[4] E. Landau, *Einführung in die Elementare und Analytische Theorie der Algebraischen Zahlen und der Ideale*, second edition, Leipzig and Berlin 1927.

[5] B. V. Levin, *Estimates from below for the number of nearly-prime integers belonging to some general sequence*, Vestnik Leningrad Univ. 15, no. 7 (1960), pp. 48-65 (Russian).

[6] H. B. Mann, *Introduction to algebraic number theory*, Columbus, Ohio 1955.

[7] K. Prachar, *Primzahlverteilung*, Berlin-Gottingen-Heidelberg 1957.

[8] A. Selberg, *On an elementary method in the theory of primes*, Norske Vid. Selsk. Forh., Trondheim 19, no. 18 (1947), pp. 64-67.

[9] — *The general sieve-method and its place in prime number theory*, Proceedings of the International Congress of Mathematicians, Cambridge, Mass., 1950, vol. 1, pp. 286-292. Amer. Math. Soc., Providence, R. I., 1952.

[10] — *On elementary methods in prime number-theory and their limitations*, Den 11te Skandinaviske Matematikerkongress, Trondheim, 1949, pp. 13-22, Oslo 1952.

[11] Y. Wang, *On sieve methods and some of their applications*, Sci. Record (n. S.) 1 (1957), no. 3, pp. 1-5.

Reçu par la Rédaction le 10.4.1963

ACTA ARITHMETICA
X(1964)

The general sieve

by

N. C. Ankeny and H. Onishi (Cambridge, Mass.)

Introduction. The sieve is a method used to derive bounds on the number of elements in a set of integers which are not divisible by any prime number in another set.

Let us suppose we are given a set S of integers, a set T of prime numbers, and $M(S,T)$ denotes the number of integers in S not divisible by any prime in T. We would now like to derive bounds on $M(S,T)$. For example, if $S_N = \{m \mid m \leqslant N\}$, $T_{\sqrt{N}} = \{p \mid p \leqslant \sqrt{N}\}$ where p ranges over all primes and N is positive, then $M(S_N, T_{\sqrt{N}})$ equals the number of prime numbers $> \sqrt{N}$ and $\leqslant N$.

To formulate the problem more precisely, define

(i) S_N as a set of N integers, for every positive integer N,

(ii) T as an infinite set of primes, T_Y as the set of primes in T less than a real number Y.

We are prepared now to observe the behavior of the function $M(S_N, T_{N^\lambda})$ for some fixed $\lambda > 0$, as $N \to \infty$. In order to do this we impose restrictions on the sets S_N, T_Y. These restrictions cover not only the classical cases of the sieve, but also several new cases.

Let d denote a square free integer all of whose prime factors are in T. We require the following assumptions:

(A) *For each N, there exists a real valued positive multiplicative function $f_N(d)$ such that*

$$\sum_m 1 = Nf_N(d)^{-1} + R_d(N), \qquad m \in S_N, \quad d \mid m$$

(*i.e.* $f(d_1 d_2) = f(d_1) f(d_2)$ *when* $(d_1, d_2) = 1$).

(B) *There exist positive real numbers* a, δ, C_1, C_2 *such that* $f_N(p)^{-1} < 1-\delta$ *for all* $p \in T$,

$$\sum_{p<X} p f_N(p)^{-1} < C_1 X (\log X)^{-1} \quad \textit{for} \quad X \leqslant \log N,$$

$$\Big|\sum \big(p f_N(p)^{-1} - a\big)\Big| < C_2 X (\log X)^{-2} \quad \textit{for} \quad \log N < X < Y.$$

(C) *$\exists$ a β_1 such that if $\nu(d)$ denote the number of prime factors of d, then*

$$\sum C_3^{\nu(d)}|R_d| = O\big(N(\log N)^{-a-2}\big), \qquad d < N^{\beta_1},\ C_3 = 6\delta^{-2}.$$

We let $\limsup \beta_1 = \beta$.

First, a few remarks about these assumptions should be made. Although the assumptions are listed separately they are interdependent and must be considered simultaneously. For example, (A) would be meaningless without (C).

Roughly speaking, $f(d)^{-1}$ could be considered as the probability measure that an element in S is divisible by d. The condition that $f(d)$ is multiplicative indicates that, if $(d_1, d_2) = 1$, the probability of being divisible by d_1 is independent of the probability of being divisible by d_2.

Specifically however, the presence of the "error term" $R_d(N)$ indicates that the problem cannot be formulated in a completely probabilistic manner. Hence, the ordinary probability argument is not applicable, except in a heuristic manner. This is then the basis for the sieve, to modify the probabilistic method to dampen the error terms $R_d(N)$.

The interpretation of (C) is first to bound the probability $f_N(p)^{-1}$ away from 1, and to guarantee some uniformity of the probability density $\sum f_N(p)^{-1}$.

