ARITHMETIC LINEAR TRANSFORMATIONS AND ABSTRACT PRIME NUMBER THEOREMS

S. A. AMITSUR

1. Introduction. Shapiro and Forman have presented in **(4)** an abstract formulation of prime number theorems which includes the various prime number theorems; for primes in arithmetic progressions, for prime ideals in ideal classes etc. The methods of proofs are "elementary" and follow closely Shapiro's proof for the primes in arithmetic progression (for reference see bibliography in **(4)**).

The author has followed in **(1)** some ideas of Yamamoto **(5)** on arithmetic linear transformations to introduce a symbolic calculus in dealing with arithmetic functions. This calculus proved to be very useful in unifying many of the "elementary" proofs in the behaviour of arithmetic functions. In **(6)** Yamamoto has extended his theory to ideals in algebraic number fields, and with this extension the symbolic calculus of **(1)** can be extended to cover the abstract case of prime number theorem in countable free abelian groups as discussed in **(4)**. Furthermore, a more careful study of the behaviour of certain "remainders" yields a more general result in the direction given by Beurling **(3)**.

Shapiro and Forman have considered the following situation. Let G be a free abelian group on a countable number of generators p_i $(i = 1, 2, \ldots,)$. $N{:}G \to \mathfrak{N}$ be a homomorphism of G into the multiplicative group of all integers $\mathfrak{N}$, with the kernel G' such that G/G' is finite. If H is a generic class of G/G', and w is an integral word in G, then the main result of **(4)** is deriving from the condition

$$\sum_{\substack{Nw \leqslant x \\ w \in H}} 1 = c_H x + R_H(x)\,;\, c_H \geqslant 0,\ \sum c_H > 0 \tag{1.1}$$

a "prime number theorem" for the class H:

$$\pi_H(x) = \sum_{\substack{Np \leqslant x \\ p \in H}} 1 = d_H \frac{x}{\log x} + o\left(\frac{x}{\log x}\right). \tag{1.2}$$

A complete analysis of the coefficient of d_H was given in **(4)** for the case $R_H(x) = O(x^\theta)$ with $1 > \theta \geqslant 0$. The methods developed in the present paper will show that the same results are valid even if $R_H(x) = O(x/\log^\gamma x)$ with $\gamma > 2$. A result of a similar nature, though in a completely different situation, was given by Beurling **(3)** for $\gamma > \frac{3}{2}$.

It is quite surprising that for $\gamma > 3$ (and in certain cases for $\gamma > 4$) the methods and the results of **(1)** can be carried over to the abstract case almost

Received April 27, 1959.

without any change whereas $3 > \gamma > 2$ involves many refinements of the methods and of the main "elementary proof" of **(1)**. In fact, some of the equivalent forms of the prime number theorem cannot be proved by our methods for $4 > \gamma > 2$; though others can be shown for these values of γ relatively easily, their classical proofs of implying the prime number theorem breaks down if $\gamma < 3$.

We take this opportunity to present also the symbolic calculus of **(1)** in a more general and in what we hope is a simplified form. An application is also given to show (by elementary methods) that $\zeta(1 + it) \neq 0$ for $t \neq 0$ and that $\sum p^{-1+it}$ converges.

2. The semi-group W and its characters. In the present context we prefer to consider the semi-group W of all integral words in G, and similarly $W' = W \cap G'$. In this way the group K of all characters of G/G' **(4)** is replaced by a finite group of characters of W. To be more precise, we assume the following:

Let W be a free abelian multiplicative semi-group generated by a countable number of generators p_i. Let N be a homomorphism of W into the multiplicative semi-group $\mathfrak{N}$ of all positive integers, that is, Nw is an integer and $N(w_1w_2) = Nw_1 \,.\, Nw_2$.

Let K be a finite group of characters of W. By a character $\chi \in K$, we mean a homomorphism of W into the complex numbers. The unit $\chi_0 \in K$ is defined as $\chi_0(w) = 1$ for all $w \in W$. Multiplication in K is given by:

$$(\chi\eta)(w) = \chi(w)\eta(w). \tag{2.1}$$

Let K be a finite group of order h, then it follows readily by (2.1) that $\chi(w)$ is an hth root of unity. Furthermore, each $w \in W$ determines a character of K by setting $\bar{w}(\chi) = \chi(w)$. Thus the mapping $w \rightarrow \bar{w}$ is a homomorphism of W into the group $\hat{K}$ of all characters of K. Let W' be the kernel of this map, that is,

$$W' = \{w;\, w \in W,\, \chi(w) = 1 \text{ for all } \chi \in K\}.$$

This readily implies that W/W' is a finite group of order $\leqslant$ order of K = order of $\hat{K} = h$. Now the classes H of W/W' are determined by the group of characters K; that is, u, v belong to the same class H if and only if $\chi(u) = \chi(v)$ for all $\chi \in K$, or in other words if and only if $\bar{u} = \bar{v}$. On the other hand, K is readily seen to induce a group of characters on the finite group W/W', and from the definition of the classes of the latter it follows that different characters of K induce different characters of W/W'. Consequently h = order of $K \leqslant$ order of W/W'. Combining this with the previous result, we obtain:

PROPOSITION 1. *W/W' is a finite group of order h, and K can be considered as the group of all characters of W/W'.*

In many cases the converse situation is preferred. Namely, given $W' \subseteq W$ such that W/W' is finite, we define K to be the group of all characters of W/W', and then $\chi(w)$ is defined to be $\chi(H)$ where $w \in H$ the class in W/W'. In any case, we shall always use the notation $\chi(H)$ and $\chi(w)$ for the same character χ.

Now the standard relation between characters yields:

$$\sum_H \chi(H)\eta(H) = \begin{cases} 0 & \text{if} \quad \chi \neq \bar{\eta} \\ h & \text{if} \quad \chi = \eta \end{cases} \tag{2.2}$$

$$\sum_\chi \chi(u)\chi(v) = \begin{cases} 0 & \text{if } u, v \text{ belong to different classes of } W/W' \\ h & \text{if } u, v \text{ belong to the same class.} \end{cases} \tag{2.3}$$

Next we assume that for any class

$$B_H(x) = \sum_{\substack{Nw \leqslant x \\ w \in H}} 1 = C_H x + O(x/\log^\gamma x);\, C_H \geqslant 0,\, \sum C_H > 0. \tag{2.4}$$

We define

$$\psi_H(x) = \sum_{\substack{Np^i \leqslant x \\ p^i \in H}} \log Np;\; \pi_H(x) = \sum_{\substack{Np \leqslant x \\ p \in H}} 1. \tag{2.5}$$

Analogous to the results of Shapiro and Forman **(4)**, we shall show that the character can be distributed into three classes Γ_1, Γ_2, Γ_3. Γ_1 will contain all character for which $A_\chi = \sum_H \chi(H)\, C_H \neq 0$, Γ_2 and Γ_3 will be defined later in §8. Our first result is:

Theorem A. *If* $\gamma > 2$, *then*

$$\sum_{Np^i \leqslant x} \chi(p^i) \log Np = \begin{cases} x + o(x) & \text{if} \quad \chi \in \Gamma_1 \\ o(x) & \text{if} \quad \chi \in \Gamma_2 \\ -x + o(x) & \text{if} \quad \chi \in \Gamma_3 \text{ and } \gamma > 3. \end{cases}$$

Let $U = \{w;\, \chi(w) = 1 \text{ for all } \chi \in \Gamma_1\}$, and $U^* = \{w;\, w \in U,\, \chi(w) = 1$ for all $x \in \Gamma_3\}$. Then $W' \subseteq U^* \subseteq U$ and as in **(4**, Theorem 3.1**)**:

Theorem B. *If* (2.4) *holds then:*

$$\pi_H(x) = d_H \frac{x}{\log x} + o\left(\frac{x}{\log x}\right),$$

if (a) $\Gamma_3 = \emptyset$, $\gamma > 2$, *where:*

$$d_H = \begin{cases} 0 & \text{for } H \notin U/W' \\ h^{-1} \text{ order } \Gamma_1 & \text{for } H \in U/W' \end{cases}$$

or (b) $\Gamma_3 \neq \emptyset$, $\gamma > 3$ *where:*

$$d_H = \begin{cases} 0 \text{ for } H \notin U/W' \text{ or } H \in U^*/W' \\ 2h^{-1} \text{ order } \Gamma_1 \quad \text{for } H \in U/W' \text{ or } H \notin U^*/W'. \end{cases}$$

3. The ring C(W). Let $C(W)$ be the set of all complex valued functions of W. As in **(5**, p. 42**)**, $C(W)$ is a ring with respect to the addition

$$(f + g)\,(w) = f(w) + g(w);\, w \in W, \tag{3.1}$$

https://doi.org/10.4153/CJM-1961-008-1 Published online by Cambridge University Press

and the convolution:

$$(3.2)\qquad (f*g)(w) = \sum_{uv=w} f(u)g(v).$$

We shall also use the ordinary multiplication:

$$(3.3)\qquad (fg)\ (w) = f(w)g(w).$$

The ring $C(W)$ is in fact a commutative ring with the unit ϵ defined.

$$(3.4)\qquad \epsilon(1) = 1,\ \epsilon(w) = 0 \text{ for } w = 1 \text{ (the unit of } W).$$

As in the classical case $W = \mathfrak{N}$ the integers, **(1, 5)** it is easily shown that the invertible function $f \in C(W)$ are those for which $f(1) \neq 0$, and in this case $f^{-1}(w)$ is defined by induction on the length of the words.

$$(3.5)\quad f^{-1}(1) = 1/f(1);\ f^{-1}(w) = -\left[\sum_{u|w} f^{-1}(u)f(wu^{-1})\right]\Big/ f(1),\ u \neq w.$$

Let E be the "one" function defined $E(w) = 1$ for all $w \in W$, then its inverse $E^{-1} = \mu_w = \mu$ is the Mobius function for W:

$$(3.6)\qquad \begin{array}{l}\mu(w) = (-1)^r \text{ if } w \text{ is the product of } r \text{ distinct generators}\\ \mu(p) = 1, \text{ and zero otherwise.}\end{array}$$

A function f is said to be *multiplicative* if:

$$(3.7)\qquad f(uv) = f(u)f(v) \text{ for } (u, v) = 1$$

where $(u, v) = 1$ means that u, v have no common divisor $\neq 1$ in W. If (3.7) holds for all u, v without any restriction, then we say that f is *factorable* or f is a *character*. Another type of functions which we meet are the *additive* functions which satisfy:

$$(3.8)\qquad f(u\,v) = f(u) + f(v).$$

In the general case of arbitrary semi-group W as in the case of the integer **(5)** we have:

PROPOSITION 2. *If f is a character, then the mapping: $g \to gf$ is an isomorphism of $C(W)$ into itself. In particular: $(g*h)f = (gf)*(hf)$.*

If f is an additive function, then the mapping: $g \to gf$ is a derivation of $C(W)$. In particular: $(g*h)f = (gf)*h + g*(hf)$.

