T_j by

$$f \mapsto Tf - \sum_{i=1}^{j} \nu_i(f, d_i)d_i,$$

we can prove that $\|T_j - S_j\| \ll (\log x)^{-\delta}$ for some positive constant δ which is independent of j, and of x ($\geqslant 2$).

As a version of Lemma 2 for the operator T on $L^\alpha(\mathbb{C}^s)$ we have

LEMMA 11. *Assume that for some eigenvalue ν_j of the discrete operator T we have $\|Tg - \nu_j g\| = \varepsilon$. Then*

$$\left\| g - (g, d_j)d_j \right\| \ll \varepsilon, \quad (g, d_i)\|d_i\| \ll \varepsilon \qquad \text{for } i \neq j.$$

PROOF. I give only a sketch, since the method is the same as that employed in the estimation of the d_i. With $r \geqslant j$ and $w(f) = \sum_{i=1}^{r} \nu_i(f, d_i) d_i$, we write the hypothesis in the form

$$\|(\nu_j I - T_r)g - w(g)\| = \varepsilon.$$

The resolvent operator $(\nu_j I - T_r)^{-1}$ is bounded, and $T_r w(g) = 0$, so that

$$\left\| \nu_j g - \sum_{i=1}^{r} \nu_i(f, d_i) d_i \right\| \ll \varepsilon.$$

Since T is bounded, operating on the function estimated in this norm yields

$$\left\| \nu_j^k g - \sum_{i=1}^{r} \nu_i^k(g, d_i) d_i \right\| \ll \varepsilon$$

for each fixed k, $1 \leqslant k \leqslant r - 1$. Elimination now gives the desired result.

Some form of this argument can be applied when $\varepsilon = \|Tg - \lambda g\|$ and λ isn't an eigenvalue of T. Looking even farther back I give a version of Lemma 1.

LEMMA 12. *Let $\beta > 0$. For any $|\lambda| \geqslant \beta$ and function g in $L^\alpha(\mathbb{C}^s)$ there is an eigenvalue ν_i of T so that*

$$|\nu_i - \lambda| \, \|g\| \ll \|Tg - \lambda g\|,$$

the implied constant depending only upon β and α.

PROOF. This follows the proof of Lemma 11. One reaches an estimate

$$D\|g\| \ll \|Tg - \lambda g\|,$$

where D is the van der Monde determinant $\det(\nu_i^j)$, $0 \leqslant i \leqslant r$, $1 \leqslant j \leqslant r + 1$, with $\nu_0 = \lambda$, and r chosen so that $\beta > 1/(r + 1)$. This determinant satisfies

$$|D| = \left| \prod_{i=0}^{r} \nu_i \prod_{0 \leqslant i < j \leqslant r} (\nu_i - \nu_j) \right| > \gamma \prod_{j=1}^{r} |\lambda - \nu_j|$$

for some positive γ depending at most upon r, and so β. If some term in the last product is less than $1/4r$, then all the other terms are at least as large as $1/4r$. This gives the desired bound.

Altogether the operators T, T_i behave on $L^\alpha(\mathbb{C}^s)$ much as they do on $L^2(\mathbb{C}^s)$.

7. Intersection spaces. Let L^α be a space of μ-measurable functions, with corresponding norm

$$\|f\|_\alpha = \left(\int |f|^\alpha \, d\mu \right)^{1/\alpha}.$$

With a view to applications to the study of T, it is appropriate to consider the spaces $L^\alpha \cap L^\beta$ with norm $\|f\|_\alpha + \|f\|_\beta$, and their duals. More exactly, if Y denotes the space of μ-measurable functions f for which $\|f\|_\alpha$ and $\|f\|_\beta$ are both defined (finite), then $L^\alpha \cap L^\beta$ is the completion of Y with respect to the norm $\|f\|_\alpha + \|f\|_\beta$.

LEMMA 13. *Let S be the space $L^\alpha \cap L^\beta$ with norm $\|f\|_\alpha + \|f\|_\beta$, $1 < \beta \leqslant \alpha$. Let y in S' be given by*

$$f \mapsto \int fg\,d\mu,$$

where $\|g\|_{\alpha'} + \|g\|_{\beta'}$ is finite. Suppose that the norm of y is at most η. Then

$$\left(\int_{|g|\leqslant\eta} |g|^{\beta'}\,d\mu\right)^{1/\beta'} + \left(\int_{|g|>\eta} |g|^{\alpha'}\,d\mu\right)^{1/\alpha'} \leqslant 12^{\beta'/\alpha'}\eta.$$

Moreover, for any value of η the majorized expression is an upper bound for the norm of y.

