CARLESON'S CONVERGENCE THEOREM FOR
DIRICHLET SERIES

Håkan Hedenmalm and Eero Saksman

CARLESON'S CONVERGENCE THEOREM FOR DIRICHLET SERIES

Håkan Hedenmalm and Eero Saksman

A Hilbert space of Dirichlet series is obtained by considering the Dirichlet series $f(s) = \sum_{n=1}^{\infty} a_n n^{-s}$ that satisfy $\sum_{n=0}^{\infty} |a_n|^2 < +\infty$. These series converge in the half plane $\mathrm{Re}\, s > \frac{1}{2}$ and define a functions that are locally L^2 on the boundary $\mathrm{Re}\, s = \frac{1}{2}$. An analog of Carleson's celebrated convergence theorem is obtained: Each such Dirichlet series converges almost everywhere on the critical line $\mathrm{Re}\, s = \frac{1}{2}$. To each Dirichlet series of the above type corresponds a "trigonometric" series $\sum_{n=1}^{\infty} a_n \chi(n)$, where χ is a multiplicative character from the positive integers to the unit circle. The space of characters is naturally identified with the infinite-dimensional torus $\mathbb{T}^{\infty}$, where each dimension comes from a a prime number. The second analog of Carleson's theorem reads: The above "trigonometric" series converges for almost all characters χ.

1. Introduction.

The study of Dirichlet series of the form $\sum_{n=1}^{\infty} a_n n^{-s}$ has a long history beginning in the nineteenth century, and the interest was due mainly to the central role that such series play in analytic number theory. The general theory of Dirichlet series was developed by Hadamard, Landau, Hardy, Riesz, Schnee, and Bohr, to name a few. As regards the modern development, we mention the work of Helson and Kahane. Helson [9, 10] should probably be credited for pioneering modern harmonic analysis methods in the theory of Dirichlet series. As a sample of Kahane's work, we mention the papers [15, 14]. Recently, in [7], Hedenmalm, Lindqvist, and Seip considered a natural Hilbert space $\mathcal{H}$ of Dirichlet series and began a systematic study thereof. The elements of $\mathcal{H}$ are analytic functions on the half-plane $\mathrm{Re}\, s > \frac{1}{2}$ of the form

$$(1.1) \qquad f(s) = \sum_{n=1}^{\infty} a_n n^{-s}$$

where the coefficients $a_1, a_2, a_3, \ldots$ are complex numbers subject to the norm boundedness condition

$$\|f\|_{\mathcal{H}} = \left(\sum_{n=1}^{\infty} |a_n|^2 \right)^{\frac{1}{2}} < +\infty.$$

In [7], the pointwise multipliers of $\mathcal{H}$ were characterized, and the result was applied to a problem of Beurling concerning 2-periodic dilation bases in $L^2([0,1])$. The reader is referred to [8] for some historical comments on the topic. In [5], Gordon and Hedenmalm followed up by characterizing the bounded composition operators of $\mathcal{H}$.

The convergence and analyticity of $f \in \mathcal{H}$ given by the series (1.1) in the half-plane $\operatorname{Re} s > \frac{1}{2}$ is a simple consequence of the Cauchy-Schwarz inequality. A deeper fact is that the boundary values of f on the 'critical' line $\operatorname{Re} s = \frac{1}{2}$ are locally L^2-functions (see [17, formula (29), p. 140] or [7, Theorem 4.11]).

Here, we establish for Dirichlet series the counterpart of the celebrated Carleson convergence theorem [1] for square summable Fourier series ($\mathbb{R}$ is the set of all real numbers):

Theorem 1.1. *Let* $\sum_{n=1}^{\infty} |a_n|^2 < +\infty$. *Then the series*

$$\sum_{n=1}^{\infty} a_n \, n^{-\frac{1}{2}+it}$$

converges for almost every $t \in \mathbb{R}$.

Let us digress on Carleson's convergence theorem in the context of a square summable Taylor series

$$(1.2) \qquad f(z) = \sum_{n=0}^{\infty} \widehat{f}(n) \, z^n, \qquad z \in \mathbb{D},$$

where $\mathbb{D}$ is the open unit disk and the coefficients satisfy

$$\|f\|_{H^2} = \left(\sum_{n=0}^{\infty} |\widehat{f}(n)|^2 \right)^{\frac{1}{2}} < +\infty.$$

The Hilbert space of such functions f is denoted by $H^2(\mathbb{D})$, and in a natural fashion, it is a closed subspace of $L^2(\mathbb{T})$, the space of (equivalence classes of) Lebesgue square summable functions g on the unit circle $\mathbb{T}$, supplied with the norm

$$\|g\|_{L^2} = \left(\int_{\mathbb{T}} |g(z)|^2 \, d\sigma(z) \right)^{\frac{1}{2}},$$

where $d\sigma$ is normalized arc length measure on $\mathbb{T}$: $d\sigma(e^{i\theta}) = d\theta/(2\pi)$. By Carleson's theorem, we have convergence in (1.2) for almost all $z \in \mathbb{T}$. And

since a general Fourier series decomposes into an analytic and an antianalytic component, the statement that Taylor series for $H^2(\mathbb{D})$ functions converge almost everywhere on $\mathbb{T}$ implies the almost everywhere convergence on $\mathbb{T}$ for Fourier series of $L^2(\mathbb{T})$ functions. We shall need the *Taylor series maximal function* associated with $f \in H^2(\mathbb{D})$,

$$\mathbf{M}f(z) = \sup_j \left| \sum_{n=j}^{\infty} \widehat{f}(n)\, z^n \right|, \qquad z \in \mathbb{T},$$

where j runs over $\{0, 1, 2, \dots\}$, and the infinite sum is interpreted in the sense of the identity

$$\sum_{n=j}^{\infty} \widehat{f}(n)\, z^n = f(z) - \sum_{n=1}^{j-1} \widehat{f}(n)\, z^n.$$

The Taylor series maximal function operator $\mathbf{M}$ is a nonlinear operator from $H^2(\mathbb{D})$ to the Lebesgue measurable functions on $\mathbb{T}$ with values in $[0, +\infty]$. Carleson's convergence theorem is a consequence of the following estimate of $\mathbf{M}$, due to Hunt [11]. The proof is essentially a modification of Carleson's original argument, which, as Hunt explains in [12], is equivalent to a weak type estimate for $\mathbf{M}$. It should be mentioned that whereas Carleson's proof is based on a careful analysis of individual square summable functions, the later proof of Fefferman [4] concentrates on analyzing the linearized maximal operator.

Theorem 1.2 (Carleson-Hunt [1, 11]). *There exists an absolute constant $A = A_{\mathrm{CH}}$ such that for every $f \in H^2(\mathbb{D})$,*

$$\int_{\mathbb{T}} |\mathbf{M}f(z)|^2\, d\sigma(z) \leq A \int_{\mathbb{T}} |f(z)|^2\, d\sigma(z).$$

A natural question is whether the absolute constant A_{CH} is astronomical. We have been assured that it is not: It can be chosen at the order of magnitude of 100. Note that in particular, $\mathbf{M}$ maps $H^2(\mathbb{D})$ into $L^2(\mathbb{T})$. To obtain the convergence theorem from the above statement, we approximate in the $H^2(\mathbb{D})$-norm the function $f \in H^2(\mathbb{D})$ by another function $g \in H^2(\mathbb{D})$ which is C^∞-smooth up to the boundary. The function g has a nicely convergent Taylor series on $\mathbb{T}$. The maximal function estimate is then applied to the difference $f - g$, and it shows that the partial sums of the Taylor series for f are uniformly close in the index parameter to those of g. Writing this down carefully, we see that the desired convergence assertion follows.

Let $K(x, \theta)$ be the kernel function

$$K(x, \theta) = \frac{e^{-i\theta x}}{ix}, \qquad x \in \mathbb{R} \setminus \{0\}, \quad \theta \in \mathbb{R},$$

which for fixed θ is interpreted as a distribution on $\mathbb{R}$ in the principal value sense.

The following dual reformulation of Theorem 1.2 appears to be essentially due to Vinogradov [**22**].

Theorem 1.3 (Strong Hilbert inequality). *Suppose φ is a compactly supported C^∞-smooth function on the set $]-\pi, \pi[\times[0, +\infty[$. Then the following estimate holds:*

$$\int_0^{+\infty} \int_0^{+\infty} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} K\big(x - y, \max(\theta_1, \theta_2)\big)\, \varphi(x, \theta_1)\, \overline{\varphi}(y, \theta_2)\, dxdy\, d\theta_1 d\theta_2$$

$$\leq B \int_{-\pi}^{\pi} \left(\int_0^{+\infty} |\varphi(x, \theta)|\, d\theta \right)^2 dx,$$

where B is an absolute constant related to the constant A_{CH} of the Carleson-Hunt theorem (Theorem 1.2).