Let N be the homomorphism of W into the semi-group of all integers $\mathfrak{N}$. We shall refer to Nw as the *norm* of w. Since N is a homomorphism, N is a character, and consequently the log-function L, defined thus:

$$(3.9)\qquad L(w) = \log Nw$$

is an additive function. Thus, it follows from Proposition 2 that

$$(3.10)\qquad (f*g)\ L = fL*g + f*gL \text{ for all } f, g \in C(W).$$

We shall use the notation L^m to mean $L^m(w) = \log^m(Nw)$. With the aid of $\mu = \mu_w$ we define as in **(6**, p. 44**)** the Mangolt-function $\Lambda = \Lambda_w = \mu * L$ the Selberg-function $\Lambda_2 = \mu * L^2$ and higher types $\Lambda_m = \mu * L^m$. We recall that

$$\Lambda(p^e) = \log Np \text{ and } \Lambda(w) = 0 \text{ if } w \neq p^e \text{ for a generator } p \in W. \tag{3.11}$$

$$\Lambda_2(p^e) = (2e - 1)\log Np;\ \Lambda_2(p^e q^f) = 2\log Np \log Nq; \quad \Lambda_2(w) = 0 \text{ for } w \neq p^e q^f. \tag{3.12}$$

To every $f \in C(W)$ we define an arithmetic function $Nf \in C(\mathfrak{N})$ by setting

$$(Nf)(n) = \sum_{Nw=n} f(w) \qquad \text{for every integer } n > 0, \tag{3.13}$$

and if there are no $w \in W$ satisfying, $Nw = n$ then we set $(Nf)(n) = 0$.

Thus $(NE)(n)$ is the number of elements of W whose norm is n. It is not difficult to show

THEOREM 1. *The mapping $f \to Nf$ is a homomorphism of $C(W)$ into the ring of all arithmetic function $C(\mathfrak{N})$.*

4. The ring of arithmetic linear transformations.

Let F be the linear space of all complex valued functions $\Phi(x)$ defined for all real $x \geqslant 1$. To each $f \in C(W)$, we make correspond (as in **(1, 5)**) a linear transformation S_f of F, defined by

$$(S_f\Phi)(x) = \sum f(w)\Phi(x/Nw);\ \Phi \in F \text{ and all } x \geqslant 1. \tag{4.1}$$

The following is then easily verified.

PROPOSITION 3.

$$S_{f+g} = S_f + S_g;\ cS_f = S_{cf};\ S_{f*g} = S_f S_g.$$

That means that the correspondence: $f \to S_f$ is a homomorphism of $C(W)$ into the ring of all linear transformations of F.

Definition (4.1) is valid for all semi-groups, in particular for $W = \mathfrak{N}$ (the integers) where in the semi-group of integer the norm is to be the identity map. Then clearly we have, by (3.13),

PROPOSITION 4.

$$S_f \Phi = S_{Nf}\Phi.$$

For practical purposes we prefer to substitute for S_f a different operator I_f defined by

$$(I_f\Phi)(x) = \sum_{Nw \leqslant x} \frac{f(w)}{Nw}\Phi\left(\frac{x}{Nw}\right) = (S_{fN^{-1}}\Phi)(x) \tag{4.2}$$

where

$$(fN^{-1})(w) = f(w)/Nw.$$

https://doi.org/10.4153/CJM-1961-008-1 Published online by Cambridge University Press

As we remarked above, N and therefore also N^{-1} are characters of W, hence it follows readily by Proposition 2 that Propositions 3 and 4 will hold also for I_f. For further references we formulate this result in the following proposition which includes also an additional simple fact.

PROPOSITION 5.

$$I_f + I_g = I_{f+g}; \quad cI_f = I_{cf}; \quad I_f I_g = I_{f*g}; \quad I_f \Phi = I_{Nf}\Phi,$$

and

$$I_f(\Phi \log x) = \log x \,.\, I_f\Phi - I_{fL}\Phi.$$

5. The space $\mathfrak{L}$. In the present section we extend the formalism introduced in **(1)** to cover the general case dealt with in the present paper.

Let $\mathfrak{L}$ be the space of all polynomials $\phi(\log x) = \sum a_\nu \log^\nu x$ in the function $\log x$. We introduce the formal derivation $D = d/d \log x$ with all its positive and negative powers by writing

$$\begin{aligned} D^m \log^n x &= (n)_m \log^{n-m} x \text{ for all } n \geqslant m,\ n \geqslant 0, \\ &= 0 \qquad\qquad\qquad \text{if } m > n, \end{aligned} \tag{5.1}$$

where $(n)_m = n!/(n-m)!$ if $n \geqslant m$ and $n \geqslant 0$; m can be positive or negative. Thus, D^0 is the identity. For completeness we set $(n)_m = 0$ if $m > n$. Now D^m acts on $\Phi(\log x)$ by setting: $D^m(\sum a_\nu \log^\nu x) = \sum (\nu)_m a_\nu \log^{\nu-m} x$.

Let $\alpha_{-p}, \alpha_{-p+1}, \ldots, \alpha_0, \ldots,$ be a sequence of complex numbers, then the symbol

$$F(D) = \sum_{\nu=-p}^{\infty} \alpha_\nu D^\nu$$

will be considered as a linear operator on $\mathfrak{L}$, by putting

$$F(D) \log^n x = \sum_{\nu=-p}^{\infty} \alpha_\nu D^\nu \log^n x = \sum_{\nu=-p}^{n} (n)_\nu \alpha_\nu \log^{n-\nu} x. \tag{5.2}$$

Let $f \in C(W)$, $F(D)$ be as above. Then we denote by $R_n(x; f, F)$ the remainder element defined by the relation

$$I_f \log^n x = F(D) \log^n x + R_n(x; f, F). \tag{5.3}$$

That is, in view of (5.2)

$$R_n(x; f, F) = \sum_{Nw \leqslant x} \frac{f(w)}{Nw} \log^n \frac{x}{Nw} - \sum_{\nu=-p}^{n} (n)_\nu \alpha_\nu \log^{n-\nu} x. \tag{5.4}$$

As in **(1)** we shall write

$$I_f = F(D) + O(\varphi_n) \tag{5.5}$$

to mean

$$R_n(x; f, F) = O(\varphi_n(x)), \text{ for all } n \geqslant 0.$$

The notations $R_n(x)$, $R_n(x; f)$ and $R_n(f)$ will replace $R_n(x; f, F)$ when no confusion will be involved.

For further references we fix

$$F(D) = \sum_{\nu=-p}^{\infty} \alpha_\nu D^\nu, \; G(D) = \sum_{\mu=-q}^{\infty} \beta_\mu D^\mu.$$

The following is easily verified.

THEOREM 2. (1) $\alpha R_n(x;f, F) + \beta R_n(x;g,G) = R_n(x;\alpha f + \beta g, \alpha F + \beta G)$
(2) $R_n(x;fL, -F') = \log x \cdot R_n(x;f, F) - R_{n+1}(x;f, F)$

where $F' = \sum \nu\alpha_\nu D^{\nu-1}$ *is the formal derivative with respect to* D.

The proof of (2) follows as in the proof of (4) of (**1**, Theorem 4.1).

Another simple result which is of great importance in the present paper is

THEOREM 3.

$$R_n(x;g{*}f, GF) = I_g R_n(x;f, F) + \sum_{j=0}^{n+p} (n!/j!)\alpha_{n-j} R_j(x;g, G)$$
$$- \sum_{t=0}^{q-1} \sum_{s=0}^{t+1} \alpha_{n+t+s}\beta_{-s}(n!/t!) \log^t x.$$

This will be used mainly in the following form, (noting that $\alpha_{-p} \neq 0$)

$$R_{n+p}(x;g, G) = cI_g R_n(x;f, F) + \sum_{j=0}^{n+p-1} c_j R_j(x;g, G) + \sum_{t=0}^{q-1} c_{n,t} \log^t x + dR_n(x;g{*}f, GF). \tag{5.6}$$

for some constants c, c_j, $c_{n,t}$, d. In both formulas if $n + p < 0$, the term containing $R_j(x_i\, g, G)$ does not appear, and if $q - 1 < 0$ the last term is not to be considered.

We note also that $G(D)F(D)$ is the formal product of the two power series in D and not the product of the operator G and F; the two products are not always equal as can be seen by: $1 = (D^{-1}D)1 \neq D^{-1}(D1) = 0$.