To derive our bounds on $M(S_N, T_{N^\lambda})$ we demand that the numbers $a, \beta, \delta, C_1, C_2$, are independent of N, although the functions $f_N(d)$ may change with N.

In order to state the main result, let $\Gamma(a)$ denote the Γ function, γ is Euler's constant, and

$$B_a(N) = \Gamma(a)\prod\big(1-f(p)^{-1}\big)(1-p^{-1})^{-a}, \quad p\in T \text{ and } p < N^\lambda.$$

THEOREM 1. *Suppose that the sets $\{S_N\}$ and $\{T_{N^\lambda}\}$ satisfy* (A), (B), *and* (C). *Then as $N\to\infty$,*

$$M(S_N, T_{N^\lambda}) \leqslant B_a(N)\,N(\log N)^{-a}\lambda^{-a}J_a(\tfrac12\beta\lambda^{-1})^{-1}\big(1+o(1)\big),$$

$$M(S_N, T_{N^\lambda}) \geqslant \big(\Gamma(a)e^{a\gamma}\big)^{-1}B_a(N)\,N(\log N)^{-a}\lambda^{-a}\{1-G_a(\tfrac12\beta\lambda^{-1})\}\big(1+o(1)\big).$$

Also there exists a constant K such that

$$(\log\log N)^{-K} < B_a(N) < (\log\log N)^{K}.$$

The functions $J_a(u)$ and $G_a(u)$ are continuous functions of u (see Chapter II), and

$$\lim_{u\to\infty} J_a(u) = \Gamma(a)e^{a\gamma}, \qquad \lim_{u\to\infty} G_a(u) = 0.$$

If $0 < u \leqslant 1$, then $J_a(u) = a^{-1}u^a$.

For a fixed $a > 0$, the function $1-G_a(x)$ has a unique simple zero $x = \zeta_a$. Hence, $M(S_N, T_{N^\lambda}) > 0$ for $\lambda^{-1} > 2\beta^{-1}\zeta_a$ and N sufficiently large. If $x > \zeta_a$, we can easily rewrite

$$1-G_a(X) = \big(\Gamma(a)e^{a\gamma}\big)ax^{-a}\int_{\zeta_a}^{x} J_a(u-\tfrac12)^{-1}u^{a-1}\,du$$
$$> \big(\Gamma(a)e^{a\gamma}\big)J_a(x-\tfrac12)^{-1}\big(1-(\zeta_a x^{-1})^a\big).$$

Hence, for a λ with the property that $\frac12\beta\lambda^{-1} > \zeta_a+1$, the upper and lower bounds are fairly close.

Thus, to complete our knowledge of the lower bound of the sieve we must have at least an upper bound on ζ_a. In Chapter II we prove

$$\lim a^{-1}\zeta_a = 1.22\ldots$$

For $a = 1, 1.5, 2, 2.5$, and 3 we have computed ζ_a (see tables). The tables indicate that $a^{-1}\zeta_a$ rapidly approaches 1.22… In Chapter II we also have proved a uniform upper bound on ζ_a.

Theorem 1 has thus reduced the sieve bounds to fairly simple formulae, where the only invariants are a, β, and $B_a(N)$. Also, $B_a(N)$ is independent of N in most applications.

Some results can be derived if condition (C) is weakened. For example,

(C_1) *If* $\sum_{p<x} f_N(p)^{-1} = o\big(X(\log X)^{-1}\big)$, *for* $\log N < X < N^\lambda$, *then* $M(S_N, T_{N^\lambda}) > 0$ *for any fixed $\lambda > 0$ and N sufficiently large.*

(C_2) *If* $\limsup_x\big(\sum_{p<x} f_N(p)^{-1}X^{-1}(\log X)\big) = a$, $\log N < X < N^\lambda$, *then Theorem 1 is still applicable.*

Because of the paper's length, it might be useful to give a brief outline.

1) Chapter I gives the method of A. Selberg to derive an upper bound on $M(S, T)$ in terms of sums involving $f_N(n)$. We prove that this immediately gives a lower bound involving these sums.

2) Chapter II studies properties of a family of functions $\tau_a(u)$ defined by a difference-differential equation. Mainly we need information about the asymptotic behavior of $\tau_a(u)$, and this is applied to yield our results concerning $J_a(u)$, $G_a(u)$, and ζ_a. This chapter is independent of the previous results and is of some interest in itself.

3) Chapter III combines the results of the two previous chapters, and completes the proof of Theorem 1.

4) Chapter IV gives several applications of Theorem 1. In the applications it is often possible to sharpen the results by using both the lower and upper bound. (See especially the first example of Chapter IV). It is then possible to derive a stronger result than by only using the lower bound.