In the above statement, the integration against the kernel in the (x, y) coordinates is to be interpreted in the weaker sense of singular integral operator theory. For instance, we can treat the integration in y as a convolution of a distribution and a smooth function, so that the rest of the integrations are well-defined in the Lebesgue sense. In this guise, we may view the Carleson-Hunt's theorem as a far-reaching generalization of Hilbert's inequality [**23**], which results by letting the functions $\varphi(x, \theta)$ tend to $\phi(x)\, \delta_0(\theta)$, where ϕ is a smooth function and δ_0 is the unit point mass at 0. For the sake of keeping the presentation as self-contained as reasonably possible, the short and simple proof of Theorem 1.3 is reproduced later on in the paper.

We return to the main topic, Dirichlet series. The point with reformulating the Carleson-Hunt theorem as Theorem 1.3 is the following. In the same fashion as the Carleson-Hunt theorem has a dual formulation, the analogous maximal function statement for Dirichlet series in the spirit of Theorem 1.1 has a similarly dual formulation with a different kernel $K^D(x, \theta)$. Surprisingly, the kernels K and K^D prove to be so similar that the inequality of Theorem 1.3 for K immediately entails the corresponding inequality with K^D (and vice versa).

We have yet another convergence result for Dirichlet series, but this time it concerns the typical convergence behavior of a function $f \in \mathcal{H}$. Given a function f of the form (1.1), we form the functions

$$(1.3) \qquad\qquad f_\chi(s) = \sum_{n=1}^{\infty} a_n\, \chi(n)\, n^{-s},$$

where $\chi(n)$ is a *character*, which means that $\chi(1) = 1$, $\chi(n) \in \mathbb{T}$ for all n, and $\chi(mn) = \chi(m)\, \chi(n)$ for all m and n. The functions f_χ are known as the *vertical limit functions* for f. The terminology is explained by the fact that

$f_\chi(s)$ is obtained from f as a limit of some of the vertical translates $f(s-it)$, with $t \in \mathbb{R}$. Each character is determined uniquely by its values on the set of primes $\mathcal{P} = \{2, 3, 5, 7, 11, \dots\}$, and the values at different primes may be chosen independently of each other. The set of all characters is denoted by Ξ, and we realize that it can be equated with the infinite-dimensional polycircle $\mathbb{T}^\infty$ by identifying each dimension with a prime number (see [**7**] for details). The polycircle $\mathbb{T}^\infty$ has a natural product probability measure defined on it, denoted $d\varpi$, the product of the normalized arc length measure $d\sigma$ in each dimension. The set of characters Ξ constitutes the dual group of the multiplicative group of positive rationals $\mathbb{Q}_+$, if the latter is given the discrete topology. The Haar probability measure on the compact group Ξ coincides with $d\varpi$. A natural question arises: Given $f \in \mathcal{H}$, what is the almost sure convergence behavior of the series (1.3) for $f_\chi(s)$, where s is a point in the complex plane, and χ is a character? The words "almost sure" refer to the Haar probability measure $d\varpi$ on Ξ. It is mentioned in [**7**] that for almost all χ, $f_\chi(s)$ extends to a holomorphic function on the right half plane $\operatorname{Re} s > 0$, and that this is best possible. In fact, in [**10**] (see also [**7**], Theorem 4.4), Helson shows that for almost all χ, the Dirichlet series (1.3) actually converges in the half-plane $\operatorname{Re} s > 0$. By Theorem 4.1 of [**7**], the function $f_\chi(it)$ makes sense as a locally L^2 summable function on the real line, for almost all χ. This makes us suspect that we have convergence in (1.3) for almost all s on the line $\operatorname{Re} s = 0$ and almost all χ. That is confirmed by the following theorem.

Theorem 1.4. *Let $f \in \mathcal{H}$ be of the form* (1.1)*, and let $f_\chi \in \mathcal{H}$ be defined by* (1.3)*. Then the series*

$$f_\chi(it) = \sum_{n=1}^\infty a_n \, \chi(n) \, n^{-it}$$

converges for almost all characters χ and almost all reals t.

An equivalent formulation of this result reads as follows.

Theorem 1.5. *Let $f \in \mathcal{H}$ be of the form* (1.1)*. Then the series*

$$\widetilde{f}(\chi) = \sum_{n=1}^\infty a_n \, \chi(n), \qquad \chi \in \Xi,$$

converges almost everywhere.

Theorem 1.5 states that square summable infinite-dimensional Taylor series converge almost everywhere on the polycircle $\mathbb{T}^\infty$ with respect to a certain order of summation (the space of such Taylor series is known as the Hardy space $H^2(\mathbb{D}^\infty)$). Here is how that works. Let p_j be the j-th prime,

and write $z_j = \chi(p_j)$; then the infinite-dimensional power series

$$F(z_1, z_2, \dots) = \sum_{n=1}^{\infty} a_n \, z_{k_1}^{\nu_1} z_{k_2}^{\nu_2} \dots z_{k_r}^{\nu_r},$$

converges for almost all $(z_1, z_2, z_3, \dots) \in \mathbb{T}^\infty$, where k_j, ν_j, and n are related via the prime number factorization of n: $n = p_{k_1}^{\nu_1} p_{k_2}^{\nu_2} \dots p_{k_r}^{\nu_r}$. The order of summation, of course, is dictated by the index n: The condition of summing up to index N is expressed by

$$\log n = \nu_1 \log p_{k_1} + \nu_2 \log p_{k_2} + \dots + \nu_r \log p_{k_r} \le \log N,$$

so that the summation is cut by a single hyperplane in the index plane. This permits us to apply a technique devised by Fefferman [**3**] for finitely many cuts in a finite-dimensional setting to the infinite-dimensional case, and obtain the L^2 maximal function estimate here as well. Fefferman's idea is to reduce the situation to the one-dimensional Fourier series case, where the Carleson-Hunt theorem (Theorem 1.2) applies.

Theorem 1.5 yields as a consequence nontrivial estimates for the almost sure growth behavior of partial sums of random characters, a question which was considered in [**7**]. It follows that almost surely,

$$\sum_{n=1}^{N} \chi(n) = O\left(\sqrt{N \log N} \left(\log \log N\right)^{1/2+\varepsilon}\right), \qquad \text{as } N \to +\infty.$$

Finding the best possible growth bound for the almost sure behavior of these partial sums has an unmistakable Erdös-type flavor, in its combination of probability and number theory. And sure enough, in [**2**, pp. 251-252], Erdös states as a problem to determine the almost sure growth of the analoguous sums, where the $\chi(p)$ for prime indices p are replaced by independent random variables assuming the values ± 1 with equal probabilities $\frac{1}{2}$. Erdös looks to compare the growth of the partial sums with the classical law of the iterated logarithm, where all the terms $\chi(n)$ are independent and take values ± 1 with equal probabilities $\frac{1}{2}$. In Erdös' problem, as in ours, the characters have the multiplicative property $\chi(mn) = \chi(m)\chi(n)$, which reduces the randomness and introduces a number-theoretic ingredient. A complete solution should thus shed light on the multiplicative structure of the integers.

2. Carleson-Hunt's theorem and duality.

The Hardy space $H^2(\mathbb{D})$ was introduced earlier in terms of Taylor coefficients. Here we mention that a function f holomorphic in $\mathbb{D}$ is in $H^2(\mathbb{D})$ if and only if

$$\|f\|_{H^2} = \sup_{0 < r < 1} \left(\int_{\mathbb{T}} |f(r\zeta)|^2 \, d\sigma(\zeta)\right)^{\frac{1}{2}} < +\infty,$$

and that the values of f are well-defined almost everywhere on the boundary $\mathbb{T}$. The norm of f in $H^2(\mathbb{D})$ then equals the $L^2(\mathbb{T})$ norm of the boundary function. Moreover, a function in $L^2(\mathbb{T})$ is in $H^2(\mathbb{D})$ if and only if its harmonic extension (via the Poisson integral) to the interior is holomorphic.

To simplify the notation, we identify the unit circle $\mathbb{T}$ with the interval $]-\pi, \pi]$, where topologically the endpoints are tied together. Our point of departure is the fundamental Carleson-Hunt estimate of the maximal function, as stated in Theorem 1.2. We intend to find a dual reformulation of this result. To that end, take an $f \in H^2(\mathbb{D})$, with Taylor series expansion

$$f(z) = \sum_{n=0}^{\infty} \widehat{f}(n)\, z^n, \qquad z \in \mathbb{D},$$

and observe that by the $l^1 - l^\infty$ duality, we may write, for a given fixed $x \in]-\pi, \pi]$,

$$\mathbf{M}f(x) = \sup_{\phi_j} \operatorname{Re} \sum_{j=0}^{\infty} \overline{\phi}_j \sum_{n=j}^{\infty} \widehat{f}(n)\, e^{inx} = \sup_{\phi_j} \operatorname{Re} \sum_{n=0}^{\infty} \widehat{f}(n)\, e^{inx} \sum_{j=0}^{n} \overline{\phi}_j,$$

where the supremum runs over all complex-valued sequences $\{\phi_j\}_j$ with

$$(2.1) \qquad \sum_{j=0}^{\infty} |\phi_j| \le 1.$$

We now let x vary and allow the ϕ_j to depend on x. Keeping in mind that the function f is fixed, it is easy to check that given an $\varepsilon > 0$, we may find Borel measurable functions $\phi_j \in L^\infty(]-\pi, \pi])$ with the analog of (2.1),