Proof.

$$I_{g*f} \log^n x = I_g(I_f \log^n x) = I_g\left[\sum_{\nu=-p}^{n} (n)_\nu \alpha_\nu \log^{n-\nu} x + R_n(x;f, F)\right]$$
$$= I_g R_n(x;f, F) + \sum_{\nu=-p}^{n} (n)_\nu \alpha_\nu I_g \log^{n-\nu} x$$
$$= I_g R_n(x;f, F) + \sum_{\nu=-p}^{n} (n)_\nu \alpha_\nu R_{n-\nu}(x;g, G)$$
$$+ \sum_{\nu=-p}^{n} \sum_{\mu=-q}^{n-\nu} (n)_\nu \alpha_\nu (n-\nu)_\mu \beta_\mu \log^{n-\nu-\mu} x$$
$$= A + B + C = (GF) \log^n x + R_n(x;g{*}f, GF).$$

The terms A, B appear in the statement of Theorem 3 (by setting $j = n - \nu$)

and if $n + p < 0$, we do not get B. To complete the proof of Theorem 3 we have to compare

$$[G(D)F(D)] \log^n x = \sum_{k=-(p+q)}^{n} \Big(\sum_{\nu+\mu=k} \alpha_\nu \beta_\mu \Big) (n)_k \log^{n-k} x$$

with

$$C = \sum_{\nu=-p}^{n} \sum_{\mu=-q}^{n-\nu} (n)_\nu (n-\nu)_\mu \alpha_\nu \beta_\mu \log^{n-\nu-\mu} x = \sum_{k=-(p+q)}^{n} \Big(\sum_{\nu+\mu=k}{}' \alpha_\nu \beta_\mu \Big) (n)_k \log^{n-k} x,$$

which is obtained by setting $\nu + \mu = k$. The difference between the two is that in $(GF) \log^n x$, the sum ranges over all $\nu \geqslant -p$, $\mu \geqslant -q$, whereas in the second sum it ranges only over: $n \geqslant \nu \geqslant -p$, $n - \nu \geqslant \mu \geqslant -q$. Comparing the two we observe that they have common range as long as $\min(n, k+q) \geqslant \nu \geqslant -p$ with $k = \nu + \mu \leqslant n$. Thus the terms for which $k + q \geqslant \nu > n$ show that:

$$C - (GF)\log^n x = - \sum_{k=n-q+1}^{n} \Big(\sum_{\nu+\mu=k}{}'' \alpha_\nu \beta_\mu \Big) (n)_k \log^{n-k} x$$
$$= \sum_{t=0}^{q-1} \sum_{s=1}^{t+1} (n!/t!) \alpha_{n+t+s} \beta_{-s} \log^t x,$$

since in $\sum''$, $\nu > n$; hence, the last form is obtained by setting $t = n - k$ and $s = -\mu$, as then $\nu = k - \mu = n + t + s$. (If $q = 0$, this term does not appear, since $k + q \leqslant n$.)

The relation (5.6) is very useful in computing $R_n(x; f^{-1}, F^{-1})$ by induction, since it provides us with a recursive formula for $R_n(x; f^{-1}, F^{-1})$ as will be used later.

Another formula for $R_n(g*f)$ has been obtained in **(2)** following the Dirichlet hyperbola method for summation. This result has been proved only for the integers (Theorem 1 of **(2)**), and we formulate it here for the semi-group W, but the two results are equivalent as is readily seen by the equality $I_f = I_{Nf}$, which leads to the relation: $R_n(x; f, F) = R_n(x; Nf, F)$ for all $f \in C(W)$. We quote that result in the following theorem.

THEOREM 4. Let $yz = x$; $1 \leqslant y \leqslant x$ *then*

$$R_n(x; g*f, GF) = \sum_{Nw \leqslant y} \frac{g(w)}{Nw} R_n\Big(\frac{x}{Nw}; f, F\Big) + \sum_{Nw \leqslant z} \frac{f(w)}{Nw} R_n\Big(\frac{x}{Nw}; g, G\Big)$$
$$- \sum_{j=0}^{n} \binom{n}{j} R_{n-j}(y; g, G) R_j(z; f, F)$$
$$+ \sum_{i=1}^{p} \sum_{j=1}^{i} \frac{n!}{(i-j)!(n+j)!} \alpha_{-i} R_{n+j}(y; gG) \log^{i-j} z +$$
$$\sum_{i=1}^{q} \sum_{j=1}^{i} \frac{n!}{(i-j)!(n+j)!} \beta_{-i} R_{n+j}(z; f, F) \log^{i-j} y,$$

and the respective terms do not appear if the power series of $F(D)$, $G(D)$ *do not have negative powers of* D (that is, $p \leqslant 0$ or $q \leqslant 0$).

The following lemma will be used extensively in §6.

LEMMA 1. *Let*

$$\Phi(x) = \sum_{Nw \leqslant x} f(w),$$

let $\varphi(x)$ *be a differentiable function. Then*

$$\sum_{Nw \leqslant x} f(w)\varphi(Nw) = \Phi(x)\varphi(x) - \int_1^x \Phi(t)d\varphi(t),$$

or more generally

$$\sum_{y < Nw \leqslant x} f(w)\varphi(Nw) = \Phi(x)\varphi(x) - \Phi(y)\varphi(y) - \int_y^x \Phi(t)d\varphi(t).$$

This lemma follows immediately from (7, Theorem 421, p. 346) noting that

$$\sum_{Nw \leqslant x} f(w) = \sum_{n \leqslant x} (Nf)(n).$$

6. Approximating I_f. In the following two sections we consider functions $f \in C(W)$ with properties

(6.1) $$S_f 1 = \sum_{Nw \leqslant x} f(w) = \alpha x + O(x/\log^\gamma x), \qquad \gamma = 1 + \delta > 0,$$

(6.2) $$S_{|f|} 1 = \sum_{Nw \leqslant x} |f(w)| = Ax + O(x/\log^\gamma x),$$

or the weaker condition:

(6.2*) $$\sum_{Nw \leqslant x} |f(w)| = O(x).$$

For later applications we shall introduce the assumption

(6.3) f^{-1} exists in $C(W)$ and $|f^{-1}(w)| \leqslant K|f(w)|$ for some $K > 0$, and all $w \in W$

These function will satisfy

PROPOSITION 6.

(6.4) $$I_f 1 = \sum_{Nw \leqslant x} (Nw)^{-1} f(w) = \alpha \log x + \alpha_0 + \rho(x), \qquad \rho(x) = O(\log^{-\delta} x)$$

(6.5a) $$S_{fL} 1 = \sum_{Nw \leqslant x} f(w) \log Nw = \alpha x \log x - \alpha x + \alpha + O(x \log^{-\delta} x)$$

(6.5b) $$S_{fL^2} 1 = \sum_{Nw \leqslant x} f(w) \log^2 Nw = \alpha x \log^2 x - 2\alpha x \log x + 2\alpha x - 2\alpha + O(x \log^{1-\delta} x).$$

The proof follows immediately from (6.1) by applying Lemma 1. We observe also that, since $\delta = 1 - \gamma > 0$.

$$\alpha_0 = \alpha + \int_1^\infty O(t^{-1}\log^{-\gamma} t)dt; \ \rho(t) = O(\log^{-\gamma} x) + \int_x^\infty O(t^{-1}\log^{-\gamma} t)dt$$
$$= O(\log^{-\delta} x).$$

In what follows we determine an approximation of I_f assuming only the validity of (6.4), and to simplify results, we assume henceforth that $\delta = \gamma - 1$ is not an integer.

From Lemma 1 we obtain, for $n > 0$,

$$I_f \log^n x = \sum_{Nw \leqslant x} (Nw)^{-1} f(w) \log^n (x/Nw) = -\int_1^x (\alpha \log t + \alpha_0 + \rho(t)) \, d\log^n (x/t)$$
$$= (n+1)^{-1}\alpha \log^{n+1} x + \alpha_0 \log^n x - \sum_{\nu=1}^n \binom{n}{\nu}(-1)^\nu \log^{n-\nu} x \int_1^x \rho(t) d\log^\nu t$$
$$= (n+1)^{-1}\alpha \log^{n+1} x + \alpha_0 \log^n x + \sum_{1\leqslant \nu < \delta} \frac{(-1)^{\nu-1}}{(\nu-1)!}\int_1^\infty \rho(t)t^{-1}\log^{\nu-1} t\, dt \cdot \frac{n!}{(n-\nu)!}\log^{n-\nu} x$$
$$+ \sum_{\nu<\delta} (-1)^\nu \binom{n}{\nu} \log^{n-\nu} x \cdot \int_x^\infty \rho(t) d\log^\mu t - \sum_{\nu>\delta} (-1)^\nu \binom{n}{\nu}\log^{n-\nu} x \int_1^x \rho(t) d\log^\nu t.$$

This is true since, for $\nu < \delta$,

$$\int_1^\infty \rho(t) d\log^\nu t = \nu \int_1^\infty t^{-1}\rho(t)\log^{\nu-1} t\, dt < \infty$$

as $\rho(t) = O(\log^{-\delta} x)$. If $n < \delta$, we disregard the last term. Put

$$(6.6) \qquad \begin{cases} F(D) = \sum_{\nu=-1}^\infty \alpha_\nu D^\nu \text{ with } \alpha_{-1} = \alpha, \alpha_0 \text{ as given in (6.4)} \\ \alpha_\nu = O \text{ for } \nu > \delta, \\ \alpha_\nu = \dfrac{(-1)^{\nu-1}}{(\nu-1)!}\displaystyle\int_1^\infty \frac{\rho(t)\log^{\nu-1} t}{t} dt \text{ for } 1 \leqslant \nu < \delta. \end{cases}$$

Thus we have obtained that $I_f \log^n x = F(D)\log^n x + R_n(x; f, F)$ where

$$(6.7) \quad R_n(x;f,F) = \sum_{\nu<\delta} (-1)^\nu \binom{n}{\nu} \log^{n-\nu} x \int_x^\infty \rho(t) d\log^\nu t - \sum_{\nu>\delta} (-1)^\nu \binom{n}{\nu} \log^{n-\nu} x \int_1^x \rho(t) d\log^\nu t$$

and for the case $n = 0$, we have clearly by (6.4)

$$(6.7^*) \qquad R_0(x; f, F) = \rho(x).$$

If $n < \delta$ we can obtain a better form for R_n, namely

$$R_n(x; f, F) = \sum_{\nu=1}^{n} (-1)^{\nu} \binom{n}{\nu} \log^{n-\nu} x \cdot \int_x^{\infty} \rho(t) d \log^{\nu} t \tag{6.8}$$
$$= \int_x^{\infty} \rho(t) d \log^n (xt^{-1}) = (-1)^n \int_0^{\infty} \rho(xe^u)\, du^n,$$

where the latter is obtained by setting $u = -\log (xt^{-1})$.

If (6.3) holds for all $\delta > 0$ (that is, (6.1) is valid for all $\gamma > 0$) then we define α_ν by the integral of (6.6) for all $\nu \geqslant 1$. Furthermore, it follows readily from (6.8) that, for all $n < \delta$,

$$R_n(x; f, F) = \pm \int_0^{\infty} \rho(xe^u) du^n = O(\log^{n-\delta} x). \tag{6.9}$$

Thus we have

THEOREM 5. *If*

$$\sum_{Nw \leqslant x} (Nw)^{-1} f(w) = \alpha_{-1} \log x + \alpha_0 + O(\log^{-\delta} x)$$

for all $\delta > 0$ *then* $I_f = F(D) + O(\log^{-\delta} x)$ *for all* $\delta > 0$, *and* $F(D)$ *is as given in* (6.6).