I. Selberg's sieve

§ 1. The upper bound. Let $S = S_N$, $T = T_Y$ be sets satisfying assumptions (A), (B), and (C). We shall define a set of variables $\{\varrho_d\}$ in a region defined by "Möbius inequalities" (1.1) and (1.2). An upper bound for $M(S, T)$ can then easily be stated in terms of ϱ_d (1.3). Our problem is then to find a minimum of the linear function $\sum_d \varrho_d f(d)^{-1}$ consistent with our inequalities. There are several ways of doing this, the most effective is that of A. Selberg. Namely, we replace $\{\varrho_d\}$ by new variables $\{\lambda_d\}$ *via* a quadratic transformation. Condition (1.1) is then automatically satisfied, and (1.2) is replaced by a stronger condition (1.5). The linear function $\sum_d \varrho_d f(d)^{-1}$ becomes a positive definite quadratic form in the variables $\{\lambda_d\}$. The minimum of the quadratic form subject to the conditions (1.5) can then be found by standard methods.

The lower bound, or more sophisticated sieves, can then be automatically computed from the upper bound, and these bounds are derived in the latter part of this chapter. Thus, whatever fault there is in the upper bound (replacing (1.2) by (1.5)) is compounded for the lower bound. The methods are still effective, but whatever the best possible results are remain a mystery except in a few cases.

Let $\{\varrho_d\}$, $d|P$, $P = \prod_{p \in T_Y} p$, denote a set of variables satisfying

$$\varrho_1 = 1, \quad \sum_{d|t} \varrho_d \geqslant 0, \quad \text{for all } t|P, \tag{1.1}$$

$$\varrho_d = 0 \quad \text{for} \quad d > z, \tag{1.2}$$

where z will be chosen later.

We shall let d, d_1, d_2, s, t be positive integers dividing P, m runs over all elements of S. Then

$$M(S, T) = \sum_{(m,P)=1} 1 = \sum_m \sum_{d|m} \mu(d) \leqslant \sum_m \sum_{d|m} \varrho_d = \sum_d \varrho_d \sum_{d|m} 1 \tag{1.3}$$
$$= \sum_d \varrho_d\{Nf(d)^{-1} + R_d\} \leqslant N \sum_d \varrho_d f(d)^{-1} + E$$

where $E = \sum_d |\varrho_d R_d|$.

Define the variables $\{\varrho_d\}$ in terms of new variables $\{\lambda_d\}$ by

$$\varrho_d = \sum_{d_1, d_2} \lambda_{d_1}\lambda_{d_2}, \quad [d_1, d_2] = d \tag{1.4}$$

($[d_1, d_2]$ denotes the least common multiple of d_1 and d_2) where we impose the condition

$$\lambda_1 = 1, \quad \lambda_d = 0 \quad \text{for} \quad d > \sqrt{z}. \tag{1.5}$$

Now

$$\sum_{d|t} \varrho_d = \sum_{d|t} \sum_{[d_1,d_2]|d} \lambda_{d_1}\lambda_{d_2} = \Big(\sum_{d|t} \lambda_d\Big)^2 \geqslant 0.$$

Also if $d > z$, $[d_1, d_2] = d$, then either $d_1 > \sqrt{z}$ or $d_2 > \sqrt{z}$. Hence, $\varrho_d = 0$ for $d > z$. Thus in terms of $\{\lambda_d\}$ the $\{\varrho_d\}$ defined by (1.4) automatically satisfy (1.1) and (1.2).

We now wish to minimize (1.3).

If $s = (d_1, d_2)$ (the greatest common divisor of d_1 and d_2), then $f([d_1, d_2]) = f(d_1)f(d_2)f(s)^{-1}$. Define

$$f'(t) = \sum_{s|t} \mu(t/s) f(s).$$

By the Möbius inversion formula,

$$f(s) = \sum_{t|s} f'(t).$$

Hence,

$$\sum_d \varrho_d f(d)^{-1} = \sum_d f(d)^{-1} \sum_{[d_1,d_2]=d} \lambda_{d_1}\lambda_{d_2} = \sum_{d_1,d_2} \lambda_{d_1}\lambda_{d_2} f([d_1, d_2])^{-1} \tag{1.6}$$
$$= \sum_{d_1,d_2} \lambda_{d_1} f(d_1)^{-1} \lambda_{d_2} f(d_2)^{-1} f((d_1, d_2))$$
$$= \sum_{d_1,d_2} \lambda_{d_1} f(d_1)^{-1} \lambda_{d_2} f(d_2)^{-1} \sum_{t|(d_1,d_2)} f'(t)$$
$$= \sum_t f'(t) \Big\{\sum_{t|d} \lambda_d f(d)^{-1}\Big\}^2.$$

Note for a fixed s,

$$\sum_{s|t} \mu(t) \sum_{t|d} \lambda_d f(d)^{-1} = \sum_{s|d} \lambda_d f(d)^{-1} \sum_{s|t|d} \mu(t) = \mu(s)\lambda_s f(s)^{-1}. \tag{1.7}$$