$$(2.2) \qquad \sum_{j=0}^{\infty} |\phi_j(x)| \le 1, \qquad x \in]-\pi, \pi],$$

such that

$$\operatorname{Re} \sum_{n=0}^{\infty} \widehat{f}(n)\, e^{inx} \sum_{j=0}^{n} \overline{\phi}_j(x) = \operatorname{Re} \sum_{j=0}^{\infty} \overline{\phi}_j(x) \sum_{n=j}^{\infty} \widehat{f}(n)\, e^{inx} \ge (1-\varepsilon)\, \mathbf{M}f(x)$$

almost everywhere on $]-\pi, \pi]$. It follows that we may calculate the $L^2(]-\pi, \pi])$-norm of the maximal function in the following fashion:

$$(2.3) \qquad \|\mathbf{M}f\|_{L^2} = \sup_{g} \operatorname{Re} \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbf{M}f(x)\, \overline{g}(x)\, dx$$

$$= \sup_{g, \phi_j} \operatorname{Re} \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{n=0}^{\infty} \widehat{f}(n)\, e^{inx} \sum_{j=0}^{n} \overline{\phi}_j(x)\, \overline{g}(x)\, dx$$

$$= \sup_{g, \phi_j} \operatorname{Re} \sum_{n=0}^{\infty} \widehat{f}(n) \sum_{j=0}^{n} \overline{\widehat{\phi_j\, g}(n)},$$

where g ranges over all functions in $L^2(]-\pi,\pi])$ with norm ≤ 1, and ϕ_j ranges over all sequences of $L^\infty(]-\pi,\pi])$ functions with (2.2). To obtain the "norm" $\|\mathbf{M}\|$ of the nonlinear operator $\mathbf{M}$, which by the Carleson-Hunt theorem is bounded by an absolute constant ($\sqrt{A_{\mathrm{CH}}}$), we form the supremum of $\|\mathbf{M}f\|_{L^2}$ over all functions $f \in H^2(\mathbb{D})$ of the norm ≤ 1. Taking the supremum over all such f of the last expression in (2.3),

$$\mathrm{Re} \sum_{n=0}^{\infty} \widehat{f}(n) \sum_{j=0}^{n} \overline{\widehat{\phi_j\, g}(n)},$$

we obtain

$$\left(\sum_{n=0}^{\infty} \left| \sum_{j=0}^{n} \widehat{\phi_j\, g}(n) \right|^2 \right)^{\frac{1}{2}},$$

and hence the assertion of the Carleson-Hunt theorem reads

$$(2.4) \qquad \sup_{g,\phi_j} \sum_{n=0}^{\infty} \left| \sum_{j=0}^{n} \widehat{\phi_j\, g}(n) \right|^2 \leq A_{\mathrm{CH}},$$

where the supremum is, as before, taken over all $g \in L^2(]-\pi,\pi])$ of unit norm and all sequences of bounded measurable functions ϕ_j with (2.2). Now consider functions $\psi_j \in L^2(]-\pi,\pi])$, such that

$$(2.5) \qquad \int_{-\pi}^{\pi} \left(\sum_{j=0}^{\infty} |\psi_j(x)| \right)^2 \frac{dx}{2\pi} \leq 1.$$

Put $g = \sum_j |\psi_j| \in L^2(]-\pi,\pi])$, which then has norm at most 1 in $L^2(]-\pi,\pi])$, and observe that the functions ϕ_j defined by the relations $\phi_j(x) = \psi_j(x)/g(x)$ if $g(x) > 0$, and $\phi_j(x) = 0$ otherwise, meet the requirement (2.1). It follows from (2.4) that we have

$$(2.6) \qquad \sup_{\psi_j} \sum_{n=0}^{\infty} \left| \sum_{j=0}^{n} \widehat{\psi_j}(n) \right|^2 \leq A_{\mathrm{CH}},$$

the supremum being taken over all sequences ψ_j with (2.5). Clearly, (2.6) generalizes (2.4). By dropping the scaling restriction (2.5), we end up with the following reformulation of Theorem 1.2.

Theorem 2.1. *For all sequences ψ_j of functions in $L^2(]-\pi,\pi])$, we have the inequality*

$$\sum_{n=0}^{\infty} \left| \sum_{j=0}^{n} \widehat{\psi_j}(n) \right|^2 \leq A \int_{-\pi}^{\pi} \left(\sum_{j=0}^{\infty} |\psi_j(x)| \right)^2 \frac{dx}{2\pi},$$

where $A = A_{\mathrm{CH}}$ is the constant of the Carleson-Hunt theorem.

In the above theorem, the assertion is trivial if the right-hand side assumes the value $+\infty$. It is clear that we have obtained an honest reformulation of the Carleson-Hunt theorem, because we may run the argument backwards and obtain Theorem 1.2 out of Theorem 2.1. The only thing we were sloppy about was that we did not justify that we could change the order of summation at a certain point of the argument. However, by first restricting our attention to, say polynomials f, the general case follows by approximation.

We carry on to write the above theorem in the form mentioned in the introduction (Theorem 1.3). Recall that in the statement, φ is a C^∞-smooth compactly supported function on $]-\pi, \pi[\times[0, +\infty[$. For $j = 0, 1, 2, \ldots$, we let the functions ψ_j be defined by

$$\psi_j(x) = \int_j^{j+1} \varphi(x, \theta)\, d\theta, \qquad x \in [-\pi, \pi],$$

so that by Theorem 2.1,

$$(2.7) \qquad \int_{[0,+\infty[} \left| \int_{-\pi}^{\pi} e^{-itx} \int_0^{t+1} \varphi(x, \theta)\, d\theta\, \frac{dx}{2\pi} \right|^2 d\#(t)$$

$$\leq A_{\mathrm{CH}} \int_{-\pi}^{\pi} \left(\sum_{j=0}^{\infty} \left| \int_j^{j+1} \varphi(x, \theta)\, d\theta \right| \right)^2 \frac{dx}{2\pi},$$

where $d\#$ stands for the counting measure on the nonnegative integers $\{0, 1, 2, \ldots\}$. By Minkowski's inequality,

$$\sum_{j=0}^{\infty} \left| \int_j^{j+1} \varphi(x, \theta)\, d\theta \right| \leq \int_0^{+\infty} |\varphi(x, \theta)|\, d\theta,$$

so that (2.7) implies that

$$(2.8) \qquad \int_{[0,+\infty[} \left| \int_{-\pi}^{\pi} e^{-itx} \int_0^{t+1} \varphi(x, \theta)\, d\theta\, \frac{dx}{2\pi} \right|^2 d\#(t)$$

$$\leq A_{\mathrm{CH}} \int_{-\pi}^{\pi} \left(\int_0^{+\infty} |\varphi(x, \theta)|\, d\theta \right)^2 \frac{dx}{2\pi}.$$

Potentially this statement might be weaker than the estimate of Theorem 2.1. However, by making the smooth function $\varphi(x, \theta)$ suitably approximate the sum of point masses

$$\sum_{j=0}^{\infty} \psi_j(x)\, \delta_j(\theta),$$

where δ_j is the unit point mass at j, we see that the two inequalities are of equal strength.

By expanding the square and changing the order of integration in (2.8), the inequality may be written in the formally equivalent form

$$\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{0}^{+\infty}\int_{0}^{+\infty}\int_{]\max(\theta_1,\theta_2)-1,+\infty[} e^{-it(x-y)}d\#(t)$$
$$\cdot\,\varphi(x,\theta_1)\,\overline{\varphi}(x,\theta_2)\,d\theta_1\,d\theta_2\,\frac{dxdy}{4\pi^2}$$
$$\leq A_{\mathrm{CH}}\int_{-\pi}^{\pi}\left(\int_{0}^{+\infty}|\varphi(x,\theta)|\,d\theta\right)^2\frac{dx}{2\pi}.$$

However, it is not obvious that the change of order in the integration is permitted. In addition, the innermost integral calls for an interpretation in the sense of distribution theory. In order to avoid such complications, we introduce a smoothing parameter ε, $0 < \varepsilon < \frac{1}{2}$. We write

$$d\#_\varepsilon(t) = e^{-\varepsilon t}\,d\#(t) = \sum_{j=0}^{\infty} e^{-\varepsilon j}\,d\delta_j(t),$$

and observe that since $d\#_\varepsilon$ is smaller than $d\#$, (2.8) implies that

$$(2.9) \qquad \int_{[0,+\infty[}\left|\int_{-\pi}^{\pi} e^{-itx}\int_{0}^{t+1}\varphi(x,\theta)\,d\theta\,\frac{dx}{2\pi}\right|^2 d\#_\varepsilon(t)$$
$$\leq A_{\mathrm{CH}}\int_{-\pi}^{\pi}\left(\int_{0}^{+\infty}|\varphi(x,\theta)|\,d\theta\right)^2\frac{dx}{2\pi}.$$

In this case, we may appeal to the Fubini theorem, and arrive at

$$(2.10) \qquad \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{0}^{+\infty}\int_{0}^{+\infty}\int_{]\max(\theta_1,\theta_2)-1,+\infty[} e^{-it(x-y)}d\#_\varepsilon(t)$$
$$\cdot\,\varphi(x,\theta_1)\overline{\varphi}(y,\theta_2)\,d\theta_1 d\theta_2\,\frac{dxdy}{4\pi^2}$$
$$\leq A_{\mathrm{CH}}\int_{-\pi}^{\pi}\left(\int_{0}^{+\infty}|\varphi(x,\theta)|\,d\theta\right)^2\frac{dx}{2\pi}.$$