In many cases we can obtain a better bound for $R_n(x; f, F)$.

If $\rho(x) = O(x^{-\vartheta})$, $\vartheta > 0$, then one readily obtains from (6.8) that

$$R_n(x; f, F) = O\left(\int_0^{\infty} x^{-\vartheta} e^{-\vartheta u} du^n\right) = O(x^{-\vartheta}).$$

COROLLARY. *If*

$$\sum_{Nw \leqslant x} f(w) = \alpha x + O(x^{1-\vartheta})$$

then $I_f = F(D) + O(x^{-\vartheta})$.

Now, if (6.9) is valid only for a bounded δ, then we can only show

THEOREM 6. $R_n(x; f, F) = O(\log^{n-\delta} x)$ *with* $F(D)$ *as given in* (6.6).

Indeed, for $\nu < \delta$,

$$\int_x^{\infty} \rho(t) d \log^{\nu} t = O(\log^{\delta-1} x)$$

and for $\nu > \delta$

$$\int_1^{x} \rho(t) d \log^{\nu} t = O(\log^{\delta-1} x).$$

Thus our theorem follows immediately from (6.7).

Applying Theorem 2 to this approximation of I_f yields

THEOREM 7.

$$I_{fL} = - \sum_{\nu=-1}^{\infty} \nu \alpha_\nu D^{\nu-1} + O(\log^{n+1-\delta} x),$$

and generally

$$I_{fLm} = (-1)^m F^{(m)}(D) + O(\log^{n+m-\delta} x),$$

where

$$F(D) = \sum_{\nu=-1}^{\infty} \alpha_\nu D^\nu$$

is given in (6.6) *and* $F^{(m)}(D)$ *denotes the* mth *formal derivative of* $F(D)$ *with respect to* D.

The coefficients $\alpha_{-1} = \alpha$, α_0, α_1, A are defined in (6.2) and (6.6). We have to deal separately with the following cases.

Case I $\alpha_{-1} \neq 0$ (this implies that $A \neq 0$).

Case II $\alpha_{-1} = 0$, $\alpha_0 \neq 0$, and $A \neq 0$.

Case III $\alpha_{-1} = \alpha_0 = 0$ (this will imply that $A \neq 0$).

Case IV $\alpha_{-1} = 0$, $\alpha_0 \neq 0$ and $A = 0$

Our first purpose is to show that

THEOREM 8.

$$I_{f^{-1}} = F^{-1}(D) + O(1) + \begin{cases} O(\log^{n-\delta} x) \text{ in Cases I, IV} \\ O(\log^{n+1-\delta} x) \text{ in Case II} \\ O(\log^{n+2-\delta} x) \text{ in Case III.} \end{cases}$$

and the four cases contain all possible conditions on the coefficients.

We shall need the following lemma.

LEMMA 2. *Let* $h(w) \geqslant 0$ *satisfying*

$$S_h 1 = \sum_{Nw \leqslant x} h(w) = O(x).$$

Let $g(x) = O(\log^r x)$ *be a non-negative bounded function in finite intervals, then* $I_h g = O(1) + O(\log^{r+1} x)$. *If* $S_h 1 = O(x/\log^r x), r > 1$ *then*: $I_h g = O(1) + O(\log^r x)$.

Indeed, let $|g(x)| < K \log^r x$ for $x > a$, then

$$|I_h g| \leqslant \sum_{xa^{-1} < Nw \leqslant x} (Nw)^{-1} h(x) |g(x/Nw)| + K \sum_{Nw \leqslant xa^{-1}} (Nw)^{-1} h(w) \log^r (x/Nw)$$

$$\leqslant \sup_{1 \leqslant t \leqslant a} |g(t)| (xa^{-1})^{-1} \cdot O(x) + K(xa^{-1})^{-1} \log^r a \cdot O(x)$$

$$- K \int_1^{xa^{-1}} O(t) d[t^{-1} \log^r xt^{-1}] = O(1) + O(\log^{r+1} x)$$

as can readily be obtained by substituting $u = x/t$ in the integral.

The second result follows similarly if we observe that

$$\int_1^{xa^{-1}} O(t \log^{-r} t) d[t^{-1} \log^r xt^{-1}] \leqslant K \int_1^s d[t^{-1} \log^r xt^{-1}] +$$
$$+ L \int_s^{xa^{-1}} t \log^{-r} t d[t^{-1} \log^r xt^{-1}]$$

for some $1 \leqslant s \leqslant a^{-1}$, and some constants $K, L > 0$. Clearly, the first integral is $O(\log^r x)$ and the second is:

$$= O\left(\int_s^{xa^{-1}} t^{-1} \log^{-r} t \log^r (xt^{-1}) dt\right) = O(\log^r x).$$

As a special case, if $f(w)$ satisfies (6.2) and (6.3) and $g(x)$ is as above, then we have

COROLLARY 2.

$$I_{f^{-1}}g = O(I_{|f|}|g|) = O(1) + O(\log^{r+1} x) \qquad \text{if } A \neq 0,$$
$$I_{f^{-1}}g = O(I_{|f|}|g|) = O(1) + O(\log^r x) \qquad \text{if } A = 0.$$

Indeed, by (6.3) it follows that $|f^{-1}(w)| \leqslant k|f(w)|$ for some $K > 0$. Thus, $I_{f^{-1}}g = O(I_{|f|}|g|)$ and the rest follows by the preceding lemma.

Next we prove

PROPOSITION 7. *If f satisfies* (6.1)–(6.3) *and $\delta > 1$ (that is, $\gamma > 2$), then one of the coefficients $\alpha_{-1}, \alpha_0, \alpha_1$ of* (6.6) *is not zero. Furthermore, if $\alpha_{-1} = \alpha_0 = 0$ then $A \neq 0$.*

For, let $\alpha_{-1} = \alpha_0 = 0$, then from Theorem 6 we deduce that $I_f \log x = \alpha_1 + O(\log^{1-\delta} x)$. Applying $I_{f^{-1}}$ on both sides and using the preceding corollary, we obtain

$$\log x = I_{f^{-1}} I_f \log x = \alpha, I_{f^{-1}} 1 + O(\log^{2-\delta} x) + O(1) = \alpha_1 I_{f^{-1}} 1 + o(\log x)$$

since $2 - \delta < 1$, and this can only be true if $\alpha_1 \neq 0$. Moreover, in this case, it follows in view of (6.3) that:

$$|\alpha_1^{-1} \log x + o(\log x)| \leqslant |I_{f^{-1}} 1| \leqslant K I_{|f|} 1 = AK \log x + O(1)$$

hence $A \neq 0$. This proves also that the four cases described in Theorem 8 cover all possible cases. (Note, that at this point only in Case IV we assumed $\gamma > 2$.)

Remark. If in (6.2), we assume that $A = 0$, then clearly $\alpha_{-1} = 0$, since $|I_f 1| \leqslant I_{|f|} 1$, and in this case it follows that $\alpha_0 \neq 0$. The latter is then true even for $\delta > 0$, since we can use the better bound given in Corollary 2 for

$$I_{f^{-1}} O(\log^{1-\delta} x).$$

So that if $\alpha_0 = 0$ we would have

$$\log x = I_{f^{-1}} I_f \log x = \alpha_1 I_{f^{-1}} 1 + I_{f^{-1}} O(\log^{1-\delta} x) = \alpha_1 I_{f^{-1}} 1 + O(\log^{1-\delta}),$$

https://doi.org/10.4153/CJM-1961-008-1 Published online by Cambridge University Press

which implies that $\alpha_1 \neq 0$ even for $\delta > 0$. But then

$$|\alpha_1^{-1} \log x + o(\log x)| = |I_{f^{-1}}1| \leqslant I_{|f|}1 = O|\log^{-\delta} x)$$

which is a contradiction. Hence, $\alpha_0 \neq 0$.

We are now in position to prove Theorem 8.

Case I. Since $\alpha = \alpha_{-1} \neq 0$, $F^{-1}(D) = \alpha^{-1}D + \ldots$ and therefore

$$I_{f^{-1}}1 = F^{-1}(D)1 + R_0(x; f^{-1}, F^{-1}).$$

That is,

$$R_0(x; f^{-1}, F^{-1}) = I_{f^{-1}}1.$$

To evaluate this element, consider the following.

$$\begin{aligned} 1 = S_{f^{-1}}S_f 1 &= S_{f^{-1}}[\alpha x + O(x \log^{-\gamma} x)] = \alpha x I_{f^{-1}}1 + x I_{f^{-1}} O(\log^{-\gamma} x) \\ &= \alpha x R_0(x; f^{-1}, F^{-1}) + x O(\log^{-\gamma+1} x) + x O(1). \end{aligned}$$

Since $S_f x = x I_f 1$, now since $-\delta = 1 - \gamma$, we have shown that $R_0(x; f^{-1}, F^{-1}) = O(1) + O(\log^{-\delta} x)$. We complete the proof of this case by induction on n. Observing that:

$$O \equiv R_n(x; \epsilon, 1) = R_n(x; f^{-1} * f, F^{-1}F)$$

we obtain by (5.6), (where $p = 1$) in view of Theorem 6 and Corollary 2,

$$\begin{aligned} R_{n+1}(x; f^{-1}, F^{-1}) &= O[I_{f^{-1}} R_n(x; f, F)] + O\left(\sum_{j=0}^{n} |R_j\,(x; f^{-1}, F^{-1})|\right) \\ &= \theta[I_{|f|} O(\log^{n-\delta} x)] + \sum_{j=0}^{n} O(\log^{j-\delta} x) + O(1) = O(\log^{n+1-\delta} x) + O(1), \end{aligned}$$

which completes the proof of this case.