If we apply Schwarz's inequality to (1.7) for $s = 1$, we have using (1.5)

$$1 = \Big\{\sum_t \mu(t) \sum_{t|d} \lambda_d f(d)^{-1}\Big\}^2$$
$$= \Big\{\sum_t \mu(t) f'(t)^{-1/2} f'(t)^{1/2} \sum_{t|d} \lambda_d f(d)^{-1}\Big\}^2$$
$$\leqslant \Big(\sum_{t \leqslant \sqrt{z}} f'(t)^{-1}\Big) \sum_t f'(t) \Big\{\sum_{t|d} \lambda_d f(d)^{-1}\Big\}^2$$

or

$$\sum_t f'(t) \Big\{\sum_{t|d} \lambda_d f(d)^{-1}\Big\}^2 \geqslant A \tag{1.8}$$

for

$$A = \Big\{\sum_{t<\sqrt{z}} f'(t)^{-1}\Big\}^{-1}.$$

Conversely, if we let

(1.9) $$\lambda_d f(d)^{-1} = \mu(d)\Big(\sum_{d|s} f'(s)^{-1}\Big)A, \qquad s<\sqrt{z/d},$$

then, for a fixed t,

$$\sum_{t|d}\lambda_d f(d)^{-1} = \sum_{t|d}\mu(d)\Big(\sum_{d|s} f'(s)^{-1}\Big)A = A\sum_{t|s} f'(s)^{-1}\sum_{t|d|s}\mu(d) = A\mu(t)f'(t)^{-1}.$$

Thus, for λ_d defined by (1.9),

(1.10) $$\sum_t f'(t)\Big(\sum_{t|d}\lambda_d f(d)^{-1}\Big)^2 = \Big(\sum_t f'(t)^{-1}\Big)A^2 = \Big\{\sum_{t<\sqrt{z}} f'(t)^{-1}\Big\}^{-1}.$$

Also by (1.9)

$$|\lambda_d| = f(d)f'(d)^{-1}\Big(\sum_{s<\sqrt{z/d}} f'(d)^{-1}\Big)A \le f(d)f'(d)^{-1}.$$

And

$$\Big|\sum_{[d_1,d_2]=d}\lambda_{d_1}\lambda_{d_2}\Big| \le \sum_{d_1|d}|\lambda_{d_1}|\sum_{t|d_1}|\lambda_{dt/d_1}|$$
$$\le f(d)f'(d)^{-1}\sum_{d_1|d}\sum_{t|d_1} f(t)f'(t)^{-1}$$
$$= f(d)f'(d)^{-1}2^{\nu(d)}\sum_{t|d} f(t)f'(t)^{-1}2^{-\nu(t)}$$
$$= f(d)f'(d)^{-1}2^{\nu(d)}\prod_{p|d}\big(1+\tfrac{1}{2}f(p)f'(p)^{-1}\big)$$
$$< \prod_{p|d} 2\big(1+f(p)^{-1}\big)\big(1+\tfrac{1}{2}f(p)f'(p)^{-1}\big)$$
$$< (6\delta^{-2})^{\nu(d)} = C_3^{\nu(d)}$$

by the definition of C_3 in Assumption (C). Hence,

(1.11) $$E_N(z) = \sum_{d<z}|\varrho_d||R_d| < \sum_{d<z} C_3^{\nu(d)}|R_d|.$$

To state our result in final form let

$$Q(T) = \Big\{n \,|\, n=\prod_j p_j^{e_j},\ p_j\in T,\ e_j\ge 0\Big\},\qquad f(n)=\prod_j f(p_j)^{e_j}.$$

Then

$$\sum_{t<\sqrt{z}} f'(t)^{-1} = \sum_t f(t)^{-1}\prod_{p|t}\big(1-f(p)^{-1}\big)^{-1} \ge \sum_{n<\sqrt{z}} f(n)^{-1},\qquad n\in Q.$$

Combining (1.6) and (1.10) with (1.11) yields

(1.12) $$M(S,T)\le N\Big\{\sum f(n)^{-1}\Big\}^{-1}+E(z)$$

where $n<\sqrt{z}$, $n\in Q$; $d<z$, $d|P$.

Let us now utilize assumption (C). As $T=\{p\,|\,p<Y\}$ and $Y<N$, if we let $z=N^{\beta_1}$ in (1.12), we have

THEOREM 1.1. *$M(S_N,T_N)\le N\{\sum_n f(n)^{-1}\}^{-1}+O\big(N(\log N)^{-a-2}\big)$ where $n\le N^{\beta_1/2}$, $n\in Q$.*

§ 2. We shall now utilize (1.12) to derive other sieve bounds. The sets S and T are as before.