PROOF. Consider the function

$$g_+ = \begin{cases} \mathrm{Re}(g) & \text{if } \mathrm{Re}(g) > \eta/4, \\ 0 & \text{otherwise}. \end{cases}$$

Then $f \mapsto \int fg_+\,d\mu$ gives a map on S considered as a real space of real functions, with norm at most $\|y\|$. For, defining

$$f_+ = \begin{cases} f & \text{if } \mathrm{Re}(g) > \eta/4, \\ 0 & \text{otherwise}, \end{cases}$$

and writing N for $\|y\|$, we see that for real f in $L^\alpha \cap L^\beta$

$$\left|\int gf_+\,d\mu\right| = \left|\int f_+g\,d\mu\right| \leqslant N\|f_+\| \leqslant N\|f\|.$$

For a complex space S we may similarly treat if, for f purely imaginary, and obtain an analogous result with $2N$ in place of N.

If we set $f = g_+^{\alpha'-1}$, then we have by Hölder's inequality (in an equality case)

$$\|f\|_{\alpha'}^{\alpha'} = \left|\int fg_+\,d\mu\right| \leqslant N\big(\|f\|_\alpha + \|f\|_\beta\big).$$

Here

$$\|f\|_\beta = \left(\int |g_+|^{\beta(\alpha'-1)}\,d\mu\right)^{1/\beta} \leqslant \left(\frac{\eta}{4}\right)^{-1+\alpha'/\beta'} \|g_+\|_{\alpha'}^{\alpha'/\beta}$$

and

$$\|f\|_\alpha = \left(\int |g_+|^{\alpha(\alpha'-1)}\,d\mu\right)^{1/\alpha} = \|g_+\|_{\alpha'}^{\alpha'/\alpha},$$

so that

$$\|g_+\|_{\alpha'}^{\alpha'} \leqslant N\big(\|g_+\|_{\alpha'}^{\alpha'/\alpha} + (\eta/4)^{-1+\alpha'/\beta'}\|g_+\|_{\alpha'}^{\alpha'/\beta}\big).$$

If the first of the two terms in the majorant exceeds $\frac{1}{2}\|g_+\|_{\alpha'}^{\alpha'}$, then

$$\left(\int |g_+|^{\alpha'}\,d\mu\right)^{1/\alpha'} \leqslant 2N.$$

Otherwise

$$\tfrac{1}{2}\| g_+ \|_{\alpha'}^{\alpha'} \leqslant N(\eta/4)^{-1+\alpha'/\beta'}\| g_+ \|_{\alpha}^{\alpha'/\beta},$$

which leads to a similar upper bound, but with $8^{\beta'/\alpha'}N$ in place of $2N$.

By considering the function

$$g_- = \begin{cases} -\operatorname{Re}(g) & \text{if } \operatorname{Re}(g) < -\eta/4, \\ 0 & \text{otherwise,} \end{cases}$$

and the analogues of g_+, g_- involving the imaginary part of g, and noting that when $|g| > \eta$

$$|g| \leqslant 4\max\big(|g_+|, |g_-|, |(ig)_+|, |(ig)_-|\big),$$

we obtain the bound

$$\left(\int_{|g|>\eta} |g|^{\alpha'}\, d\mu \right)^{1/\alpha'} \leqslant 10^{\beta'/\alpha'}N.$$

The integrals involving the condition $|g| \leqslant \eta$ may be likewise reduced to the case when g is real, and typically satisfies $0 \leqslant g \leqslant \eta$. In this case one sets $f = g^{\beta'-1}$, and applies Hölder's inequality as an equality, to obtain

$$\| g \|_{\beta'}^{\beta'} \leqslant N\big(\| f \|_\alpha + \| f \|_\beta\big).$$

Since $\alpha(\beta' - 1) \geqslant \beta'$, we have

$$\| f \|_\alpha \leqslant \left(\eta^{\alpha(\beta'-1)-\beta'}\| g \|_{\beta'}^{\beta'} \right)^{1/\alpha}.$$

From these we readily derive the bound

$$\left(\int_{|g|\leqslant\eta} |g|^{\beta'}\, d\mu \right)^{1/\beta'} \leqslant 2\eta,$$

and the first assertion of the lemma is justified.

The second assertion may be justified by expressing g as the sum of the two functions, according to whether $|g| \leqslant \eta$ or not, and applying to the corresponding functional on S the appropriate form of Hölder's inequality.

Consider now the space V', where $V = L^\alpha \cap L^\beta$ has norm $\| f \|_\alpha + \| f \|_\beta$ with respect to the measure $u^{-1}\, du$ on the interval $0 < u \leqslant 1$. The above lemma shows that V' is not in general the space $L^{\alpha'} \cap L^{\beta'}$. We continue with the study of this space.

LEMMA 14. *To each functional y in $(L^\alpha \cap L^\beta)'$, $1 \leqslant \beta < \alpha$, there corresponds a function g so that y has a representation*

$$f \mapsto \int_{|g|\leqslant\|y\|} fg \frac{du}{u} + \int_{|g|>\|y\|} fg \frac{du}{u},$$

where the function

$$t(c) = \left(\int_{|g|\leqslant c} |g|^{\beta'} \frac{du}{u} \right)^{1/\beta'} + \left(\int_{|g|>c} |g|^{\alpha'} \frac{du}{u} \right)^{1/\alpha'}$$

satisfies $t(\|y\|) \leqslant 24^{\beta'/\alpha'}\|y\|$.