Case II. The proof follows by a similar application of (5.6). In this case, $p = 0$ and we need no special method for computing. As we have by (5.6)

$$R_n(x; f^{-1}, F^{-1}) = O[I_{f^{-1}} R_n(x; f, F)] + O\left(\sum_{j=0}^{n-1} |R_j\,(x; f^{-1}, F^{-1})|\right)$$

where for $n = 0$, the sum does not appear. Thus using again an induction, together with Corollary 2 and Theorem 6, we obtain that

$$R_n(x; f^{-1}, F^{-1}) = O(\log^{n+1-\delta} x) + O(1).$$

Incidentally, this provides the proof for Case IV also, since there $A = 0$ and we can use the better approximation

$$I_{f^{-1}} R_n(x; f, F) = O(\log^{n-\delta} x) + O(1)$$

which will yield in Case IV

$$R_n(x; f^{-1}, F^{-1}) = O(\log^{n-\delta} x).$$

https://doi.org/10.4153/CJM-1961-008-1 Published online by Cambridge University Press

Case III. Again we use the same procedure, but here $p = -1$. So that

$$R_{n-1}(x;f^{-1}, F^{-1}) = O[I_{f^{-1}}R_n(x;f, F)] + O\left(\sum_{j=0}^{n-2} |R_j\,(x;f^{-1}, F^{-1})|\right)$$

and we thus obtain $R_{n-1}(x;f^{-1}, F^{-1}) = O(\log^{n+1-\delta}x)$, which completes the proof of Theorem 8.

It follows now readily from Theorem 3 that

THEOREM 9. For $m \geqslant 1$:

$$I_{f^{-1}*fL^m} = (-1)^m F^{-1}(D)F^{(m)}(D) + O(1) + \begin{cases} O\log^{n+m+1-\delta}x) \text{ in Cases I–III} \\ O(\log^{n+m-\delta}x) \text{ in Case IV.}\end{cases}$$

Indeed, Theorem 3 implies

$$R_n(x;f^{-1}*fL^m) = I_f R_n(fL^m) + O\left(\sum_{j=0}^{n+p} |R_;\,(f^{-1})|\right) + O(1)$$

where D^{-p} is the first power of D appearing in $F^{(m)}(D)$, and $O(1)$ has to be added only if $\alpha_{-1} = \alpha_0 = 0$, since then $F^{-1}(D) = \alpha_{-1}{}^{-1}D^{-1} + \ldots$ (that is, $q = 1$).

Since $F(D)$ has at most one negative power of D, that is, D^{-1}, the mth derivative may have the lowest power $D^{-(m+1)}$, thus $p \leqslant m + 1$. Furthermore, in view of Corollary 2 and Theorem 7,

$$I_{f^{-1}}R_n(fL^m) = O(\log^{n+m+1-\delta}x) + O(1).$$

The other terms can get at most to this power, by Theorem 8, which proves Cases I–III. In Case IV $p = 0$ and we can apply the better bound of Corollary 2 to yield the required result.

In particular this leads to

COROLLARY 3. If

$$\sum_{Nw\leqslant x} f(w) = \alpha x + O(x \log^{-\delta}x)$$

for all $\delta > 0$ *and then*

$$I_{f^{-1}*fL^m} = (-1)^m F^{-1}(D)F^{(m)}(D) + O(1).$$

This includes the known results **(1)** about I_μ, I_Λ, I_{Λ_2} where μ, Λ, Λ_2 are the Mobius', Mangoll's, and Selberg's function for the integers, respectively. More applications will be given later.

7. Approximating characters. Let f be a character on W, then the preceding results can be further refined in the direction of the "elementary proofs" developed in **(1)**. This can be achieved relatively easily following the proofs of Theorem 9.1 and 9.2 of **(1)**—only if we assume that $\gamma > 3$ where γ is given in (6.1). We shall outline the proofs of this fact later.

In the present section we want to obtain results which will give us the proof of Theorem A and B even for $\gamma > 2$. We will be able to obtain the result that $\sum f(w)\Lambda(w) = x + o(x)$ if $\gamma > 2$ in most cases, whereas the relation $\sum f(w)Nw^{-1}\Lambda(w) = \log x + c + o(1)$ will be obtained only for $\gamma > 3$.

In the rest of this section we assume that

(A) f is a character satisfying (6.1), (6.2), and $A \neq 0$, $\gamma = 1 + \delta > 2$. Since f is a character, it follows by Proposition 2 that $f^{-1}(w) = f(w)\mu_W(w)$ which shows that f satisfies also (6.3). For these characters we show

PROPOSITION 8.

$$(7.1) \qquad S_{f\Lambda}1 = \sum_{Nw\leqslant x} f(w)\,\Lambda(w) = O(x)$$

$$(7.2) \qquad S_{|f|\Lambda_2}1 = \sum_{Nw\leqslant x} |f(w)|\,\Lambda_2(w) = 2x\log x + O(x) + O(x\log^{2-\delta}x),$$

(Selberg's formula)

$$(7.3) \qquad \Big|\sum_{x<Nw\leqslant tx} f(w)\,\Lambda(w)\Big| \leqslant 2(t-1)x + o(x) \text{ as } (t,x)\to(1,\infty).$$

Proof. It follows from Proposition 2, that since f is a character,

$$(7.4) \qquad f^{-1} = f\mu;\, f\Lambda = f(\mu*L) = f^{-1}*fL \text{ and } f\Lambda_2 = f(\mu*L^2) = f^{-1}*fL^2.$$

As the mapping $g \to gL$ is a derivation in $C(W)$, we have: $(\mu*L)L = \mu L*L + \mu*L^2 = -(\mu*\mu*L)*L + \mu*L^2 = -(\mu*L)^2 + (\mu*L^2)$. Hence:

$$(7.5) \qquad f\Lambda_2 = f\Lambda^2 + f\Lambda L.$$

Now $|f|$ is also a character, hence it follows by (6.5a) that:

$$\begin{aligned} x^{-1}\sum |f(w)|\,\Lambda(w) &= I_{|f|\Lambda}x^{-1} = I_{|f|^{-1}}I_{|f|L}x^{-1} = I_{|f|^{-1}}x^{-1}S_{|f|L}1 \\ &= I_{|f|^{-1}}x^{-1}[Ax\log x - Ax + A + O(x\log^{-\delta}x)] \\ &= AI_{|f|^{-1}}\log x - AI_{|f|^{-1}}1 + x^{-1}AS_{|f|^{-1}}1 + I_{|f|^{-1}}O(\log^{-\delta}x) \\ &= O(1) + O(\log^{1-\delta}x) = O(1) \end{aligned}$$

which follows immediately by Corollary 2 and Theorem 8. This gives the proof of (7.1). The proof of (7.2) follows similarly by use of (6.5b). Namely

$$\begin{aligned} x^{-1}\sum |f(w)|\,\Lambda_2(w) &= I_{|f|^{-1}}(I_{|f|L^2}x^{-1}) \\ &= I_{|f|^{-1}}[A\log^2 x - 2A\log x + 2A + O(\log^{1-\delta}x)] \\ &= 2\log x + O(1) + O(\log^{2-\delta}x), \end{aligned}$$

since by Theorem 2 $I_{|f|^{-1}}\log^2 x = (A^{-1}D + \ldots)\log^2 x + O(\log^{2-\delta}x)$.

The proof of (7.3) follows now by standard methods from (7.2) and (7.5). That is

https://doi.org/10.4153/CJM-1961-008-1 Published online by Cambridge University Press

$$0 \leqslant \sum_{x<Nw\leqslant xt} |f(w)|\,\Lambda(w) \leqslant \sum_{x<Nw\leqslant tx} |f(w)|\,\Lambda(w) \log Nw \log^{-1}x$$

$$\leqslant \log^{-1}x \sum_{x<Nw\leqslant tx} \Lambda_2(w) = \log^{-1}x(2tx \log tx - 2x \log x) + O(x \log^{-1}x) + O(x \log^{1-\delta}x) = (2t-1)x + o(x).$$

The "elementary proofs" lie in the following refinement of (**2**, Theorem 4).

THEOREM 10. *Let $g(w) \in C(W)$ be a non-negative function satisfying*

(g1) $$\sum_{Nw\leqslant x} g(w) = Mx \log^n x + o(x \log^n x); \qquad M > 0, n \geqslant 1.$$

Let $h(x)$ be a real a complex-valued function which satisfies

(h1) $$h(x) = O(1)$$

(h2) $$\sum_{\nu\leqslant x} \nu^{-1}h(\nu) = O(1)$$

(h3) $$h(tx) - h(x) = o(1) \quad \cdot\mathrm{s} \quad (t, x) \to (1, \infty).$$

Then the condition

(g2) $$|h(x)| \log^{n+1}x \leqslant \frac{n+1}{M} \sum_{Nw\leqslant x} \frac{g(w)}{Nw} \left|h\left(\frac{x}{Nw}\right)\right| + o(\log^{n+1}x)$$

implies

$$h(x) = o(1).$$

This theorem has been given in (**2**, Theorem 4) with the condition

$$\sum (Nw)^{-1}g(w) = a \log^{n+1}x + b \log^n x + o(\log^n x)$$

which is stronger than (g1), since (g1) implies only that

(7.6) $$I_g 1 = \sum (Nw)^{-1}g(w) = [Mx \log^n x + o(x \log^n x)]x^{-1} + \int_1^x [Mt \log^n t + o(t \log^n t]t^{-2}dt = (n+1)^{-1}M \log^{h+1}x + o(\log^{n+1}x).$$

To prove this theorem, we first observe that for given $t > 1$, we can find x_δ such that, for x, y satisfying $xy^{-1} > x_\delta$, the following is valid:

(7.7) $$\sum_{y\leqslant x/Nw\leqslant yt} (Nw)^{-1}g(w) > C \log^n(xy^{-1}) \quad \text{for some } C > 0.$$

Indeed, choose δ (to be fixed later) then there is x_δ such that for $x > x_\delta$, the absolute value of the error term in ($g1$) is $< \delta \log^n x$. Then it follows by (g1) that

$$\sum_{y\leqslant x/Nw\leqslant yt} (Nw)^{-1}g(w) \geqslant yx^{-1} \sum g(w)$$

$$\geqslant yx^{-1}[Mxy^{-1} \log^n xy^{-1} - Mx(yt)^{-1} \log^h x(yt)^{-1} - 2\delta xy^{-1}\log^n xy^{-1}]$$

$$\geqslant M(1-t^{-1}) \log^n xy^{-1} + Mt^{-1}[\log^h xy^{-1} - \log^n x(yt)^{-1}] - 2\delta \log^n xy^{-1}$$

$$> M(1 - t^{-1} - 2\delta) \log^n xy^{-1},$$

and (7.7) is true if we choose $C = M(1 - t^{-1} - 2\delta) > 0$, which can be fulfilled as $t > 1$.

By the standard method of Selberg's proof (**1**, Theorem 6.1 and **2**, Theorem 4) one can obtain the following.