DEFINITION. Let $h|P$, $\nu(h)$ the number of distinct prime divisors of h, $p_T(m)$ the largest prime divisor of m in T; then

$$S_N(h) = S(h) = \{m\,|\,h|m,\ m\in S_N\}.$$

For every $p\in T_{N^\lambda}$, consider the set of $m\in S_N$ for which $p|m$ and m has no smaller prime divisor in T_{N^λ}. The number of elements in such a set is $M\big(S(p),T_p\big)$. These sets are obviously disjoint for distinct p, and the union is the set of m which have at least one prime divisor in T_{N^λ}. Hence,

(1.13) $$M(S_N,T_{N^\lambda}) = N-\sum_p M\big(S(p),T_p\big),\qquad p\in T_{N^\lambda}.$$

If $(d,p)=1$, $d|P$, then

$$\sum_{\substack{m\in S(p)\\ d|m}} 1 = \sum_{\substack{m\in S\\ dp|m}} 1 = \big(Nf(p)^{-1}\big)f(d)^{-1}+R_{pd}.$$

Thus, (A) holds for the sets $S(p)$ and T_p. We can now apply (1.12), changing z into zp^{-1}, and yielding,

(1.14) $$M(S_p,T_p)\le Nf(p)^{-1}\Big\{\sum_n f(n)^{-1}\Big\}^{-1}+\sum_d|R_{dp}|C_3^{\nu(d)}$$

where $n<\sqrt{z/p}$, $n\in Q(T_{N^\lambda})$, $d<zp^{-1}$, $p_T(d)<p$. We note

$$\sum_p\sum_{d<zp^{-1}}|R_{dp}|C_3^{\nu(d)} \le \sum_{d<z}|R_d|\,\nu(d)C_3^{\nu(d)} < \Big(\sum_{d<z}|R_d|C_3^{\nu(d)}\Big)(\log z).$$

Hence, if we sum (1.14) over all $p\in T_{N^\lambda}$, and use (1.13) we have proved for $z=N^{\beta_1}$, THEOREM 1.2. *We have*

$$M(S_N, T_{N^\lambda}) \geqslant N\Big\{1-\sum_p f(p)^{-1}\Big(\sum_n f(n)^{-1}\Big)^{-1}\Big\}+O\big(N(\log N)^{-a-1}\big)$$

where $p \in T_{N^\lambda}$, $n \leqslant (N^{\beta_1}p^{-1})^{1/2}$, $n \in Q(T_p)$.

For more complicated sieves we would proceed as follows. Let $M_r(S, T)$ denote the number of elements in S having at most r prime factors in T;

$$T(h) = \{p \mid p \in T_{N^\lambda},\ p \nmid h\};$$

$$T'(h) = \{p \mid p \in T_{N^\lambda}, p \nmid h, p < p_T(h)\}.$$

It is then easy to prove that, for $h \mid P$,

$$M_r(S, T) = \sum_h M\big(S(h),\ T(h)\big);\qquad \nu(h) \leqslant r, \tag{1.15}$$

$$M(S, T) = \sum_h M\big(S(h),\ T^1(h)\big);\qquad \nu(h) = r+1. \tag{1.16}$$

We could then utilize the upper bound (1.12) to derive upper and lower bounds on $M_r(S, T)$ by the above method. A slight revision is needed in the definition of β_1, as our error term is slightly larger. We leave these proofs to the reader.

Our immediate task will be to evaluate the sums appearing in the above theorems, and particularly to find a criterion in Theorem 1.2 to guarantee $M(S, T) > 0$. This objective will be pursued in the following two chapters, and the final results appear at the end of Chapter III.

II. The functions $\tau_a(u)$

§1. The functions $\tau_a(u)$. This chapter concerns the family of functions $\{\tau_a(u)\}$ defined below, and various functions derived from $\tau_a(u)$. We shall couple these results with the sieve in Chapter III, but this chapter will be independent from Chapter I. Only an elementary knowledge of analysis is required of the reader, except for the evaluation of $\int_0^\infty \tau_a(t)\,dt$ which is derived by Laplace transforms.

DEFINITION. For every $a > 0$ define the function

$$\tau_a(u) = \begin{cases} 0 & \text{if} \quad u \leqslant 0, \\ u^{a-1} & \text{if} \quad 0 < u \leqslant 1; \end{cases} \tag{2.1}$$

$$\tau_a'(u) = -u^{-1}\{a\tau_a(u-1)-(a-1)\tau_a(u)\}. \tag{2.2}$$

($\tau_a(u)$ is to be continuous at $u = 1$.)

If $a > 1$, $\tau_a(u)$ is a continuous function everywhere and differentiable except at $u = 0$. If $a \leqslant 1$, $\tau_a(u)$ is continuous except at $u = 0$, and differentiable except at $u = 0$ and 1.

Thus our functions $\tau_a(u)$ are integrable over any finite interval.