We sum the geometric series which actually appears on the left-hand side of the above inequality, and obtain

$$\int_{]\max(\theta_1,\theta_2)-1,+\infty[} e^{-it(x-y)}d\#_\varepsilon(t) = \frac{e^{-[\max(\theta_1,\theta_2)](\varepsilon+i(x-y))}}{1-e^{-(\varepsilon+i(x-y))}},$$

where $[\cdot]$ denotes the operation of taking the integer part. Let the kernel $K_\varepsilon^{\#}$ be given by

$$K_\varepsilon^{\#}(x,\theta) = \frac{e^{-[\theta](\varepsilon+ix)}}{1-e^{-(\varepsilon+ix)}},$$

in terms of which (2.10) simplifies:

(2.11)
$$\int_0^{+\infty} \int_0^{+\infty} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} K_\varepsilon^\# \big(x - y, \max(\theta_1, \theta_2)\big)\, \varphi(x, \theta_1)\overline{\varphi}(y, \theta_2)\, \frac{dxdy}{4\pi^2}\, d\theta_1 d\theta_2$$
$$\le A_{\mathrm{CH}} \int_{-\pi}^{\pi} \left(\int_0^{+\infty} |\varphi(x, \theta)|\, d\theta \right)^2 \frac{dx}{2\pi}.$$

The left-hand side of (2.9) increases as ε decreases, so that the left-hand side of (2.11) must increase with decreasing ε as well. Letting ε tend to 0, we therefore obtain the strongest form of (2.11). The kernel $K_\varepsilon^\#$ then tends to the distribution

$$K^\#(x, \theta) + \pi \sum_{n=-\infty}^{\infty} \delta_{2\pi n}(x),$$

where $\delta_{2\pi n}$ is the unit point mass at $2\pi n$, and

$$K^\#(x, \theta) = \ \mathrm{pv}\ \frac{e^{-i[\theta]x}}{1 - e^{-ix}} = e^{-i[\theta]x}\ \mathrm{pv}\ \frac{1}{1 - e^{-ix}},$$

the "pv" standing for the principal value operation (in the x variable). Given the smoothness of $\varphi(x, \theta)$, and the fact that in the x variable, it is supported inside $]-\pi, \pi[$, the integral

$$\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} K_\varepsilon^\# \big(x - y, \max(\theta_1, \theta_2)\big)\, \varphi(x, \theta_1)\overline{\varphi}(y, \theta_2)\, \frac{dxdy}{4\pi^2}$$

tends to the expression

$$\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} K^\# \big(x - y, \max(\theta_1, \theta_2)\big)\, \varphi(x, \theta_1)\overline{\varphi}(y, \theta_2)\, \frac{dxdy}{4\pi^2}$$
$$+ \int_{-\pi}^{\pi} \varphi(x, \theta_1)\, \overline{\varphi}(x, \theta_2)\, \frac{dx}{4\pi}$$

as $\varepsilon \to 0$, where the singular integral is well-defined as either a distributional convolution in x first, and then an ordinary Lebesgue integral in y, or we may reverse the order of x, y; it doesn't matter which we choose to do, because the integral comes out the same. The above convergence (as $\varepsilon \to 0$) is easily seen to be uniform in (θ_1, θ_2), and because of the assumption that φ has compact support, the integral on the left-hand side of (2.11) tends to

$$\int_0^{+\infty} \int_0^{+\infty} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} K^\# \big(x - y, \max(\theta_1, \theta_2)\big)\, \varphi(x, \theta_1)\overline{\varphi}(y, \theta_2)\, \frac{dxdy}{4\pi^2}\, d\theta_1 d\theta_2$$
$$+ \int_{-\pi}^{\pi} \left| \int_0^{+\infty} \varphi(x, \theta)\, d\theta \right|^2 \frac{dx}{4\pi}.$$

It follows that we have obtained yet another reformulation of the Carleson-Hunt theorem.

Theorem 2.2. *Suppose φ is a compactly supported C^∞-smooth function on the set $]-\pi, \pi[\times[0, +\infty[$. Then the following estimate holds:*

$$\int_0^{+\infty} \int_0^{+\infty} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} K^\#\big(x - y, \max(\theta_1, \theta_2)\big) \varphi(x, \theta_1) \overline{\varphi}(y, \theta_2) \frac{dxdy}{4\pi^2} \, d\theta_1 d\theta_2$$

$$\leq A \int_{-\pi}^{\pi} \left(\int_0^{+\infty} |\varphi(x, \theta)| \, d\theta \right)^2 \frac{dx}{2\pi},$$

where $A = A_{\mathrm{CH}}$ is the absolute constant of Theorem 1.2.

We remark that when we derive the Carleson-Hunt theorem from Theorem 2.2, we get a slightly worse constant, because we dropped a term for æsthetic reasons.

The property of the kernel $K^\#$ which ensures that the integral expression on the left-hand side of the inequality of the above theorem is real-valued can be formalized: It is the symmetry requirement of a given kernel $K(x, \theta)$ that

$$\overline{K}(x, \theta) = K(-x, \theta).$$

Let $K(x, \theta)$ have the above symmetry property and suppose that it is uniformly within finite distance from $K^\#$ on a symmetric interval:

$$\big|K(x, \theta) - K^\#(x, \theta)\big| \leq C, \qquad (x, \theta) \in \,]-2\alpha, 2\alpha[\times[0, +\infty[,$$

where α and C are a positive real numbers. Just as $K^\#$, we define K as a principal value distribution on $]-2\alpha, 2\alpha[$. Let α be confined to the interval $0 < \alpha < \frac{1}{2}\pi$, so that we have the above inequality fulfilled for the particular choice

$$K(x, \theta) = \mathrm{pv} \, \frac{e^{-i\theta x}}{ix} = e^{-i\theta x} \, \mathrm{pv} \, \frac{1}{ix}, \qquad x \in \mathbb{R} \setminus \{0\}, \quad \theta \in \mathbb{R}.$$

Then since

$$\int_0^{+\infty} \int_0^{+\infty} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |\varphi(x, \theta_1)| \, |\varphi(y, \theta_2)| \frac{dxdy}{4\pi^2} \, d\theta_1 d\theta_2$$

$$= \left(\int_0^{+\infty} \int_{-\pi}^{\pi} |\varphi(x, \theta)| \frac{dx}{2\pi} \, d\theta \right)^2$$

$$\leq \int_{-\pi}^{\pi} \left(\int_0^{+\infty} |\varphi(x, \theta)| \, d\theta \right)^2 \frac{dx}{2\pi},$$

an application of Theorem 2.2 yields that

(2.12)

$$\int_0^{+\infty} \int_0^{+\infty} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} K\big(x - y, \max(\theta_1, \theta_2)\big) \varphi(x, \theta_1) \overline{\varphi}(y, \theta_2) \frac{dxdy}{4\pi^2} \, d\theta_1 d\theta_2$$

$$\leq (A + C) \int_{-\pi}^{\pi} \left(\int_0^{+\infty} |\varphi(x, \theta)| \, d\theta \right)^2 \frac{dx}{2\pi},$$

provided that the compact support of the function φ is contained in $]-\alpha, \alpha[\times[0, +\infty[$. An analysis of the derivation of Theorem 2.2 reveals that the estimate of that theorem, restricted to functions φ supported on $]-\alpha, \alpha[\times[0, +\infty[$, is equivalent to the following estimate for the maximal function:

$$(2.13) \qquad \int_{-\alpha}^{\alpha} |\mathbf{M}f(x)|^2 \frac{dx}{2\pi} \leq A' \int_{-\pi}^{\pi} |f(x)|^2 \frac{dx}{2\pi}, \qquad f \in H^2(\mathbb{D}),$$

for some absolute constant A'. By rotation invariance, this estimate is equipotent with the estimate of the Carleson-Hunt theorem, albeit that it leads to a worse absolute constant. In other words, we have the following reformulation of Theorem 1.2, if we accept that the size of the absolute constant is unimportant.

Theorem 2.3. *Suppose* $0 < \alpha < \frac{1}{2}\pi$*, and that* φ *is a compactly supported* C^∞*-smooth function on the set* $]-\alpha, \alpha[\times[0, +\infty[$*, extended to vanish elsewhere. Then the following estimate holds:*

$$\int_0^{+\infty} \int_0^{+\infty} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} K\big(x-y, \max(\theta_1, \theta_2)\big)\, \varphi(x, \theta_1)\, \overline{\varphi}(y, \theta_2)\, dxdy\, d\theta_1 d\theta_2$$
$$\leq B \int_{-\pi}^{\pi} \bigg(\int_0^{+\infty} |\varphi(x, \theta)|\, d\theta \bigg)^2 dx,$$

where B *is an absolute constant.*

In the formulation in the introduction, φ was allowed to have compact support in $]-\pi, \pi[\times[0, +\infty[$. To cover that case, we recall the specific choice of K there, and make a suitable dilation in both variables (x, θ).