(7.8) Given $\Delta > 0$, there exists x_Δ, $T > t > 1$ such that for $x > x_\Delta$, there is $y, x \leqslant y < yt \leqslant xT$ with the property that for all $y \leqslant z \leqslant yt$, $|h(z)| < \Delta$.

We turn now to the proof of Theorem 10, which contains only a more careful repetition of the proof of (**2**, Theorem 4).

Let $\lim\sup|h(x)| = A$. If $A > 0$, choose $\Delta = \frac{1}{2}A$ and fix x_Δ, $T > t > 1$ satisfying (7.8). For this given t we choose x_δ to satisfy (7.7). Now for given $\epsilon > 0$, let $|h(x)| < A + \epsilon$ for all $x > X_\epsilon$.

Denote by y_i the element y given in (7.8) for $x = T^i > x_\Delta$, that is, $T^i < y_i < y_i.t < T^{i+1}$ and put $\xi = \log x_0/\log T$ where $x_0 = \max(x_\epsilon, x_\Delta)$, and $\eta = \log(xx_\delta^{-1})/\log T$. Thus for each $\xi < i < \eta$, $T^i > x_0$ and $xT^{-i} > x_\delta$.

It follows now by (9.2) in view of (7.7), (7.8), and (7.6) that

$$\begin{aligned}
|h(x)| \log^{n+1}x \leqslant{}& (n+1)M^{-1} \sum_{x(Nw)^{-1} \leqslant x_0} (Nw)^{-1}g(w)|h(x/Nw)| \\
&+ (n+1)M^{-1}(A+\epsilon) \sum_{xx_0^{-1} < x(Nw)^{-1} \leqslant x} (Nw)^{-1}g(w) \\
&+ \sum_{\xi < i < \eta} \sum_{y_i < (Nw)^{-1}x < y_i t} [\Delta - (A+\epsilon)](Nw)^{-1}g(w) \\
\leqslant{}& K \sum_{xx_0^{-1} < Nw \leqslant x} (Nw)^{-1}g(w) + (n+1)M^{-1}(A+\epsilon)[M(n+1)^{-1}\log^{n+1}x \\
&+ o(\log^{n+1}x)] + [\Delta - (A+\epsilon)]C \sum_{\xi < i < \eta} \log^n(xT^{-i-1}),
\end{aligned}$$

since $\Delta - (A + \epsilon) < 0$ and $xy_i^{-1} > xT^{-i-1}$. Now, the first term is $\leqslant Kx_0x^{-1}(Mx\log^n x + o(x\log^n x)) = o(\log^{n+1}x)$. For the third term we have by Lemma 1,

$$\begin{aligned}
\sum_{\xi < i < \eta} \log^n(xT^{-1}\cdot T^{-i}) ={}& -[\xi]\log^n xT^{-\xi-1} + [\eta]\log^n xT^{-\eta-1} \\
&- \int_\xi^\eta [u]d\log^n xT^{-u-1} = (n+1)^{-1}\log^{-1}T\log^{n+1}x + o(\log^{n+1}x)
\end{aligned}$$

as follows immediately by standard method of replacing $[u]$ by u and noting that $T^\xi = x_0$, $T^\eta = xx_\delta^{-1}$.

Thus

$$|h(x)| \leqslant A + \epsilon + (\Delta - A - \epsilon)C/M\log T + o(1).$$

As $x \to \infty$ with $|h(x)| \to A$ we get $A \leqslant A + \epsilon + (\Delta - A - \epsilon)C'$, $C' > 0$. But this cannot be true for all $\epsilon > 0$ since $(\Delta - A)C' < 0$. This contradiction leads to the conclusion that $A = 0$.

We apply now Theorem 10 to the following function: $g(w) = |f(w)|\Lambda_2(w)$ and

$$(7.9)\qquad h(x) = x^{-1}\sum_{Nw\leqslant x} f(w)\Lambda(w) + \begin{cases} -1 & \text{in Case I} \\ 0 & \text{in Case II} \\ +1 & \text{in Case III} \end{cases}$$

since $2-\delta<1$, it follows by (7.2) that $|f(w)|\Lambda_2(w)$ satisfies (g1). Clearly, (7.1) means that $h(x)$ given in (7.9) satisfies (h1). Condition (h3) follows by (7.1) and (7.3),

$$|h(tx)-h(x)| \leqslant \Big|\sum_{Nw\leqslant x} f(w)\Lambda(w)(t^{-1}-1)x^{-1}\Big| + \Big|\sum_{x<Nw\leqslant tx} f(w)t^{-1}x^{-1}\Lambda(w)\Big|$$
$$\leqslant K(t^{-1}-1)+2(t-1)t^{-1}+o(1) = o(1) \quad \text{as} \quad (t,x)\to(1,\infty).$$

To obtain (h2) we put $\sigma=1$ in case I, $\sigma=0$ for case II, and $\sigma=-1$ in Case III:

$$\Big|\sum_{\nu\leqslant x}\nu^{-1}h(\nu)\Big| = \Big|\sum_{\nu\leqslant x}\nu^{-2}\sum_{Nw\leqslant\nu} f(w)\Lambda(w) - \sigma\sum_{\nu\leqslant x}\nu^{-1}\Big|$$
$$= \Big|\sum_{Nw\leqslant x} f(w)\Lambda(w)\cdot\sum_{Nw<\nu\leqslant x}\nu^{-2} - \sigma\log x + O(1)\Big|$$
$$= \Big|\sum f(w)\Lambda(w[(Nw)^{-1} - x^{-1} + O(Nw^{-2})] - \sigma\log x + O(1)$$
$$\leqslant \textstyle\sum_1 + \sum_2 + \sum_3 + O(1),$$

where

$$\textstyle\sum_2 = \Big|x^{-1}\displaystyle\sum_{Nw\leqslant x} f(w)\Lambda(w)\Big| = O(1) \text{ by (7.1), and}$$
$$\textstyle\sum_3 = \Big|\displaystyle\sum_{Nw\leqslant x}(Nw)^{-2}|f(w)|\Lambda(w)\Big| \leqslant \sum_{Nw\leqslant x}|f(w)|\log Nw(Nw)^{-2} = O(1),$$

as follows immediately by (6.5b).

We can conclude from Theorem 9 that $\sum_1 = |I_{f\Lambda} - \sigma\log x| = O(1) + O(\log^{2-\delta}x)$ which is $O(1)$ if $2-\delta<0$, that is, $\delta>2$ or $\gamma>3$. From this we can conclude Theorem A for $\gamma>3$. To obtain our result for $\gamma>2$ we need a refinement of Theorem 9, which we can carry out only in the following form.

THEOREM 11. *If f is a character satisfying* (6.1) *and* (6.3) *with* $A\neq 0$, *then*

$$I_{f\Lambda} = -F^{-1}(D)F'(D) + O(1) + \begin{cases} O(\log^{n-\delta}x) & \text{in Case I} \\ O(\log^{n+1-\delta}x) & \text{in Case II} \\ O(\log^{n+2-\delta}x) & \text{in Case III.} \end{cases}$$

Before proceeding with the proof of this theorem, we observe that with these results it follows now that $\sum_1 = O(1)$ if $\delta>1$ in Cases I and II, and only in Case III we have to assume that $\delta>2$. To complete the proof of our first main theorem, we establish

THEOREM 12. *If f is a character satisfying* (6.1) *and* (6.2) *and* $A \neq 0$, *then*

$$\sum_{Nw \leqslant x} f(w)\,\Lambda(w) = \begin{cases} x + o(w) & \text{in Case I and } \gamma > 2, \\ o(x) & \text{in Case II and } \gamma > 2, \\ -x + o(x) & \text{in Case III and } \gamma > 3. \end{cases}$$

We still have to prove the validity of (g2). Indeed by (6.5a)

$$I_f h(x) = I_f(x^{-1}S_{f\Lambda}1) = x^{-1}S_f S_{f\Lambda}1 - \sigma I_f 1$$
$$= x^{-1}S_{fL}1 - \sigma I_f 1 = \begin{cases} O(\log^{-\delta} x) - \alpha_0 & \text{in Cases II and III} \\ O(\log^{-\delta} x) - \alpha_1 - \alpha_0 & \text{in Case I} \end{cases}$$

since $I_f x^{-1} = x^{-1}S_f$ (as operators) and $S_{f\Lambda} = S_{f^{-1}}S_{fL}$. Hence if $a = -\alpha_0$ or $a = \alpha_{-1} - \alpha_0$,

$$I_{f\mu} \log x\, I_f h(x) = aI_{f\mu} \log x + I_{f\mu}O(\log^{1-\delta} x)$$
$$= O(1) + O(\log^{2-\delta} x) = o(\log x),$$

by Theorem 8 and Corollary 2. We now proceed similarly to (**1**, Lemma 6.1)

$$I_{f\mu} \log x\, I_f = I_{f\mu}(I_f \log x + I_{fL}) = \log x + I_{f\Lambda}.$$

It follows therefore that

$$|h(x)|\log x - I_{|f|\Lambda}|h(x)| \leqslant |(\log x + I_{f\Lambda})h(x)| = |I_{f\mu}\log x\, I_f h(x)| = o(\log x).$$

As in (**1**, Lemmas 6.3 and 6.5) we obtain

$$(\log^2 x - I_{|f|\Lambda_2})|h(x)| = (\log x + I_{|f|}\ )(\log x - I_{|f|}\ )|h(x)|$$
$$\leqslant (\log x + I_{|f|}\ )o(\log x) = o(\log^2 x).$$

That is,

$$|h(x)|\log^2 x \leqslant I_{|f|\Lambda_2}|h(x)| + o(\log^2 x)$$

which proves (g2), after verifying easily as in (**1**, Lemma 6.3) with the aid of Theorem 11 that $I_{|f|\Lambda}o(\log x) = o(\log^2 x)$.

We return now to the proof of Theorem 11. From Lemma 2 and (7.1) it follows that

$$I_{|f|\Lambda}O(\log^r x) = O(1) + O(\log^{r+1} x). \tag{7.9}$$

The proof is similar to the proof of Theorem 8. It follows by (6.5a) that

$$\alpha x \log x - \alpha x + \alpha + O(x \log^{-\delta} x) = S_{fL}1 = S_{f\Lambda}S_f 1$$
$$= S_{f\Lambda}[\alpha x + o(x/\log^{1+\delta} x)] = \alpha x I_{f\Lambda}1 + xI_{f\Lambda}O(\log^{-1-\delta} x)$$
$$= \alpha x I_{f\Lambda}1 + xO(\log^{-\delta} x) + O(x).$$

Thus, if $\alpha \neq 0$, we have: $I_{f\Lambda}1 = \log x + O(1) + O(\log^{-\delta} x)$ which yield $R_0(x; f\Lambda, -F'F^{-1}) = O(1) + O(\log^{-\delta} x)$.