Moreover, any function g for which t(c) is finite for some positive c defines a functional on $L^\alpha \cap L^\beta$, with a norm not exceeding t(c).

PROOF. Let y be a functional on V, with norm N. If $0 < \delta < 1$, and we extend the functions on $(\delta, 1)$ by having them assume the value zero on the remaining interval $0 < u < \delta$, then y induces a functional on $L^\alpha(\delta, 1) \cap L^\beta(\delta, 1)$ with norm at most N.

Since $u^{-1} du$ assigns a finite measure to the interval $(\delta, 1)$, $L^\beta(\delta, 1)$ may be viewed as a subspace of $L^\alpha(\delta, 1)$, and an application of Hölder's inequality shows that

$$\| f \|_\beta \leqslant \| f \|_\alpha (-\log \delta)^{(\alpha - \beta)/\alpha\beta}.$$

Moreover, the polynomials on $(\delta, 1)$ belong to both $L^\beta(\delta, 1)$ and $L^\alpha(\delta, 1)$, and are dense in the latter. Then y induces a functional on $L^\alpha(\delta, 1)$, where it has a representation

$$f \mapsto \int_\delta^1 gf \, \frac{du}{u},$$

the function g being essentially determined on $(\delta, 1)$. By varying δ we define a function g on the whole unit interval $0 < u \leqslant 1$.

For any function f on $0 < u \leqslant 1$ define

$$f_\delta = \begin{cases} f & \text{if } \delta \leqslant u \leqslant 1, \\ 0 & \text{if } 0 < u < \delta. \end{cases}$$

The functional

$$z : f \mapsto \int_\delta^1 f g_n \, \frac{du}{u},$$

where

$$g_n = \begin{cases} g & \text{if } |g| \leqslant n, \ n \text{ an integer,} \\ 0 & \text{otherwise,} \end{cases}$$

is bounded on V, and of norm at most N. Indeed, if $\theta_n = 1$ when $|g| \leqslant n$, and is zero otherwise, then in the usual notation

$$\langle f, z \rangle = \int_0^1 f_\delta \theta_n g \, \frac{du}{u} = \langle f_\delta \theta_n, y \rangle,$$

so that

$$|\langle f, z \rangle| \leqslant N \| f_\delta \theta_n \| \leqslant N \| f \|.$$

Clearly $(g_n)_\delta$ belongs to $L^\alpha(0, 1)$ and $L^\beta(0, 1)$, and we apply our previous lemma to obtain

$$\left(\int_{\substack{\delta \\ |g_n| \leqslant N}}^1 |g_n|^{\beta'} \, \frac{du}{u} \right)^{1/\beta'} + \left(\int_{\substack{\delta \\ |g_n| > N}}^1 |g_n|^{\alpha'} \, \frac{du}{u} \right)^{1/\alpha'} \leqslant \gamma N,$$

with $\gamma = 12^{\beta'/\alpha'}$.

The sets E_n of points on which $|g| \leq n$, $n = 1, 2, \ldots$, form an increasing nest. For any $r \geq N$, and all $n \geq r$,

$$\left(\int_{\delta \atop |g| \leq N}^{1} |g|^{\beta'} \frac{du}{u} \right)^{1/\beta'} + \left(\int_{\delta \atop u \in E_r}^{1} |g_n|^{\alpha'} \frac{du}{u} \right)^{1/\alpha'} \leq \gamma N.$$

We obtain the bound on $t(N)$, and so the first part of the lemma, by applying Fatou's lemma as $n \to \infty$, the dominated convergence theorem as $r \to \infty$, and again Fatou's lemma, as $\delta \to 0^+$. Thus

$$\langle f, y \rangle = \lim_{\delta \to 0^+} \langle f_\delta, y \rangle = \lim_{\delta \to 0^+} \int_{\delta}^{1} fg \frac{du}{u} = \int_{0}^{1} fg \frac{du}{u}.$$

The second part of Lemma 14 is straightforward.

With this integral representation to hand we can once again prove that T is formally self-dual. If the functional y corresponds to g, then $T'y$ corresponds to Tg. Since we already know that T is compact on L^α, L^β and so on $L^\alpha \cap L^\beta$, by a theorem of Schauder T' (and so T) is compact on $(L^\alpha \cap L^\beta)'$. A similar argument is valid for T_1.

The operator T is compact when viewed on each of the spaces L^α, $L^\alpha \cap L^\beta$, $(L^\alpha \cap L^\beta)'$, and has the same spectrum.