Theorem 2.3 suggests possible generalizations of the maximal function estimate of the Carleson-Hunt theorem. For instance, what if we replace the L^1 norm on the right-hand side by a slightly weaker norm expression?

Remark 2.4. In the sharp constant form, Hilbert's inequality states that

$$\bigg| \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \mathrm{pv}\, \frac{1}{i(x-y)}\, \varphi(x)\, \overline{\varphi}(y)\, dxdy \bigg| \leq \pi \int_{-\pi}^{\pi} |\varphi(x)|^2\, dx,$$

and there is a variant which applies to the bilinear form with two functions φ, ψ. As we apply that version to the left-hand side of the expression in Theorem 2.3, we obtain

$$\bigg| \int_0^{+\infty} \int_0^{+\infty} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} K\big(x-y, \max(\theta_1, \theta_2)\big)\, \varphi(x, \theta_1)\, \overline{\varphi}(y, \theta_2)\, dxdy\, d\theta_1 d\theta_2 \bigg|$$
$$\leq \pi \bigg(\int_0^{+\infty} \bigg(\int_{-\pi}^{\pi} |\varphi(x, \theta)|^2\, dx \bigg)^{\frac{1}{2}} d\theta \bigg)^2,$$

which estimate differs from that of Theorem 2.3 in that the L^2 in the x variable is taken first, and the L^1 norm in θ second. Of course this estimate is worse (if we forget about the size of the constant, that is), because the triangle inequality guarantees that taking the L^1 norm in θ first and then the L^2 norm in x produces a smaller quantity.

3. Proof of Theorem 1.1.

We assume that

$$\sum_{n=1}^{\infty} |a_n|^2 < +\infty,$$

where $a_1, a_2, a_3, \dots$ is a complex-valued sequence, and study the convergence properties of the Dirichlet series

$$(3.1) \qquad f(t) = \sum_{n=1}^{\infty} a_n \, n^{-\frac{1}{2}-it}, \qquad t \in \mathbb{R}.$$

We have changed the notation a little: Previously, the above function would have been denoted by $f(\frac{1}{2} + it)$. We mentioned in the introduction that f makes sense as a locally square-integrable function on $\mathbb{R}$ (see [7]). The corresponding Dirichlet series maximal function is defined by

$$\mathbf{M}_D f(t) = \sup_j \left| \sum_{n=j}^{\infty} a_n \, n^{-\frac{1}{2}-it} \right|, \qquad t \in \mathbb{R},$$

where j ranges over $\{1, 2, 3, \dots\}$, and the infinite sum is interpreted as

$$\sum_{n=j}^{\infty} a_n \, n^{-\frac{1}{2}-it} = f(t) - \sum_{n=1}^{j-1} a_n \, n^{-\frac{1}{2}-it}.$$

By translation invariance, we need only prove the almost everywhere convergence on the interval $[-1, 1]$. We will do this by obtaining the following maximal function estimate:

$$(3.2) \qquad \int_{-1}^{1} \left| \mathbf{M}_D f(t) \right|^2 dt \leq B \sum_{n=1}^{\infty} |a_n|^2,$$

where B is an absolute constant. The derivation of the convergence statement then follows from a standard argument (see [21, Proposition 6.2]), which was used for Taylor series as well: Approximate f by a Dirichlet polynomial g such that the norm of the difference is small in $\mathcal{H}$, and apply the maximal function estimate to the difference, whence the assertion follows by inspection.

We proceed exactly as in the proof of the first dual formulation (Theorem 2.1) of the Carleson-Hunt theorem. We obtain

(3.3)

$$\left(\int_{-1}^{1} \left| \mathbf{M}_D f(t) \right|^2 dt \right)^{\frac{1}{2}} = \sup_g \int_{-1}^{1} \mathbf{M}_D f(t)\, \overline{g}(t) dt$$

$$= \sup_{g,\phi_j} \operatorname{Re} \int_{-1}^{1} \sum_{n=1}^{\infty} a_n\, n^{-\frac{1}{2}} e^{it \log n} \sum_{j=1}^{n} \phi_j(t)\, \overline{g}(t)\, dt$$

$$= \sup_{g,\phi_j} \operatorname{Re} \sum_{n=1}^{\infty} a_n \sum_{j=1}^{n} n^{-\frac{1}{2}} \overline{\widehat{\phi_j g}}(\log n),$$

where g is restricted by the condition that it is Lebesgue measurable and

$$\int_{-1}^{1} |g(t)|^2 dt \le 1,$$

and ϕ_j is a sequence of bounded Lebesgue measurable functions with

$$\sum_{j=1}^{\infty} |\phi_j(t)| \le 1, \qquad t \in [-1, 1].$$

These functions are extended to vanish off $[-1, 1]$, and the Fourier transform appearing in the above formula is given by

$$\widehat{\varphi}(\xi) = \int_{-\infty}^{+\infty} e^{-it\xi}\, \varphi(t)\, dt, \qquad \xi \in \mathbb{R}.$$

As in the previous Taylor series case, we have some difficulty motivating why we may change the order of summation, and we resolve the difficulty by assuming that only finitely many of the coefficients a_n are different from 0. Such functions (Dirichlet polynomials) are dense in $\mathcal{H}$, and it suffices to obtain the maximal function estimate (3.2) for them. It follows from the Cauchy-Schwarz inequality applied to (3.3) that we just need to obtain the estimate

$$\sum_{n=1}^{\infty} \frac{1}{n} \left| \sum_{j=1}^{n} \widehat{\phi_j g}(\log n) \right|^2 \le B,$$

assuming that g and ϕ_j are as above. As before we find that this is equivalent to establishing for sequences $\psi_1, \psi_2, \psi_3, \ldots$ of Lebesgue measurable functions on $[-1, 1]$ the inequality

$$\sum_{n=1}^{\infty} \frac{1}{n} \left| \sum_{j=1}^{n} \widehat{\psi_j}(\log n) \right|^2 \le B \int_{0}^{1} \left(\sum_{j=1}^{\infty} |\psi_j(t)| \right)^2 dt.$$

We again proceed as in Section 3, with the counting measure $d\#$ on the nonnegative integers replaced by the sum of point masses

$$\sum_{n=1}^{\infty} \frac{1}{n} \delta_{\log n},$$

and obtain that the above estimate is equivalent to having, for C^{∞}-smooth compactly supported functions φ on $]-1,1[\times[0,+\infty[$,

(3.4)

$$\int_0^{+\infty} \int_0^{+\infty} \int_{-1}^1 \int_{-1}^1 K^D\big(x-y, \max(\theta_1, \theta_2)\big)\, \varphi(x, \theta_1)\, \overline{\varphi}(y, \theta_2)\, dx dy\, d\theta_1 d\theta_2$$

$$\leq B' \int_0^1 \left(\int_0^{+\infty} |\varphi(x, \theta)|\, d\theta \right)^2 dx,$$

where B' is some absolute constant. The kernel K^D is essentially the distributional limit of the kernels K_ε^D as $0 < \varepsilon \to 0$ (we want K^D to be a principal value distribution, and the limit of the kernels K_ε^D contains a point mass at the origin, which needs to be removed), where, for ε, $0 < \varepsilon < +\infty$, K_ε^D is given by the expression

$$K_\varepsilon^D(x, \theta) = \sum_{n=[e^\theta]}^{\infty} n^{-1-\varepsilon-ix}.$$

We approximate the sum by the corresponding integral expressions,

$$K_\varepsilon^D(x, \theta) - \frac{e^{-\theta(\varepsilon+ix)}}{\varepsilon + ix}$$

$$= \int_{[e^\theta, +\infty[} t^{-1-\varepsilon-ix} \{d\#(t) - dt\}$$

$$= e^{-(1+\varepsilon+ix)\theta} \big(e^\theta - [e^\theta]'\big) - (1+\varepsilon+ix) \int_{e^\theta}^{+\infty} t^{-2-\varepsilon-ix} \big(t - [t]\big)\, dt,$$

where we rely on integration by parts. The notation $[x]'$ stands for the integer part of x, except that if x is an integer, we get $x-1$ instead of x. It follows that we have

$$\left| K_\varepsilon^D(x, \theta) - \frac{e^{-\theta(\varepsilon+ix)}}{\varepsilon + ix} \right| \leq 4, \qquad 0 \leq \theta < +\infty, \ -2 \leq x \leq 2.$$

As $0 < \varepsilon \to 0$, the expression

$$\frac{e^{-\theta(\varepsilon+ix)}}{\varepsilon + ix}$$

tends to the distribution $K(x, \theta) + \pi \delta_0(x)$, where K is as before:

$$K(x, \theta) = \mathrm{pv} \, \frac{e^{-i\theta x}}{ix} = e^{-i\theta x} \, \mathrm{pv} \, \frac{1}{ix}.$$

Let K_*^D be the distributional limit of the kernels K_ε^D as $0 < \varepsilon \to 0$, and put

$$K^D(x, \theta) = K_*^D(x, \theta) - \pi \delta_0(x),$$

so that K^D becomes a principal value distribution near $x = 0$. It suffices to obtain (3.4) for this particular kernel. By the above, K^D is uniformly close to the kernel K:

$$\left| K^D(x, \theta) - K(x, \theta) \right| \le 4, \qquad (x, \theta) \in [-2, 2] \times [0, +\infty[.$$

The desired boundedness is now a consequence of Theorem 2.3.