It follows now from the relation $fL = f*f\Lambda$ and by (5.6) and Theorem 7 using induction that

$$R_{n+1}(x;f\Lambda, -F^{-1}F') = cI_{f\Lambda}R_n(x;f, F) + O\left(\sum_{j=0}^{n} |R_j\,(x;f\Lambda, -F'F^{-1})|\right)$$

$$+ O(1) + R_n(x;f\Lambda*f, -F^{-1}F'\cdot F)$$

$$= I_{f\Lambda}O(\log^{n-\delta}x) + O(1) + R_n(x;fL, -F') + O(\log^{n-\delta}x)$$

$$= O(\log^{n+1-\delta}x) + O(1),$$

which prove the first case of Theorem 11.

The other cases follow as in Theorem 8:

$$R_{n+p}(x;f\Lambda, -F^{-1}F') = cI_{f\Lambda}R_n(x;f, F) + O\left(\sum_{j=0}^{n+p-1} |R;\,(x;f\Lambda, -F^{-1}F')|\right)$$

$$+ O(1) + R_n(x;fL, -F') = O(\log^{n+1-\delta}x) + O(1).$$

In Case II, $p = 0$ and in Case III, $p = -1$, which readily imply by induction the other two cases of Theorem 11.

We conclude this section with the last case $\alpha = A = 0$ (which implies $\alpha_0 \neq 0$). Here we do not have to use Theorem 10. The proof of (7.1) which leads to (7.9) holds in this case, and consequently, Theorem 11 (Case II) is also valid. Writing

$$h(x) = x^{-1} \sum_{Nw\leqslant x} \Lambda(w) = x^{-1}S_{f\Lambda}1.$$

As in the first part of the proof of Theorem 12, we obtain

$$h(x)\log x + I_{f\Lambda}h(x) = I_{\mu f}\log x\, I_f(x^{-1}S_{f\Lambda}1) = I_{\mu f}x^{-1}\log x S_{fL}1$$

$$= I_{\mu f}O(\log^{1-\delta}x) = O(1) + O(\log^{2-\delta}x).$$

The power series in D corresponding to $I_{|f|}$ will be of the form $G(D) = A_0 + A_1D + \ldots$, and $A_0 \neq 0$ (by Case IV of Theorem 8). Thus it follows from Theorem 11 that

$$I_{|f|\Lambda}1 = -A_0{}^{-1}A_1 + O(\log^{1-\delta}x).$$

Thus, since $h(x) = O(1)$

$$|h(x)|\log x \leqslant I_{|f|\Lambda}|h(x)| + O(\log^{2-\delta}x) + O(1)$$

$$\leqslant O(1).I_{|f|\Lambda}1 + O(\log^{2-\delta}x) + O(1) = O(\log^{2-\delta}x) + O(1).$$

Consequently $|h(x)| \leqslant O(\log^{1-\delta}x) + O(\log^{-1}x)$. That is,

THEOREM 12*. *If f is a character satisfying*

$$\sum_{Nw\leqslant x} f(w) = O(x/\log^{1+\delta}x)$$

and

$$\sum_{Nw \leqslant x} |f(w)| = O(x/\log^{1+\delta} x),$$

then

$$\sum_{Nw \leqslant x} f(w)\,\Lambda(w) = O(x/\log^{1-\delta} x) + O(x/\log x).$$

8. Proofs of Theorems A and B. We return now to the situation of §2. Let H denote a generic class W/W' and let $e_H(w)$ be the characteristic function of H, that is, $e_H(w) = 1$ if $w \in H$ and zero otherwise. From the properties of characters (2.2) and (2.3) we have the relations

$$\chi = \sum_H \chi(H) e_H, \qquad e_H = h^{-1} \sum_\chi \bar{\chi}(H)\chi. \tag{8.1}$$

We assumed in (2.4) that

$$S_{e_H} 1 = \sum_{\substack{Nw \leqslant x \\ w \in H}} 1 = c_H x + O(x/\log^\gamma x), \qquad \sum c_H > 0. \tag{8.2}$$

Thus

$$S_\chi 1 = \sum_H \chi(H) S_{e_H} 1 = A_\chi x + O(x/\log^\gamma x), \qquad A_\chi = \sum_H \chi(H) c_H,$$

and for the identity $x_0 = E$. Also

$$S_{\chi_0} 1 = cx + O(x/\log^\gamma x), \qquad A_{\chi_0} = c = \sum c_H > 0.$$

The characters are thus functions of the type which were dealt with in the preceding sections. Let

$$L_\chi(D) = \sum_{\nu=-1}^{\infty} L_\nu(\chi) D^\nu; \qquad L_{-1}(\chi) = A_\chi$$

be the polynomial corresponding to I_χ in (6.6), then we distribute the characters of K in three classes

$$\Gamma_1 = \{\chi; \chi \in K, A_\chi = L_{-1}(\chi) \neq 0\},$$
$$\Gamma_2 = \{\chi; \chi \in K, A_\chi = 0, L_0(\chi) \neq 0\},$$
$$\Gamma_3 = \{\chi; \chi \in K, A_\chi = 0, L_0(\chi) = 0\}.$$

Theorem 8 now implies that (all characters are in our case subjected to CasesI–III):

Corollary 3.

$$I_{\chi\mu} = L_\chi^{-1}(D) + O(1) + \begin{cases} O(\log^{n-\delta} x), & \chi \in \Gamma_1 \\ O(\log^{n+1-\delta} x), & \chi \in \Gamma_2 \\ O(\log^{n+2-\delta} x), & \chi \in \Gamma_3. \end{cases}$$

From Theorem 12 we now obtain Theorem A.

https://doi.org/10.4153/CJM-1961-008-1 Published online by Cambridge University Press

Theorem A and (8.1) yield

$$\sum_{Nw\leqslant x} e_H(w)\,\Lambda(w) = \sum_{\substack{Nw\leqslant x\\ w\in H}} \Lambda(w) = h^{-1}\sum_{Nw\leqslant x} \chi(H)\,\bar{\chi}(w)\,\Lambda(w)$$

$$= h^{-1}\left[\sum_{x\in\Gamma_1} \chi(H) - \sum_{x\in\Gamma_3} \chi(H)\right] x + o(x) = d_H x + o(x).$$

From here we can follow the ideas developed in **(4)**, but replacing the "Dirichlet density" k of a set S, defined there, by the sum

$$\sum_{w\in S} d(w)^{-1}\Lambda(w) = k\log x + O(1),$$

by dealing in a parallel way with the sum

$$\sum_{w\in S} \Lambda(w) = kx + o(x)$$

and by calling k the Dirichlet density of the set S. As the reasoning is identical with that of **(**4, p. 602**)** as well as the passage from

$$\psi_H(x) = \sum_{\substack{Nw\leqslant x\\ w\in H}} \Lambda(w)$$

to

$$\pi_H(x) = \sum_{\substack{Np\leqslant x\\ p\in H}} 1$$

we just quote the final results.

(a) Γ_1 is a subgroup of index 1 or 2 in the group $\Gamma_1 \cup \Gamma_3$.

(b) Let $U = \{w; \chi(w) = 1, \chi \in \Gamma_1\}$, then $W' \subseteq U \subseteq W$ and U/W' has Γ_1 as the group of characters. Put $U^* = \{w; w \in U, \chi(w) = 1 \text{ for all } \chi \in \Gamma_3\}$ then $U \supseteq U^* \supseteq W'$ and Theorem B is valid for these groups U and U^*. (Compare with **(4**, Theorem 3.1**)**.)

9. Other "elementary results." In the present section we shall outline the extensions of the elementary proofs of **(1)** to our case. These lie in the following extension of **(1**, Theorem 9.2**)**.

THEOREM 13. *Let* $g \in C(W); G(D) = \gamma_{-q}D^{-q} + \gamma_{-q+1}D^{-q+1} + \ldots.$ *Let* f *be a character satisfying* (6.1) *and* (6.2). *Then for* $n < \delta - q - 1$

$$I_{f^{-1}*g} \log^n x = [F^{-1}(D)G(D)] \log^n x + o(1),$$

where $F(D)$ *is given in* (6.6), *provided that the following conditions hold:*

(i) $I_{f^{-1}}R_\nu(x; g, G) = O(1)$ for $\nu \leqslant n + 1$,

(ii) $I_{|f|} \log x\, R_n(x; g, G) = o(\log x)$

(iii) $I_{|h|}1 = O(x^\theta)$ where $h = f^{-1}*g$ and $\theta < 1$

https://doi.org/10.4153/CJM-1961-008-1 Published online by Cambridge University Press

(iv) $\sum_{x<Nw\leqslant tx} (Nw)^{-1}|h(w)| = o(1)$ as $(t, x) \to (1, \infty)$

(v) Case I and $q \leqslant 2$; Case II and $q \leqslant 1$; Case III and $q \leqslant 0$.

Before proceeding with this proof we give here some examples as applications.

Example 1. $g = \epsilon$ is the identity, and $G(D) = 1$. Thus $R_\nu(x; \epsilon, 1) = O(1)$ for all ν which yields (i) and (ii) trivially. Here, $h = f^{-1}*\epsilon = f^{-1}$ hence $I_{|h|}1 = O(\log x)$, by (6.3) and (6.2) which proves (iv). To prove (v) we obtain by (6.3) and (6.9) that for some $K > 0$,

$$O \leqslant \sum_{x<Nw\leqslant tx} (Nw)^{-1}|f^{-1}(w)| \leqslant K \sum_{x<Nw\leqslant tx} (Nw)^{-1}|f(w)| = KA \log t + O(\log^{-\delta} x)$$
$$= o(1)$$

as $(t, x) \to (1, \infty)$. Consequently,

COROLLARY 4.

$$I_{f^{-1}} \log^n x = F^{-1}(D)\log^n x + o(1) \quad \text{for} \quad n < \delta - 1.$$

In particular this yields, for $\delta > 1$:

$$I_{f^{-1}}1 = \sum_{Nw\leqslant x} \frac{\mu(w)f(w)}{Nw} = \begin{cases} o(1) & \text{in Case I} \\ \alpha_0^{-1} + o(1) & \text{in Case II} \\ \alpha_1^{-1}\log x + \alpha_2\alpha_1^{-1} + o(1) & \text{in Case III.} \end{cases}$$

This includes for the case $f(w) = 1$, one of the equivalent forms of the prime number theorem, but our effort to follow the classical proof of that theorem from this result failed, and we have obtained the prime number theorem in §8 in a different way.