We may restate the difference-differential equation (2.2) in the equivalent form

$$ua^{-1}\tau_a(u) = \int_{u-1}^{u} \tau_a(t)\,dt. \tag{2.3}$$

The equivalence of (2.2) and (2.3) is seen by taking the derivative of both sides of (2.3), noting it satisfies (2.2), and that the two functions agree for $0 \leqslant u \leqslant 1$.

Next we prove $\tau_a(u) > 0$ for all $u > 0$. If not, let u_1 be the greatest lower bound of $u > 0$ for which $\tau_a(u) \leqslant 0$. By (2.1), $u_1 \geqslant 1$, and by continuity of our function for $u > 1$, $\tau_a(u_1) = 0$. However, by (2.3),

$$0 = u_1 a^{-1}\tau_a(u_1) = \int_{u_1-1}^{u_1} \tau_a(t)\,dt.$$

This implies that $\tau_a(u) \leqslant 0$ for some $u < u_1$, a contradiction to the choice of u_1, thus u_1 is nonexistent.

If $a \leqslant 1$ then, by (2.2), $\tau_a'(u) \leqslant 0$ for $u > 0$, or $\tau_a(u)$ is monotonically decreasing for $u > 0$. By (2.1) this is not obviously not the case when $a > 1$. By (2.3),

$$\tau_a(a) = \int_{a-1}^{a} \tau_a(t)\,dt. \tag{2.4}$$

Hence, $\tau_a'(u)$ has at least one zero in the range $(a-1, a)$. We prove

LEMMA 2.1. *If* $a > 1$, $\tau_a'(u) = 0$ *has a unique simple zero. Call this root* u_a. *Then* $\max(1, a-1) < u_a < a$.

Proof. Differentiating (2.2) gives

$$\tau_a''(u) = -u^{-1}\{a\tau_a'(u-1)-(a-2)\tau_a'(u)\}. \tag{2.5}$$

Let u_a be the least zero of $\tau_a'(u)$, we shall prove that it is a unique and simple root. If u_a were a multiple root, then by (2.5), $\tau_a'(u_a-1) = 0$, a contradiction. Now let v_1 be the next smallest zero of $\tau_a'(u)$. If $v_1 \geqslant u_a+1$, substitute $u = v_1$ in (2.2), giving $0 = (a-1)\tau_a(v_1)-a\tau_a(v_1-1)$, or $\tau_a(v_1) > \tau_a(v_1-1)$. This is false as $\tau_a(u)$ is decreasing in the range $[u_a, v_1]$ and $u_a \leqslant v_1-1 < v_1$. If $u_a < v_1 < u_a+1$, substitute $u = v_1$ in (2.5), giving $\tau_a''(v_1) = -av_1^{-1}\tau_a'(v_1-1)$. As $\tau_a'(u) > 0$ for $u < u_a$, we have $\tau_a''(v_1) < 0$. But $\tau_a'(u) < 0$ for $u_a < u < v_1$. Thus $\tau_a''(v_1) \geqslant 0$, again proving that v_1 does not exist. This completes the proof of Lemma 2.1 by the discussion at equation (2.4). (Note that $u_a \geqslant 1$ by (2.1).)

Next we prove that $\tau_a(u)$ is integrable over the range $[0, \infty]$. By Lemma 2.1, $\tau_a(u)$ is decreasing for $u > a$, hence by (2.3) for $u \geqslant a+1$

$$\tau_a(a) \geqslant \tau_a(u-1) > \int_{u-1}^{u} \tau_a(t)\,dt = ua^{-1}\tau_a(u).$$

Hence, for $u \geqslant a+3$,

$$\tau_a(u) < a^3\tau_a(a)\{u(u-1)(u-2)\}^{-1}.$$

This immediately implies the integrability of $\tau_a(u)$ over the range $[0, \infty]$.

§ 2. The functions $F_a(u)$. Define the function $F_a(u)$ by

$$F_a(u) = \int_u^\infty \tau(t)\,dt. \tag{2.6}$$

We are going to concern ourselves with the rate of decrease of $F_a(u)$ for large u. In terms of $F_a(u)$ we may rewrite (2.3) as

$$-F_a'(u) = au^{-1}\{F_a(u-1)-F_a(u)\}. \tag{2.7}$$

Using (2.7) and integrating by parts we infer

$$\int_u^\infty F_a(t)\,dt = -uF_a(u) - \int_u^\infty tF_a'(t)\,dt$$

$$= -uF_a(u) + a\int_u^\infty \big(F_a(t-1)-F_a(t)\big)\,dt$$

$$= -uF_a(u) + a\int_{u-1}^u F_a(t)\,dt$$

or

$$a\int_{u-1}^u F_a(t)\,dt > uF_a(u).$$

If $u > a+1 > u_a+1$, then $F_a(t)$ is convex in the interval $[u-1, u]$ (i.e. $F_a'(t) < 0$, $F_a''(t) > 0$). Hence, for $u > a+1$,

$$\tfrac{1}{2}\big(F_a(u-1)+F_a(u)\big) > \int_{u-1}^u F_a(t)\,dt > a^{-1}uF_a(u)$$

or

$$\begin{aligned} -F_a'(u) &= au^{-1}\big(F_a(u-1)-F_a(u)\big) \\ &= au^{-1}\big(F_a(u-1)+F_a(u)-2F_a(u)\big) \\ &> au^{-1}\big(2a^{-1}uF_a(u)-2F_a(u)\big) \\ &= 2(1-au^{-1})F_a(u). \end{aligned} \tag{2.8}$$