4. Proof of Theorem 1.5.

As in the previous section, $a_1, a_2, a_3, \dots$ is a sequence of complex numbers, subject to the square summability condition

$$\sum_{n=1}^{\infty} |a_n|^2 < +\infty.$$

The function we shall study is the infinite power series

$$(4.1) \qquad f(\chi) = \sum_{n=1}^{\infty} a_n \, \chi(n), \qquad \chi \in \Xi,$$

which defines a square summable function on Ξ, by standard Fourier analysis (see [7]):

$$\|f\|_{L^2(\Xi)}^2 = \int_\Xi |f(\chi)|^2 \, d\varpi(\chi) = \sum_{n=1}^{\infty} |a_n|^2.$$

The space of all such f is denoted by $H^2(\Xi)$, and is in a natural way the Hardy space on the infinite-dimensional polydisk $\mathbb{D}^\infty$. The corresponding maximal function is

$$(4.2) \qquad \mathbf{M}_\Xi f(\chi) = \sup_j \left| \sum_{n=j}^{\infty} a_n \, \chi(n) \right|, \qquad \chi \in \Xi,$$

where j ranges over $\{1, 2, 3, \dots\}$, and the infinite sum is interpreted as

$$\sum_{n=j}^{\infty} a_n \, \chi(n) = f(\chi) - \sum_{n=1}^{j-1} a_n \, \chi(n).$$

As before the claimed almost convergence follows as soon as we obtain the estimate

$$
(4.3) \qquad \left\| \mathbf{M}_\Xi f \right\|_{L^2(\Xi)}^2 \leq B \left\| f \right\|_{L^2(\Xi)}^2,
$$

for some finite absolute constant B, by a standard approximation argument: We take a Dirichlet polynomial which approximates f in norm, apply the estimate (4.3) to the difference, and the almost everywhere convergence assertion follows.

For $N = 1, 2, 3, \ldots$, we denote by $\mathcal{N}_N$ the set of all positive integers whose prime factorizations contain only the first N primes $p_1, p_2, \ldots, p_N$. We define the corresponding space $H_N^2(\Xi)$ consisting of all functions $f \in H^2(\Xi)$ with series expansion (4.1) for which $a_n = 0$ unless $n \in \mathcal{N}_N$. If we put $z_j = \chi(p_j)$, and think of $f \in H_N^2(\Xi)$ as a function of $(z_1, z_2, \ldots, z_N)$, the series expansion (4.1) can be written as

$$
f(z_1, z_2, \ldots, z_N) = \sum_{k_1, \ldots, k_N = 0}^{\infty} a_n \, z_1^{k_1} \ldots z_N^{k_N}, \qquad n = p_1^{k_1} \ldots p_N^{k_N},
$$

so that $H^2(\Xi)$ may be identified with $H^2(\mathbb{D}^N)$, the Hardy space on the finite-dimensional polydisk $\mathbb{D}^N$. We shall obtain the maximal function estimate (4.3) for functions $f \in H_N^2(\Xi)$, with a constant B independent of N. The assertion (4.3) then follows in general by a standard approximation argument.

We write $z_j = e^{i\theta_j}$, where θ_j is a real parameter, and

$$
a_{k_1, \ldots, k_N} = a_n \qquad \text{provided that} \quad n = p_1^{k_1} \ldots p_N^{k_N},
$$

and extend $a_{k_1, \ldots, k_N}$ to vanish whenever one of the indices $k_1, \ldots, k_N$ is negative. Modulo a slight abuse of notation, we then have

$$
f(\theta_1, \ldots, \theta_N) = \sum_{k_1, \ldots, k_N = 0}^{\infty} a_{k_1, \ldots, k_N} \, e^{ik_1\theta_1 + \cdots + ik_N\theta_N}, \qquad n = p_1^{k_1} \ldots p_N^{k_N}.
$$

The corresponding maximal function is

$$
\mathbf{M}_N f(\theta_1, \ldots, \theta_N) = \sup_j \left| \sum_{(k_1, \ldots, k_N) \in \mathcal{G}(\log j, N)} a_{k_1, \ldots, k_N} \, e^{ik_1\theta_1 + \cdots + ik_N\theta_N} \right|,
$$

where $\mathcal{G}(T, N)$ consists of all vectors $(k_1, \ldots, k_N)$ where the entries are nonnegative integers, and the following inequality holds:

$$
T \leq k_1 \log p_1 + k_2 \log p_2 + \cdots + k_N \log p_N.
$$

We need to prove that there exists an absolute constant B such that

$$(4.4) \quad \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \left| \mathbf{M}_N f(\theta_1, \ldots, \theta_N) \right|^2 \frac{d\theta_1}{2\pi} \cdots \frac{d\theta_N}{2\pi}$$

$$\leq B \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \left| f(\theta_1, \ldots, \theta_N) \right|^2 \frac{d\theta_1}{2\pi} \cdots \frac{d\theta_N}{2\pi}$$

$$= B \sum_{k_1, \ldots, k_N = 0}^{\infty} |a_{k_1, \ldots, k_N}|^2.$$

We turn to Fefferman's idea [**3**]. It suffices to obtain (4.4) for functions f whose coefficients are nonzero for a finite number of indices only. Suppose we know that the above maximal function estimate (4.4) holds when $\mathcal{G}(T, N)$ is replaced by $\mathcal{G}_{\mathbf{x}}(T, N)$, consisting of all vectors $(k_1, \ldots, k_N)$ with nonnegative integer entries subject to

$$T \leq k_1 x_1 + k_2 x_2 + \cdots + k_N x_N,$$

for a dense (in $[0, +\infty[^N)$ collection of vectors $\mathbf{x} = (x_1, \ldots, x_N)$ with positive rational entries. By approximation, then, (4.4) holds, with the original set $\mathcal{G}(T, N)$, for all functions f, the coefficients of which are nonzero for a finite number of indices only.

We consider $\mathbf{x}$ of the form $(q_1/Q, q_2/Q, \ldots, q_N/Q)$, where q_j and Q are positive integers, and the integers q_1, q_2 are relatively prime. We may pick x_1 and x_2 as arbitrarily positive rationals, and let Q be the least possible positive integer such that $q_1 = Qx_1$ and $q_2 = Qx_2$ are integers: Then q_1 and q_2 are automatically relatively prime. The set of pairs (x_1, x_2) for which Q gets as large as we prescribe is dense in $[0, +\infty[^2$, and hence the rational vectors $\mathbf{x}$ of the above form are dense in $[0, +\infty[^N$. The maximal function for slope $\mathbf{x}$ is

$$\mathbf{M}_{N, \mathbf{x}} f(\boldsymbol{\theta}) = \sup_{0 < T < +\infty} \left| \sum_{\mathbf{k} \in \mathcal{G}_{\mathbf{x}}(T, N)} a_{\mathbf{k}} e^{i\langle \mathbf{k}, \boldsymbol{\theta} \rangle} \right|, \qquad \boldsymbol{\theta} \in \mathbb{R}^N,$$

where $\mathbf{k} = (k_1, \ldots, k_N)$, $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_N)$, and $\langle \mathbf{k}, \boldsymbol{\theta} \rangle$ is

$$\langle \mathbf{k}, \boldsymbol{\theta} \rangle = k_1 \theta_1 + \ldots k_N \theta_N.$$

The criterion that $\mathbf{k} \in \mathcal{G}_{\mathbf{x}}(T/Q, N)$ can be written

$$T \leq k_1 q_1 + k_2 q_2 + \cdots + k_N q_N = \langle \mathbf{k}, \mathbf{q} \rangle,$$

where $\mathbf{q} = (q_1, \ldots, q_N)$, which suggests calling this set $\mathcal{G}_{\mathbf{q}}(T, N)$. The maximal function for slope $\mathbf{x}$ can then be written as

$$\mathbf{M}_{N, \mathbf{x}} f(\boldsymbol{\theta}) = \sup_{0 \leq T < +\infty} \left| \sum_{\mathbf{k} \in \mathcal{G}_{\mathbf{q}}(T, N)} a_{\mathbf{k}} e^{i\langle \mathbf{k}, \boldsymbol{\theta} \rangle} \right|, \qquad \boldsymbol{\theta} \in \mathbb{R}^N.$$