Example 2. $g(w) = (fL)(w) = f(w) \log Nw$, and $G(D) = -F'(D)$. We shall consider only the case $A \neq 0$ (in Case I, $q = 2$, Cases II and III, $q \leqslant 0$). In this example $R_n(x; fL, -F') = O/\log^{n+1-\delta} x$ (Theorem 7) and thus Corollary 2 implies that

$$I_{f^{-1}}R_\nu(x; fL, -F') = O(\log^{\nu+2-\delta} x) + O(1) = O(1)$$

for all ν satisfying $\nu + 2 - \delta < 0$. This proves the validity of (i) for all $n + 3 - \delta < 0$. But then (ii) also holds since

$$I_f \log x\, R_n(x; fL, -F') = I_f O(\log^{n+2-\delta} x) = O(1) + O(\log^{n+3-\delta} x) = O(1).$$

Now $h = f^{-1}*fL = f\Lambda$ which implies the validity of (iii) by Theorem 9 when applied to $|f|$, since $I_{|f|}1 = O(\log x) + O(\log^{2-\delta} x)$. The last condition (iv) follows readily from (7.3). Consequently we obtain by applying Theorem 13 that

COROLLARY 5.

$$I_{f\Lambda} \log^n x = -[F^{-1}(D)F'(D)] \log^n x + o(1) \text{ if } \delta > n + 3.$$

https://doi.org/10.4153/CJM-1961-008-1 Published online by Cambridge University Press

In particular, if $\sigma = +1, 0, -1$ in Cases I, II, III respectively, then for $\delta > 3$:

$$\sum_{Nw \leqslant x} \frac{f(w)\,\Lambda(w)}{Nw} = \sigma \log x + \beta + o(1)$$

which is another equivalent form of the prime number theorem in the case of integers with $f(w) = 1$. Note that this is obtained only for $\delta > 3$ (that is, $\gamma > 4$) whereas the other equivalent for

$$\sum_{Nw \leqslant x} f(w)\,\Lambda(w) = \sigma x + o(x)$$

was obtained in Theorem 12 for $\gamma > 2$.

Now for the proof of Theorem 13. We wish to show that the function $h(x) = R_n(x; f^{-1}*g, F^{-1}G)$ satisfies the requirements of Theorem 10 with $g(w) = |f(w)|\Lambda_2(w)$. It was already proved in the preceding that $g(w)$ satisfies (g1) of Theorem 10 if $\delta > 1$, and we now prove the validity of (g2). It follows from Theorem 3 that

$$R_n(x; g, G) = R_n(x; f*(f^{-1}*g),$$

$$F(F^{-1}G)) = I_f R_n(x; f^{-1}*g, F^{-1}G) + \sum_{j=0}^{n+p} c_j\, R_j\,(x; f, F) + a,$$

where D^{-p} is the first power of D in $F^{-1}G$, and for some constants $c_j\, a$ ($a = 0$ if f is a character satisfying Cases II and III). Operating with $I_{f^{-1}} \log x$ on this result, we find

$$\begin{aligned} I_{f^{-1}} \log x\; I_f R_n(x; f^{-1}*g, F^{-1}G) \\ &= I_{f^{-1}} \log x R_n(x; g, G) - \sum_{j=0}^{n+p} c_j\, I_{f^{-1}} \log x R_j\,(x; f, F) - a I_{f^{-1}} \log x \\ &= o(\log x) + O(\log^{n+p+1-\delta} x) + O(1) + O(\log^{1-\delta} x) \\ &= o(\log x) \end{aligned}$$

if $n + p + 1 - \delta < 1$. Since $R_j\,(x; f, F) = O(\log^{j-\delta} x)$ by Theorem 6 the rest follows like the proof of Theorem 12, that is

$$I_{f^{-1}} \log x\; I_f\, R_n(x; f^{-1}*g) = (\log x + I_{f\Lambda}) R_n(x; f^{-1}*g) = o(\log x),$$

from which we deduce that

$$|R_n(x; f^{-1}*g)| \log^2 x \leqslant I_{|f|\Lambda_2} |R_n(x; f^{-1}*g)| + o(\log^2 x),$$

namely, (g2).

To prove conditions (h1) $-$ (h3), we first observe that

$$R_n(f^{-1}*g) = I_{f^{-1}} R_n(g) + \sum_{j=0}^{n+q} c_j\, R_j\,(f^{-1}) + a = O(1) + O(\log^{n+q-\delta} x)$$

where $a \neq 0$ if the situation is of Case III. This shows that

$$R_n(f^{-1}*g) = O(1) \text{ and } R_{n+1}(f^{-1}*g) = O(1)$$

if $n + 1 + q - \delta < 0$.

The completion of the proof follows the computation of (**1**, pp. 306–7). We do not repeat the computation but present the final result in the following proposition.

PROPOSITION 9. *Let* $l(w) \in C(W)$ *and*

$$L(D) = \sum_{\nu=-m}^{\infty} \gamma_\nu D^\nu.$$

Then

$$(9.1) \qquad \sum_{\mu\leqslant x} \mu^{-1} R_n(\mu; l, L) = (n+1)^{-1} R_{n+1}(x; l, L) + O(1) + O\Big(\sum_{Nw\leqslant x} (Nw)^{-2} |l(w)| \log^n Nw\Big)$$

and

$$(9.2) \qquad R_n(tx; l, L) - R_n(x; l, L) = \sum_{Nw\leqslant x} (Nw)^{-1} l(w) \log^n(tx/Nw) + \sum_{j=0}^{n-1} \binom{n}{j} \log^{n-j} t R_j(x; l, L) + O(\log^{m-1} x \log^{n+1} t),$$

woth the last factor omitted if $L(D)$ *does not contain negative powers of* D.

Thus in our case, $l = f^{-1} * g$, $L = F^{-1}G$ we observe that (iii) of Theorem 13 yields, by Lemma 2:

$$\sum_{Nw\leqslant x} (Nw)^{-2} |l(w)| \log^n Nw = O(x^\vartheta) x^{-1} \log^n x - \int_1^x O(t^\vartheta) d[t^{-1} \log^n t] = O(1)$$

which shows that $R_{n+1}(f^{-1}*g) = O(1)$ implies (h2). Condition (h3) of Theorem 10 follows from (iv), since in our case $R_j(x;l,L) = O(1)$ and $m = 1$ (in Case III, we have to require that $q = 0$) imply that

$$\sum_{x<Nw\leqslant tx} (Nw)^{-1} l(w) \log^n [tx/Nw] = o(1).$$

This completes the proof that if $n + q + 1 < \delta$ and (a) Case I with $q \leqslant 2$, (b) Case II, $q \leqslant 1$, or (c) Case III, $q \leqslant 0$, all conditions of Theorem 10 are fulfilled. Thus the proof of Theorem 13 is complete.

10. The character $f(n) = n^{it}$. We conclude our result with an application for the semi-group of integers and the character $f(n) = n^{it}$, $t \neq 0$ fixed.

Clearly f is a character, and satisfies $I_{|f|} 1 = \sum n^{-1} = \log x + c + O(x^{-1})$ and $I_f 1 = \sum n^{-1+it} = c_t + O(x^{-1})$ where $c_t = \zeta(1 - it)$. Thus this function f satisfies the condition of either Case II or Case III.

If (III) is valid then we would get from Theorem 13 (Example 2) that

$$I_{f^{-1}*fL} 1 = I_{f\Lambda} 1 = \sum_{n\leqslant x} \Lambda(n) n^{-1+it} = -\log x + c + o(1).$$

But since

$$\sum_{n \leqslant x} \frac{\Lambda(n)}{n} = \log x + c_0 + o(1)$$

(the case $f(n) = 1$) we have by Lemma 2

$$-\log x + c + o(1) = \sum_{n \leqslant x} \Lambda(n) n^{-1+it}$$

$$= [\log x + c_0 + o(1)]x^{it} - \int_1^x [\log u + c_0 + o(1)] du^{it}$$

$$= O(1) + \int_1^x o(1) du^{it} = o(\log x).$$

Indeed, if the function $|o(1)| < \epsilon$ for $x > \lambda$ and $|o(1)| < K$ for $x \leqslant \lambda$, then

$$\left| \int_1^x o(1) du^{it} \right| < Kt \int_1^\lambda u^{-1} du + t\epsilon \int_\lambda^x u^{-1} du < M + \epsilon t \log x. \qquad \text{Q.E.D.}$$

Thus Case III is disposed of and there remains Case II, which means that

$$\sum_{n=1}^{\infty} n^{-1+it} = \zeta(1 - it) \neq 0$$

and consequently that the series

$$\sum_{n=1}^{\infty} n^{-1+it} \Lambda(n)$$

converges for all $t \neq 0$. From this one readily proves that

$$\sum_p p^{-1+it} \log p$$

converges.

References

1. S. A. Amitsur, *On arithmetic functions*, J. d'Anal. Math., *5* (1956–7), 273–317.
2. ——— *Some results on arithmetic functions*, J. Math. Soc. Japan *11* (1959), 275–290.
3. A. Beurling, *Analyse de la loi asymptotique de la distribution des nombre premiers generalises*, Acta. Math., *68* (1937), 255–291.
4. H. Shapiro and W. Forman, *Abstract prime number theorems*, Comm. Pure and Applied Math., *7* (1954), 587–619.
5. K. Yamamoto, *Theory of arithmetic linear transformations and its applications to an elementary proof of Dirichlet's theorem*, J. Math. Soc. Japan, *7* (1955), 424–434.
6. ——— *Arithmetic linear transformations in an algebraic number field*, Mem. Fac. Sci., Kyushu Univ. Sci. A. *12* (1958), 41–66.
7. Hardy and Wright, *The theory of numbers* (Oxford, 1954).

Hebrew University