We note that there is a natural injection

(39) $$L^{\alpha'} \cap L^{\beta'} \to (L^\alpha \cap L^\beta)'$$

which takes g in the first space to the functional

$$y: f \mapsto \int_{0}^{1} fg \frac{du}{u}.$$

Here (as we have implicitly before in this section) we regard g in $L^{\alpha'} \cap L^{\beta'}$ to be the limit of functions g_m which are measurable and have finite $\| \: \|_{\alpha'}$ and $\| \: \|_{\beta'}$, and define y to be the limit as $m \to \infty$ of the functionals

$$f \mapsto \int_{0}^{1} fg_m \frac{du}{u}.$$

Since $\|y\| \leq \|g\|_{\alpha'} + \|g\|_{\beta'}$, this injection does not increase norms.

Lemma 13 remains valid, so we may construct analogues of these results using discrete spaces L^β with respect to the measure $\widehat{du}$.

8. High-Moment Inequalities. We return to the study of high-moment analogues of the Turán-Kubilius inequality. The first examples may be found in Elliott [12]:

(40)

$$\left(x^{-1} \sum_{n \leq x} |f(n) - M(f)|^\alpha \right)^{1/\alpha}$$

$$\ll \left(\sum_{q \leq x} q^{-1} |f(q)|^2 \right)^{1/2} + \left(\sum_{q \leq x} q^{-1} |f(q)|^\alpha \right)^{1/\alpha}$$

if $\alpha \geq 2$, and a similar inequality, but with the final term removed, when $0 < \alpha < 2$. I remark that Ruzsa [38] later showed that a wide class of upper-bound inequalities satisfied by sums of the random variables Y_q have

analogues involving additive arithmetic functions. These I shall not examine here.

When $1 < \alpha < 2$, the upper bound most naturally given by my original proof method is

(41)
$$\left(\sum_{\substack{q \leqslant x \\ |f(q)| \leqslant D}} \frac{1}{q}\left(1 - \frac{1}{q_0}\right)|f(q)|^2 \right)^{1/2}$$
$$+ \left(\sum_{\substack{q \leqslant x \\ |f(q)| > D}} \frac{1}{q}\left(1 - \frac{1}{q_0}\right)|f(q)|^\alpha \right)^{1/\alpha} \left(1 + \sum_{\substack{q \leqslant x \\ |f(q)| > D}} \frac{1}{q}\left(1 - \frac{1}{q_0}\right)\right),$$

valid for every positive D.

Consider now $f(q)$ as belonging to the space $(L^2 \cap L^{\alpha'})'$, $1 < \alpha \leqslant 2$, where L^β denotes a space of functions on the prime powers not exceeding x, with measure $\widehat{du}$. If $\| \ \|$ denotes the norm on this dual space, then with $D = \|f\|$ we see from Lemma 13 that

$$\sum_{\substack{q \leqslant x \\ |f(q)| > D}} \frac{1}{q}\left(1 - \frac{1}{q_0}\right) \leqslant D^{-1}\left(\int_{|f| > D} |f|^\alpha \widehat{du} \right)^{1/\alpha} \leqslant \gamma_1,$$

where γ_1 is a constant depending only upon α. Another application of that lemma shows that the expression of (41) is not more than $c_1\|f\|$ for some constant c_1 depending at most upon α. Moreover, except possibly for the value of c_1, this bound is of a correct form.

We see that for $1 < \alpha \leqslant 2$ it is natural to view the operator A_α, given by

$$(A_\alpha f)(n) = f(n) - M(f),$$

as being on $(L^2 \cap L^{\alpha'})'$ with $\widehat{du}$, and into L^α with respect to the frequency measure on the integers up to x. The content of the αth high-moment version of the Turán-Kubilius inequality is then that A_α has a norm which is bounded uniformly for all x.

For arithmetic functions which depend structurally upon their values at the prime powers, the awkwardness of dealing with their α-powers, $1 < \alpha < 2$, is somewhat explained by their being naturally members of the space $(L^2 \cap L^{\alpha'})'$, rather than L^α, with respect to $\widehat{du}$.

We can view A_α as a bounded operator in three circumstances.

$$\begin{aligned}
&\text{a:} \quad (L^2 \cap L^{\alpha'})' \to L^\alpha, \quad 1 < \alpha \leqslant 2, \\
&\text{b:} \quad L^2 \to L^\alpha, \quad\quad\quad\ 1 < \alpha \leqslant 2, \\
&\text{c:} \quad L^2 \cap L^\alpha \to L^\alpha, \quad\ \alpha > 2.
\end{aligned}$$

The natural analogue of the operator $A_2^* A_2$ which occurs in §3 is now $A_{\alpha'}' A_\alpha$. For $1 < \alpha \leqslant 2$ we obtain the composition

$$(L^2 \cap L^{\alpha'})' \xrightarrow{\text{a}} L^\alpha \simeq (L^{\alpha'})' \xrightarrow{\text{c'}} (L^2 \cap L^{\alpha'})'$$

where $\simeq$ denotes the standard isomorphism between L^α and the dual space of $L^{\alpha'}$, and c′ denotes the dual of the map A_α in the form c.

This gives

$$\left\| \check{A}'_{\alpha'} A_\alpha - I + T \right\| \ll (\log x)^{-\delta}$$

for some positive δ, and the discrete version of T, but with an underlying norm different from that of L^α considered in §6. However, the results of that section apply.