Since q_1 and q_2 are assumed relatively prime, we can find integers r_1, r_2 such that $q_1 r_2 - q_2 r_1 = 1$, so that the $N \times N$ matrix $\mathbf{A}$,

$$\mathbf{A} = \begin{pmatrix} q_1 & q_2 & q_3 & q_4 & \cdots & q_n \\ r_1 & r_2 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{pmatrix},$$

has determinant 1. The matrix $\mathbf{A}$ has integer entries, and by Cramer's rule, so does the inverse matrix $\mathbf{A}^{-1}$. Consequently, the map $\mathbf{A} : \mathbb{Z}^n \to \mathbb{Z}^n$ is one to one and onto ($\mathbb{Z}$ is the set of all integers). Putting $\mathbf{j} = \mathbf{A}\mathbf{k}$, where $\mathbf{j}, \mathbf{k}$ are thought of as column vectors, it follows that the maximal function for slope $\mathbf{x}$ can be rewritten as

$$\mathbf{M}_{N,\mathbf{x}} f(\boldsymbol{\theta}) = \sup_{0 \leq T < +\infty} \left| \sum_{\mathbf{j} \in \mathbf{A}\mathcal{G}_{\mathbf{q}}(T,N)} a_{\mathbf{A}^{-1}\mathbf{j}} e^{i\langle \mathbf{j},(\mathbf{A}^{-1})^*\boldsymbol{\theta}\rangle} \right|, \qquad \boldsymbol{\theta} \in \mathbb{R}^N,$$

where the superscript $*$ indicates that the matrix is transposed. In view of the definition of the index set $\mathcal{G}_{\mathbf{q}}(T,N)$, $\mathbf{j} \in \mathbf{A}\mathcal{G}_{\mathbf{q}}(T,N)$ means that $T \leq j_1$, where $\mathbf{j} = (j_1, \ldots, j_N)$. After the change of variables $\boldsymbol{\phi} = (\mathbf{A}^{-1})^*\boldsymbol{\theta}$, we have, because this transformation is volume-preserving,

$$\int_{]-\pi,\pi]^N} \left| \mathbf{M}_{N,\mathbf{x}} f(\boldsymbol{\theta}) \right|^2 \frac{d\boldsymbol{\theta}}{(2\pi)^N}$$

$$= \int_{]-\pi,\pi]^N} \sup_{0 \leq T < +\infty} \left| \sum_{\mathbf{j} \in \mathbf{A}\mathcal{G}_{\mathbf{q}}(T,N)} a_{\mathbf{A}^{-1}\mathbf{j}} e^{i\langle \mathbf{j},\boldsymbol{\phi}\rangle} \right|^2 \frac{d\boldsymbol{\phi}}{(2\pi)^N},$$

where $d\boldsymbol{\theta} = d\theta_1 \ldots d\theta_N$ and $d\boldsymbol{\phi} = d\phi_1 \ldots d\phi_N$. Actually, the image of $]-\pi,\pi]^N$ under the linear transformation induced by $(\mathbf{A}^{-1})^*$ probably is not $]-\pi,\pi]^N$, but in any case an equivalent domain in $\mathbb{R}^N$ modulo $(2\pi\mathbb{Z})^N$. We write $\mathbf{j} = (j_1, \mathbf{j}')$, where $\mathbf{j}' = (j_2, \ldots, j_N) \in \mathbb{Z}^{N-1}$, and do the same for $\boldsymbol{\phi}$. Then

$$\sum_{\mathbf{j} \in \mathbf{A}\mathcal{G}_{\mathbf{q}}(T,N)} a_{\mathbf{A}^{-1}\mathbf{j}} e^{i\langle \mathbf{j},\boldsymbol{\phi}\rangle} = \sum_{j_1 \geq T} \sum_{\mathbf{j}' \in \mathbb{Z}^{N-1}} a_{\mathbf{A}^{-1}\mathbf{j}} e^{i\langle \mathbf{j}',\boldsymbol{\phi}'\rangle} e^{ij_1\phi_1},$$

where we are able to extend the latter summation over the whole lattice $\mathbb{Z}^{N-1}$ by our convention that $a_{\mathbf{j}} = 0$ if any component of $\mathbf{j}$ is negative. Applying the one-dimensional maximal function estimate of Theorem 1.2 (which holds for $L^2(\mathbb{T})$ functions, not just for those in $H^2(\mathbb{D})$) with respect

to the variable θ_1, we find that for fixed $\boldsymbol{\phi}'$,

$$\int_{-\pi}^{\pi} \sup_{0 \leq T < +\infty} \left| \sum_{j_1 \geq T} \sum_{\mathbf{j}' \in \mathbb{Z}^{N-1}} a_{\mathbf{A}^{-1}\mathbf{j}}\, e^{i\langle \mathbf{j}', \boldsymbol{\phi}' \rangle}\, e^{ij_1\phi_1} \right|^2 \frac{d\theta_1}{2\pi}$$

$$\leq A_{\mathrm{CH}} \int_{-\pi}^{\pi} \left| \sum_{j_1 \in \mathbb{Z}} \sum_{\mathbf{j}' \in \mathbb{Z}^{N-1}} a_{\mathbf{A}^{-1}\mathbf{j}}\, e^{i\langle \mathbf{j}', \boldsymbol{\phi}' \rangle}\, e^{ij_1\phi_1} \right|^2 \frac{d\phi_1}{2\pi}$$

$$= A_{\mathrm{CH}} \int_{-\pi}^{\pi} \left| \sum_{\mathbf{j} \in \mathbb{Z}^N} a_{\mathbf{A}^{-1}\mathbf{j}}\, e^{i\langle \mathbf{j}, \boldsymbol{\phi} \rangle} \right|^2 \frac{d\phi_1}{2\pi},$$

where A_{CH} is the absolute constant of the Carleson-Hunt theorem. Integrating with respect to all the other parameters $\theta_2, \ldots, \theta_N$, we obtain

$$\int_{]-\pi,\pi]^N} \sup_{0 \leq T < +\infty} \left| \sum_{j_1 \geq T} \sum_{\mathbf{j}' \in \mathbb{Z}^{N-1}} a_{\mathbf{A}^{-1}\mathbf{j}}\, e^{i\langle \mathbf{j}, \boldsymbol{\phi} \rangle} \right|^2 \frac{d\boldsymbol{\phi}}{(2\pi)^N}$$

$$\leq A_{\mathrm{CH}} \int_{-\pi}^{\pi} \left| \sum_{\mathbf{j} \in \mathbb{Z}^N} a_{\mathbf{A}^{-1}\mathbf{j}}\, e^{i\langle \mathbf{j}, \boldsymbol{\phi} \rangle} \right|^2 \frac{d\boldsymbol{\phi}}{(2\pi)^N} = A_{\mathrm{CH}} \sum_{j \in \mathbb{Z}^N} \left| a_{\mathbf{A}^{-1}\mathbf{j}} \right|^2$$

$$= A_{\mathrm{CH}} \sum_{k \in \mathbb{Z}^N} \left| a_{\mathbf{k}} \right|^2,$$

as desired. Hence (4.4) holds with the constant $B = A_{\mathrm{CH}}$, where A_{CH} is the absolute constant of the Carleson-Hunt theorem for Fourier series (which may be slightly larger than for Taylor series). The proof is complete.

5. Comments on character sums.

We next comment on the relation of Theorem 1.5 to the question mentioned in the introduction — essentially the same as the problem raised by Erdös — concerning the almost sure growth of the character sums

$$S_N(\chi) = \sum_{n=1}^{N} \chi(n).$$

Erdös [**2**, pp. 251-252] studies the sum $S_N(\chi)$ in the analoguous case where for prime indices p, the random variables $\chi(p)$ are independent and take on the values ± 1 with equal probability $\frac{1}{2}$. At other positive integers, the value of the character is then determined by the multiplicative rule $\chi(mn) = \chi(m)\,\chi(n)$. He announced the existence of a positive constant c such that for almost every character χ, one has

$$S_N(\chi) = O\big(\sqrt{N} \log^c N\big), \qquad \text{as} \quad N \to +\infty.$$

A somewhat unfortunate feature of the numbers ± 1 is that the square of each is 1. That means that when n is a square, we know with certainty that $\chi(n) = 1$. This peculiarity will influence any number containing a square as a factor. This is in sharp contrast with the case when for prime indices p the character $\chi(p)$ is uniformly distributed on $\mathbb{T}$, because then $\chi(n)$ keeps the same uniform distribution on $\mathbb{T}$ for all $n = 2, 3, 4, \ldots$. Instead of switching from ± 1 to $\mathbb{T}$, there is another way to alleviate the difficulty. Namely, we decide to sum only over the square-free numbers, defined by the property that they are not divisible by any square (other than 1); these are the numbers in whose factorization each prime occurs at most once. The square-free numbers have density $6/\pi^2$ in the positive integers [**18**, p. 390]. Wintner opts to sum only over square-free numbers [**24**] (his reason is different, though: He studies the reciprocal of the zeta function, where the Dirichlet coefficients are given by the Mœbius function μ, and he wants to know what happens when μ is replaced by a random function), but as his work predates that of Erdös, we cannot expect a better estimate. Indeed, he obtain the upper estimate $O(N^{\frac{1}{2}+\varepsilon})$, but he also shows that $O(N^{\frac{1}{2}-\varepsilon})$ is almost surely false (here, ε stands for an arbitrarily small positive number). Halász [**6**] makes a considerably deeper study of the same problem. He obtains the estimate from above

$$S_N(\chi) = O\big(\sqrt{N}\, e^{c\sqrt{\log\log N \log\log\log N}}\big), \qquad \text{as} \quad N \to +\infty,$$

almost surely in χ, for some positive constant c, and also finds that there is some negative c such that the above estimate fails almost surely. Although quite sharp, these estimates leave some room for improvement. For instance, if we — as does Erdös — compare with the law of the iterated logarithm, where the sharp growth is known to be $O(\sqrt{N \log\log N})$ (see [**20**, p. 397]), we cannot say whether the multiplicativity makes the sum behave better or worse than this. There is some reason to believe that it makes it better (that is, smaller), and indeed, in the appendix of Montgomery's monograph [**17**], problem number 26, due to Halász, asks whether we have $S_N(\chi) = O(\sqrt{N})$ almost surely.