LEMMA 2.2. *If $a \leqslant 1$, then $F_a(u) \leqslant F_a(1)e^{1-u}$ for $u \geqslant 1$. If $a > 1$, then $\exists \xi_a$ such that $\xi_a < (e-1)a$ and*

$$F_a(u)e^u \leqslant F_a(\xi_a)e^{\xi_a} \quad \textit{for} \quad u \geqslant \xi_a.$$

Proof. If $a \leqslant 1$, then $\tau_a(u)$ is a decreasing function for $u > 0$. Hence, by (2.3),

$$ua^{-1}\tau_a(u) < \tau_a(u-1) \quad \text{for} \quad u > 1.$$

Thus, by (2.2), $-\tau_a'(u) \geqslant au^{-1}\tau_a(u-1) > \tau_a(u)$. Hence,

$$F_a(u) = \int_u^\infty \tau_a(t)\,dt < -\int_u^\infty \tau_a'(t)\,dt = \tau_a(u) = -F_a'(u).$$

We immediately have the first part of Lemma 2.2.

If $a > 1$, we note by (2.9) that for $u > 2a$,

$$-F_a'(u)F_a(u)^{-1} > 1.$$

If $u < 1$,

$$-F_a'(u)F_a(u)^{-1} = u^{a-1}\big(F_a(0)-a^{-1}u^a\big)^{-1},$$

or

$$-F_a'(u)F_a(u) < 1 \quad \text{for sufficiently small } u > 0.$$

Let u_1, u_2 be respectively the smallest and the largest zeros of $-F_a'(u)F_a(u)^{-1} = 1$. Let $\xi_a = \xi$ be the real number in the interval $[u_1, u_2]$ such that

$$F_a(\xi)e^\xi = \max_{u_1 \leqslant u \leqslant u_2} \big(F_a(u)e^u\big).$$

We now prove that

$$F_a(u)e^u \leqslant F_a(\xi)e^\xi \quad \text{for all } u.$$

Assume the contrary, namely let u_3 be such that $F_a(u_3)e^{u_3} > F_a(\xi)e^\xi$. By the definition of ξ, $u_3 < u_1$ or $u_3 > u_2$. If $u_3 < u_1$, then $-F_a'(u)F_a(u)^{-1} < 1$ for $u_3 < u < u_1$, and by integration this gives

$$\log\big(F_a(u_3)F_a(u_1)^{-1}\big) < u_1 - u_3$$

or

$$F_a(u_3)e^{u_3} < F_a(u_1)e^{u_1} \leqslant F_a(\xi)e^\xi.$$

The same argument holds for $u_3 > u_2$, proving the claim.

To prove $\xi < (e-1)a$, we note that

$$\int_{\xi-1}^{\xi} F_a(t)\,dt < F_a(\xi)\int_{\xi-1}^{\xi} e^{\xi-t}\,dt = F_a(\xi)(e-1). \tag{2.9}$$

Combining (2.8) and (2.9) completes the proof of Lemma 2.2.

§ 3. The functions $G_a(x)$. We need define one more function which is important in evaluating the lower bound in the sieve.

$$(2.10)\qquad G_a(x) = ax^{-a}\int_x^{\infty} u^{a-1}F_a(u-\tfrac{1}{2})\{F_a(0)-F_a(u-\tfrac{1}{2})\}^{-1}du.$$

$G_a(x)$ is clearly a decreasing function of x. To prove a positive lower bound for the sieve we would like to find ζ_a, the value for which $G_a(\zeta_a) = 1$.

THEOREM 2.1. *If* $a \geqslant 1$, $x \geqslant (e-1)a+\frac{1}{2}+\log((e-1)/(e-2))$, *then* $G_a(x) < 1$.