For any $\gamma > 0$ we may choose k large enough that both $\|T_1^k f\|_\alpha \leqslant \frac{1}{2}\gamma\|f\|_\alpha$ and $\|T_1^k f\|_2 \leqslant \frac{1}{2}\gamma\|f\|_2$ for all f in L^α, L^2 respectively (which here is of course all f). Then in terms of the norm on $L^2 \cap L^\alpha$ or of the norm on $(L^2 \cap L^{\alpha'})'$

$$\left\| T_1^k f \right\| \leqslant \left\| T_1^k f \right\|_\alpha + \left\| T_1^k f \right\|_2 \leqslant \gamma \| f \|.$$

The mechanics of §6 proceed, and given that $\|Tf - f\| = \varepsilon$ we may still conclude that the estimate (37) is valid, save that the norms are to be interpreted in the space $(L^2 \cap L^{\alpha'})'$, and the implied constant may depend on α. With d_1 the first eigenfunction of T, and $F = (f, d_1)$, we have

$$(42) \qquad \| f - Fd_1 \| \ll \| A_\alpha (f - Fd_1) \| \ll \| f - Fd_1 \|,$$

the outer norms being in $(L^2 \cap L^{\alpha'})'$ with respect to $\widehat{du}$, $1 < \alpha \leqslant 2$, and the center norm being in L^α with respect to the frequency measure on the integers up to x.

If $\alpha \geqslant 2$, we can form two chains of compositions:

$$(43) \qquad L^2 \cap L^\alpha \overset{c}{\to} L^\alpha \simeq \left(L^{\alpha'} \right)' \overset{b'}{\to} L^2,$$

and a similar chain with the last map replaced by

$$(44) \qquad \left(L^{\alpha'} \right)' \overset{a'}{\to} \left(L^2 \cap L^\alpha \right)'' \simeq L^2 \cap L^\alpha,$$

the last step by the standard isometric embedding of a space into its second dual. In this case the spaces have the same dimension, so may be identified. Thus $A'_{\alpha'} \colon L^{\alpha'} \to L^2 \cap L^\alpha$ is bounded.

Altogether, for $\alpha \geqslant 2$ we may view $A'_{\alpha'} A_\alpha$, and so T, as being a bounded operator on $L^2 \cap L^\alpha$ into itself. In particular, we may obtain the analogue of the inequalities (42), the outer norms now being in the space $L^2 \cap L^\alpha$ with respect to $\widehat{du}$.

For a general comparison, we introduce the notation

$$B_\beta(f, x) = \left(\sum_{q \leqslant x} \frac{|f(q)|^\beta}{q} \right)^{1/\beta}, \qquad D_\beta(f, x; A) = \left(\frac{1}{x} \sum_{n \leqslant x} |f(n) - A|^\beta \right)^{1/\beta},$$

$$A_0(f, x) = \sum_{q \leqslant x} \frac{f(q)}{q}, \qquad A_1(f, x; \lambda, c) = \lambda \log x + \sum_{\substack{q \leqslant x \\ |f_\lambda(q)| \leqslant c}} \frac{f_\lambda(q)}{q},$$

where $f_\lambda(q) = f(q) - \lambda \log q$. Furthermore, set

$$f'_{\lambda, c}(q) = \begin{cases} f_\lambda(q) & \text{if } |f_\lambda(q)| \leqslant c, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$f''_{\lambda,c}(q) = \begin{cases} 0 & \text{if } |f_\lambda(q)| \leqslant c, \\ f_\lambda(q) & \text{otherwise.} \end{cases}$$

Then we have the following two theorems of Hildebrand [22]:

THEOREM 15. *For each* (*real*) *additive function f, and real* $x \geqslant 1$, *there are numbers* λ_0, $c_0 \geqslant 0$ (*possibly depending upon f, x*) *so that for every* $\beta > 0$ *and real A*

(45)

$$D_\beta(f, x; A) \asymp |\lambda_0| + |A - A_1(f, x; \lambda_0, c_0)| + B_2\big(f'_{\lambda_0,c_0}, x\big) + B_\beta\big(f''_{\lambda_0,c_0}, x\big),$$

the implied constants only depending upon β.