As the peculiar difficulties with squares encountered when summing over all the positive integers vanish when we turn to having $\chi(p)$ uniformly distributed on $\mathbb{T}$, we expect that Halasz' methods will carry over and supply similar estimates in this setting. We do not pursue this idea further here. Instead, we restrict ourselves to recording a nontrivial upper bound of Erdös type as a corollary of Theorem 1.5.

Suppose $\{a_n\}_{n=1}^{\infty}$ is an arbitrary positive, decreasing, square-summable sequence. Applying summation by parts to Theorem 1.5 (in the form known

as Kronecker's lemma, see [**20**, p. 390]), we obtain the estimate

$$S_N(\chi) = O\left(\frac{1}{a_N}\right) \qquad \text{as} \quad N \to +\infty.$$

With a suitable choice of the sequence $\{a_n\}_{n=1}^{\infty}$ we have the following:

Corollary 5.1. *For any $\varepsilon > 0$, we have, for almost every character χ,*

$$\sum_{n=1}^{N} \chi(n) = O\left(\sqrt{N \log N}\,(\log\log N)^{\frac{1}{2}+\varepsilon}\right) \qquad as \quad N \to +\infty.$$

This settles a question in [**7**, p. 22], which asked for the estimate $O(\sqrt{N} \log N)$. It is better than the bounds obtained directly from the classical Menshov-Steinhaus theorem (see [**13**] or [**19**]).

We finally mention a lower bound that can be obtained in a relatively simple manner by observing the growth of the 'random zeta function'

$$\zeta_\chi(s) = \sum_{n=1}^{\infty} \chi(n)\, n^{-s}$$

as $s \to \frac{1}{2}$ along the positive real line. The proof is omitted.

Proposition 5.2. *For almost every character χ, the following estimate fails:*

$$\sum_{n=1}^{N} \chi(n) = O\left(\frac{\sqrt{N}}{\log N}\right) \qquad as \quad N \to +\infty.$$

There is (so far quite weak) reason to believe that only the randomness really contributes to growth. For instance, in the the totally deterministic case of additive characters on the positive integers, the partial sums are $O(1)$ almost everywhere on the unit circle. In contrast, the completely random case of the law of the iterated logarithm gives $O(\sqrt{N \log\log N})$. We are led to ask whether in the multiplicative case we might have almost everywhere

$$\sum_{n=1}^{N} \chi(n) = O\left(\sqrt{\frac{N \log\log N}{\log N}}\right) \qquad \text{as} \quad N \to +\infty,$$

by counting the number of primes below N and taking into account the log log contribution of the law of the iterated logarithm. Of course, if this is true, then the question of Halász mentioned above automatically gets an affirmative answer (in the setting of uniformly distributed random variables on the unit circle, that is).

References

[1] L. Carleson, *On convergence and growth of partial sums of Fourier series*, Acta Math., **116** (1966), 135-157, MR 33 #7774, Zbl 0144.06402.

[2] P. Erdös, *Some unsolved problems*, Magyar Tud. Akad. Mat. Kutató Int. Közl., **6** (1961), 221-254, MR 31 #2106, Zbl 0100.02001.

[3] C. Fefferman, *On the convergence of multiple Fourier series,* Bull. Amer. Math. Soc., **77** (1971), 744-745, MR 55 #8682, Zbl 0234.42008.

[4] ———, *Pointwise convergence of Fourier series*, Ann. of Math., **98** (1973), 551-571, MR 49 #5676, Zbl 0268.42009.

[5] J. Gordon and H. Hedenmalm, *The composition operators on the space of Dirichlet series with square summable coefficients*, Michigan Math. J., **46** (1999), 313-329, MR 2000f:47046, Zbl 0963.47021.

[6] G. Halász, *On random multiplicative functions*, in 'Hubert Delange Colloquium (Orsay, 1982)', Publ. Math. Orsay, **83**(**4**), Univ. Paris XI, Orsay, 1983, 74-96, MR 85i:11081, Zbl 0522.10033.

[7] H. Hedenmalm, P. Lindqvist and K. Seip, *A Hilbert space of Dirichlet series and systems of dilated functions in $L^2(0,1)$,* Duke Math. J., **86** (1997), 1-37, MR 99i:42033, Zbl 0887.46008.

[8] ———, *Addendum to "A Hilbert space of Dirichlet series and systems of dilated functions in $L^2(0,1)$"*, Duke Math. J., **99** (1999), 175-178, MR 2000g:42029, Zbl 0953.46015.

[9] H. Helson, *Foundations of the theory of Dirichlet series,* Acta Math., **118** (1967), 61-77, MR 35 #5944, Zbl 0171.11902.

[10] ———, *Compact groups and Dirichlet series,* Ark. Mat., **8** (1969), 139-143, MR 44 #3075, Zbl 0199.46601.

[11] R.A. Hunt, *On the convergence of Fourier series*, in 'Orthogonal Expansions and their Continuous Analogues (Proc. Conf., Edwardsville, Ill., 1967)', Southern Illinois Univ. Press, Carbondale, Ill., 1968, 235-255, MR 38 #6296, Zbl 0159.35701.

[12] ———, *Comments on Lusin's conjecture and Carleson's proof for L^2 Fourier series*, in 'Linear Operators and Approximation, II (Proc. Conf., Oberwolfach Math. Res. Inst., Oberwolfach, 1974)', Internat. Ser. Numer. Math., **25**, Birkhäuser, Basel, 1974, 235-245, MR 52 #6300, Zbl 0296.43007.

[13] S. Kaczmarz and H. Steinhaus, *Theorie der Orthogonalreihen*, Monografie Matematyczne, Warzawa, 1935.

[14] J.-P. Kahane, *The last problem of Harald Bohr*, J. Austral. Math. Soc. Ser. A, **47**(**1**) (1989), 133-152, MR 90f:30002, Zbl 0734.30003.

[15] J.-P. Kahane and S. Mandelbrojt, *Sur l'équation fonctionnelle de Riemann et la formule sommatoire de Poisson*, Ann. Sci. École Norm. Sup. (3), **75** (1958), 57-80, MR 21 #4942, Zbl 0086.06004.

[16] M. Lacey and C. Thiele, *On Calderón's conjecture*, Ann. of Math., **149** (1999), 475-496, MR 2000d:42003, Zbl 0934.42012.

[17] H.L. Montgomery, *Ten Lectures on the Interface between Analytic Number Theory and Harmonic Analysis,* CBMS Regional Conference Series in Mathematics, 84; published for the Conference Board of the Mathematical Sciences, Washington, DC, by the American Mathematical Society, Providence, RI, 1994, MR 96i:11002, Zbl 0814.11001.

[18] I. Niven, H.S. Zuckerman and H.L. Montgomery, *An Introduction to the Theory of Numbers*, Fifth edition, John Wiley & Sons, Inc., New York, 1991, MR 91i:11001, Zbl 0742.11001.

[19] A.M. Olevskiĭ, *Fourier Series with Respect to General Orthogonal Systems,* Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 86, Springer-Verlag, New York-Heidelberg, 1975, MR 57 #10347, Zbl 0321.42010.

[20] A.N. Shiryaev, *Probability,* Second edition, Graduate Texts in Mathematics, **95**, Springer-Verlag, New York, 1996, MR 97c:60003, Zbl 0835.60002.

[21] A. Torchinsky, *Real-Variable Methods in Harmonic Analysis*, Pure and Applied Mathematics, **123**, Academic Press, Inc., Orlando, Fla., 1986, MR 88e:42001, Zbl 0621.42001.

[22] S.A. Vinogradov, *Convergence almost everywhere of Fourier series of functions in L^2 and the behavior of the coefficients of uniformly convergent Fourier series,* Soviet Math. Dokl., **17** (1976), 1323-1327, MR 54 #10606, Zbl 0361.42002.

[23] H. Weyl, *Singuläre Integralgleichungen mit besonderer Berücksichtigung des Fourierschen Integraltheorems*, Inaugural Dissertation, Göttingen, 1908.

[24] A. Wintner, *Random factorizations and Riemann's hypothesis*, Duke Math. J., **11** (1944), 267-275, MR 5,255c, Zbl 0060.10510.

DEPARTMENT OF MATHEMATICS
LUND UNIVERSITY, BOX 118
S–22100 LUND
SWEDEN
E-mail address: haakan@maths.lth.se

DEPARTMENT OF MATHEMATICS
THE ROYAL INSTITUTE OF TECHNOLOGY
S-10044 STOCKHOLM
SWEDEN

DEPARTMENT OF MATHEMATICS
DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF JYVÄSKYLÄ
P.O. BOX 35 (MAD)
40014 UNIVERSITY OF JYVÄSKYLÄ
FINLAND
E-mail address: saksman@maths.jyu.fi