This appears as Theorem 1 in Hildebrand's paper. As Theorem 1* he gives

THEOREM 16.

$$D_\beta(f, x; A) \asymp \min_{\lambda \in \mathbf{R}} \big(|\lambda| + |A - A_1(f, x; \lambda, c)| + B_2\big(f'_{\lambda,c}, x\big) + B_\beta\big(f''_{\lambda,c}, x\big)\big),$$

where $c = D_\beta(f, x; A)$. *Moreover, if* $\beta \geqslant 2$, *then*

$$D_\beta(f, x; A) \asymp \min_{\lambda \in \mathbf{R}} \big(|\lambda| + |A - A_0(f, x)| + B_2(f_\lambda, x) + B_\beta(f_\lambda, x)\big).$$

These results generalize the estimate

$$x^{-1} \sum_{n \leqslant x} |f(n) - A_0(f, x)|^2 \asymp \min_{\lambda \in \mathbf{R}} \big(\lambda^2 + B^2(f_\lambda, x)\big)$$

of Ruzsa, which we derived as Theorem 4 in §3, and improve to the same depth my earlier high-moment inequalities (40). According to a remark of Ruzsa, for $\beta > 2$ similar results were obtained by Manstavičius. An essential ingredient of Hildebrand's proofs, which do not involve functional analysis, is Ruzsa's bound [36]

$$Q(x) \ll \left(\min_{\lambda \in \mathbf{R}} \left(\lambda^2 + \sum_{p \leqslant x} p^{-1} \min\big(1, |f_\lambda(p)|^2\big) \right) \right)^{-1/2}$$

for the concentration function

$$Q(x) = \sup_y x^{-1} \sum_{\substack{n \leqslant x \\ y \leqslant f(n) \leqslant y+1}} 1.$$

This last estimate Ruzsa obtained by combining ideas from the theory of probability with the full panoply of Halasz' method [20] for the quantitative Fourier inversion of Dirichlet series which have Euler products.

In view of its heavy dependence upon duality, when not combined with the method of approximate functional equations the functional analytic method given in the present paper applies to the cases $\beta > 1$. However, it gives insight into the meaning of Hildebrand's theorems, and extends them to complex-valued functions.

For all $\beta > 1$ the inequality (45) holds with λ_0 formally replaced by ψ, defined by

$$\psi \sum_{q \leqslant x} q^{-1}(\log q)^2 = \sum_{q \leqslant x} q^{-1}f(q) \log q.$$

The value of $\psi \log x$ is essentially $\sqrt{2}\,(f, \hat{\psi}_1)$. The first step is to note the representation

$$(\psi - \lambda_0) \sum_{q \leqslant x} q^{-1}(\log q)^2 = \sum_{q \leqslant x} q^{-1}f_{\lambda_0}(q) \log q.$$

In this case we may replace $A_1(f, x; \lambda_0, c_0)$ by $A_0(f, x)$. Furthermore, if $1 < \beta \leqslant 2$, then in the notation of the map

$$A_\beta \colon (L^2 \cap L^{\beta'})' \to L^\beta,$$

fortunately not with $\beta = 1!$, we have

$$\left\| A_\beta(f) \right\| \asymp |\psi| + \| f - \psi \log \|.$$

The lower bound for the left-hand norm comes from (34) and Hildebrand's inequality (45), and the upper bound from our earlier bounding of the norm of A_β, noting that by the triangle inequality

$$\| f \| \leqslant \| f - \psi \log \| + \| \psi \log \|,$$

the last being readily estimated. In this presentation the number c_0 has disappeared. In view of Lemma 13,

$$\| f - \psi \log \| \asymp B_2\big(f'_{\psi,c}, x\big) + B_\beta\big(f''_{\psi,c}, x\big)$$

with $c = \| f - \psi \log \|$ which, however explicit it might be, still depends upon β.

For $\beta > 2$ the inequalities involving the truncated functions $f'_{\psi,c}$ and $f''_{\psi,c}$ cannot be derived directly from the norm on $(L^2 \cap L^{\beta'})'$ in this manner, but they do not lie any deeper than those derived from results on $L^2 \cap L^\beta$. Thus, our functional analytic treatment gives (42), with β in place of α. If we use g to denote $f - Fd_1$, and D to denote $\| A_\beta(f - Fd_1) \|$, then in particular

$$D \ll \| g \|_2 + \left(\int_{|g| > D} |g|^\beta \widehat{du} \right)^{1/\beta} + \left(D^{\beta-2} \int_{|g| \leqslant D} |g|^2 \widehat{du} \right)^{1/\beta},$$

from which a little manipulation gives

$$D \ll \| g \|_2 + \left(\int_{|g| > D} |g|^\beta \widehat{du} \right)^{1/\beta}.$$

Moreover,

$$\int_{|g| > D} 1 \, \widehat{du} \leqslant \left(\int D^{-2}|g|^2 \widehat{du} \right)^{1/2} \ll 1$$

from another application of (42). After an application of Hölder's inequality

$$\left(\int_{|g|>D} |g|^2 \widehat{du}\right)^{1/2} \ll \left(\int_{|g|>D} |g|^\beta \widehat{du}\right)^{1/\beta},$$

and

$$D \ll \left(\int_{|g|\leqslant D} |g|^2 \widehat{du}\right)^{1/2} + \left(\int_{|g|>D} |g|^\beta \widehat{du}\right)^{1/\beta}.$$

For the study of additive functions in an L^α sense, with $\alpha > 1$, the functional analytic method thus gives a flexible alternative to the application of Halász' method in the theory of Dirichlet series. Since it yields upper bounds for sums of the type $\sum_{|f(q)-Fs(q)|>\epsilon D} 1/q$ it may conveniently be combined with the method of finite probability models [11, Chapter 3]. It brings a motivation to the study of differences of additive functions, such as $f(n + 1) - f(n)$. Here the corresponding Dirichlet series would have the form

$$\sum_{n=1}^{\infty} g(n + 1)\overline{g(n)}\, n^{-s},$$

where $g(n)$ is the complex-valued multiplicative function $\exp(itf(n))$, t real. This series has no Euler product, and Halász' method cannot be applied. We have no analogue of Ruzsa's concentration function estimate, although with the $f(n)$ in the definition of $Q(x)$ replaced by $f(n + 1) - f(n)$, such an estimate is very likely valid. For the study of differences in power mean, at present only the application of the operators A_α and their duals, together with the method of approximate functional equations, is available.

As indicated in §5, approximate functional equations may be used to study the operator T on the 'spaces' L^α with $\alpha < 1$; in fact right down to a result in measure.

Besides this, the operators $A'_{\alpha'}$ are of interest for themselves. As an example:

THEOREM 17. *If* $1 < \alpha \leqslant 2$, *the inequality*

$$\left(\sum_{q\leqslant x}' q^{\alpha-1}\left| \frac{1}{x}\sum_{\substack{n\leqslant t \\ n\cong 0\,(\mathrm{mod}\,q)}} a_n - \frac{1}{q}\left(1 - \frac{1}{q_0}\right)\frac{1}{x}\sum_{n\leqslant x} a_n \right|^\alpha\right)^{1/\alpha}$$

$$+ \left(\sum'' \text{ similar summand with } \alpha = 2\right)^{1/2}$$

$$\ll \left(\frac{1}{x}\sum_{n\leqslant x} |a_n|^\alpha\right)^{1/\alpha}$$

holds for all complex numbers a_n, $1 \leqslant n \leqslant x$. *Here the first summation is to be over those prime powers* q *for which*

$$\frac{q}{x}\left| \sum_{\substack{n\leqslant x \\ n\cong 0\,(\mathrm{mod}\,q)}} a_n - \frac{1}{q}\left(1 - \frac{1}{q_0}\right)\sum_{n\leqslant x} a_n \right| > \left(\frac{1}{x}\sum_{n\leqslant x} |a_n|^\alpha\right)^{1/\alpha},$$

and the second summation '' is to be over those q for which the opposite inequality is satisfied. When $\alpha > 2$ *a similar main inequality obtains, but with the summation conditions ' and '' removed.*

These results are very useful in the study of additive or multiplicative arithmetic functions when raised to powers $\alpha > 1$.

As $x \to \infty$, the spectral radius of $A'_{\alpha'}A_\alpha$ is $3/2 + o(1)$. This generalizes the best-constant result for $\alpha = 2$ of Kubilius [28, 29]. In that case $\|A_2^*A_2\|$ coincides with the spectral radius. For $\alpha > 1$, $\alpha \neq 2$, $\|A'_{\alpha'}A_\alpha\|$ need not do so, but it will be asymptotically at least as large as $3/2$. Thus $\|A_2\| = (1 + o(1))\sqrt{3/2}$ as $x \to \infty$, but we know only $(1 + o(1))\sqrt{3/2} \leqslant \|A_\alpha\| \ll 1$ in the general case.

It is now clear where the logarithms in the study of additive arithmetic functions come from—the first eigenspace of the operator T.

References

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

2. H. Davenport, *Multiplicative number theory* (2nd ed.), Graduate Texts in Math., vol. 74, Springer-Verlag, New York, Berlin, Heidelberg, 1980.

3. F. Dress and B. Volkmann, *Ensembles d'unicité pour les fonctions arithmétiques additives ou multiplicatives*, C. R. Acad. Sci. Paris, Sér. A **287** (1978), 43–46.

4. P. D. T. A. Elliott, *On inequalities of large sieve type*, Acta Arith. **18** (1971), 405–422.

5. _____, *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.

6. _____, *On the limiting distribution of additive arithmetic functions*, Acta Math. **132** (1974), 53–75.

7. _____, *The law of large numbers for additive arithmetic functions*, Math. Proc. Cambridge Philos. Soc. **78** (1975), 33–71.

8. _____, *General asymptotic distributions for additive arithmetic functions*, Math. Proc. Cambridge Philos. Soc. **79** (1976), 43–54.

9. _____, *The Turán-Kubilius inequality*, Proc. Amer. Math. Soc. **65** (1977), 8–10.

10. _____, *On the difference of additive functions*, Mathematika **24** (1977), 153–165.

11. _____, *Probabilistic number theory*. I, II, Grundlehren math. Wiss., vols. 239, 240, Springer-Verlag, New York, Heidelberg, Berlin, 1979/80.

12. _____, *High-power analogues of the Turán-Kubilius inequality and an application to number theory*, Canad. J. Math. **32** (1980), 893–907.

13. _____, *Arithmetic functions and integer products*, Grundlehren math. Wiss., vol. 272, Springer-Verlag, New York, Berlin, Heidelberg, Tokyo, 1984/85.

14. _____, *Cauchy's functional equation in the mean*, Adv. in Math. **51** (3) (1984), 253–257.

15. _____, *The exponential function characterized by an approximate functional equation*, Illinois J. Math. **26** (1982), 503–518.

16. P. D. T. A. Elliott and C. Ryavec, *The distribution of the values of additive arithmetical functions*, Acta Math. **126** (1971), 143–164.

17. P. Erdős, *On the distribution function of additive functions*, Ann. of Math. (2) **47** (1946), 1–20.

18. P. Erdős and M. Kac, *On the Gaussian law of errors in the theory of additive functions*, Proc. Nat. Acad. Sci. U.S.A. **25** (1939), 206–207.

19. _____, *The Gaussian law of errors in the theory of additive number-theoretic functions*, Amer. J. Math. **62** (1940), 738–742.

20. G. Halász, *Über die Mittelwerte multiplikativer zahlentheoretischer Funktionen*, Acta Math. Acad. Sci. Hungar. **19** (1968), 365–403.

21. G. H. Hardy and S. Ramanujan, *The normal number of prime factors of a number n*, Quart. J. Math. (Oxford) **48** (1917), 76–92.

22. A. Hildebrand, *Sur les moments d'une fonction additive*, Ann. Inst. Fourier **33** (1983), 1–22.

23. _____, *An asymptotic formula for the variance of an additive function*, Math. Z. **183** (1983), 145–170.

24. I. Kátai, *On a problem of P. Erdős*, J. Number Theory **2** (1970), 1–6.

25. J. P. Kubilius, *Probabilistic methods in the theory of numbers*, Uspekhi Mat. Nauk (N.S.) **11** (1956), 2(68), 31–66; Amer. Math. Soc. Transl., Vol. **19** (1962), 47–85.

26. _____, *Probabilistic methods in the theory of numbers*, Trans. Math. Monographs, vol. 11, Amer. Math. Soc., Providence, R. I., 1964.

27. _____, *On an inequality for additive arithmetic functions*, Acta Arithmetica **27** (1975), 371–383.

28. _____, *On the estimation of the second central moment for strongly additive arithmetic functions*, Liet. Mat. Rinkinys = Litovsk Mat. Sbornik **23** (1983), (1) 122–123.

29. _____, *On the estimate of the second central moment for arbitrary arithmetic functions*, Liet. Mat. Rinkinys = Litovsk Mat. Sbornik **23** (1983), (2) 110–117.

30. S. Lang, *Algebra*, Addison-Wesley, Reading, Mass., 1965.

31. B. V. Levin and N. M. Timofeev, *An analytic method in the probabilistic theory of numbers*, Uchen. Zap. Vladimirskovo Gos. Ped. Inst. Mat. **57** (2) (1971), 57–150.

32. J. Meyer, *Ensembles d'unicité pour les fonctions additives. Étude analogue dans le cas des fonctions multiplicatives*, Journées de Theorie Analytique et Elémentaire des Nombres, (Orsay, 2 et 3 Juin, 1980), Publ. Math. Orsay, pp. 50–66.

33. K. Prachar, *Primzahlverteilung*, Grundlehren math. Wiss., vol. 91, Springer-Verlag, Berlin, 1957.

34. E. D. Rainville, *Special functions*, Macmillan, New York, 1960.

35. I. Z. Ruzsa, *General multiplicative functions*, Acta Arith. **32** (1977), 313–347.

36. _____, *On the concentration of additive functions*, Acta Math. Acad. Sci. Hungar. **36** (1980), 215–232.

37. _____, *On the variance of additive functions*, Turán Memorial Volume of the J. Bolyai Math. Soc., Hungary.

38. _____, *Generalized moments of additive functions*, J. Number Theory, **18** (1984), 27–33.

39. P. Turán, *On a theorem of Hardy and Ramanujan*, J. London Math. Soc. **9** (1934), 274–276.

40. _____, *Über einige Verallgemeinerungen eines Satzes von Hardy and Ramanujan*, J. London Math. Soc. **11** (1936), 125–133.

41. E. Wirsing, *A characterization of* log *n as an additive arithmetic function*, Symposia Mathematica **IV**, 45–57, Academic Press, London and New York, 1970.

42. _____, *Characterization of the logarithm as an additive function*, Proc. Sympos. Pure Math., vol. 20, Amer. Math. Soc., Providence, R. I., 1971, pp. 375–381.

43. D. Wolke, *Bemerkungen über Eindeutigkeitsmengen additiver Funktionen*, Elem. Math. **33** (1978), 14–16.

44. K. Yosida, *Functional analysis*, Grundlehren math. Wiss., vol. 123, 3rd ed., Springer-Verlag, New York, Berlin, Heidelberg, 1971.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF COLORADO, BOULDER, COLORADO 80309