Proof. By Lemma 2.2 and the fact that $F_a(u)$ is decreasing, we have

$$G_a(x) = ax^{-a}\int_x^{\infty} u^{a-1}F_a(u-\tfrac{1}{2})\{F_a(0)-F_a(u-\tfrac{1}{2})\}^{-1}du$$

$$< ax^{-a}F_a(\xi)e^{\xi+1/2}\{F_a(0)-F_a(x-\tfrac{1}{2})\}^{-1}\int_x^{\infty} u^{a-1}e^{-u}du$$

$$< ae^{\xi+1/2}\{1-e^{\xi+1/2-x}\}^{-1}x^{-a}\int_x^{\infty} u^{a-1}e^{-u}du.$$

We note that

$$\int_x^{\infty} u^{a-1}e^{-u}\,du < x^a e^{-x}(x+1-a)^{-1}\qquad \text{for}\quad x>a.$$

Hence,

$$(2.11)\qquad G_a(x) < a(x+1-a)^{-1}e^{\xi+1/2-x}\{1-e^{\xi+1/2-x}\}^{-1}.$$

By our hypothesis,

$$e^{\xi+1/2-x}(1-e^{\xi+1/2-x})^{-1} \leqslant e-2,$$

$$a(x+1-a)^{-1} \leqslant a((e-2)a+\tfrac{1}{2}) < (e-2)^{-1}.$$

These inequalities coupled with inequality (2.11) completes the proof of Theorem 2.1.

THEOREM 2.2. *If* $a, \beta > 0$, *then*

$$\tau_{a+\beta}(u) = \Gamma(a+\beta)\Gamma(a)^{-1}\Gamma(\beta)^{-1}\int_0^u \tau_a(t)\tau_\beta(u-t)\,dt$$

where $\Gamma(a)$ *is the classical gamma function.*

Proof. By (2.1) and (2.2), we have

$$(2.16)\qquad \int_0^1 t\tau_a'(ut)\tau_\beta(u(1-t))\,dt$$

$$= -u^{-1}\Big\{a\int_0^1 \tau_a(ut-1)\tau_\beta(u(1-t))\,dt-(a-1)\int_0^1 \tau_a(ut)\tau_\beta(u(1-t))\,dt\Big\}$$

$$= -u^{-1}\Big\{a\int_{1/u}^1 \tau_a(ut-1)\tau_\beta(u(1-t))\,dt-(a-1)\int_0^1 \tau_a(ut)\tau_\beta(u(1-t))\,dt\Big\}$$

$$= -(u-1)u^{-2}a\int_0^1 \tau_a((u-1)s)\tau_\beta((u-1)(1-s))\,ds$$

$$+(a-1)u^{-1}\int_0^1 \tau_a(ut)\tau_\beta(u(1-t))\,dt,$$

by letting $ut-1 = (u-1)s$. Hence, if we let

$$g(u) = u\int_0^1 \tau_a(ut)\tau_\beta(u(1-t))\,dt,$$

we find by (2.16) that

$$g(u) = -u^{-1}\{(a+\beta)g(u-1)-(a+\beta-1)g(u)\}$$

or that $g(u)$ satisfies the same difference-differential equation (2.2) as $\tau_{a+\beta}(u)$. Also $g(u) = 0$ if $u \leqslant 0$, and if $0 < u \leqslant 1$,

$$g(u) = u\int_0^1 (ut)^{a-1}(u(1-t))^{\beta-1}dt$$

$$= u^{a+\beta-1}\int_0^1 t^{a-1}(1-t)^{\beta-1}dt$$

$$= \Gamma(a)\Gamma(\beta)\Gamma(a+\beta)^{-1}u^{a+\beta-1}.$$

Thus

$$\tau_{a+\beta}(u) = \Gamma(a+\beta)\Gamma(a)^{-1}\Gamma(\beta)^{-1}g(u)$$

$$= \Gamma(a+\beta)\Gamma(a)^{-1}\Gamma(\beta)^{-1}\int_0^u \tau_a(t)\tau_\beta(1-t)\,dt.$$

In the proof above we assumed the derivatives existed at all positive values, which is not the case when a or $\beta < 1$. The proof can be justified in these cases by suitably splitting the integral and will be left to the reader.

The last theorem is not important in itself but suggests that we might examine the Laplace transform of our function. Let

$$(2.17)\qquad L_a(z) = \int_0^\infty e^{-zt}\tau_a(t)\,dt$$

(i.e., $L_a(z)$ is the Laplace transform of $\tau_a(t)$). We know that $\tau_a(t)$ is of bounded variation and its integral is uniformly convergent. Thus

$$(2.18)\qquad \tau_a(u) = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} e^{zu}L_a(z)\,dz, \quad \sigma > 0,\ u > 0.$$

We can now explicitly evaluate $L_a(z)$ by the difference-differential equation (2.2). Namely,

$$L_a'(z) = -\int_0^\infty e^{-zt}t\tau_a(t)\,dt.$$

Hence,

$$zL_a'(z) = -\int_0^\infty e^{-zt}d\big(t\tau_a(t)\big) = \int_0^\infty e^{-zt}\{a\tau_a(t-1)-a\tau_a(t)\}\,dt$$

$$= \int_0^\infty e^{-zt}\tau_a(t-1)\,dt - aL_a(z) = a(e^{-z}-1)L_a(z),$$

by (2.1). Hence,

$$(2.19)\qquad L_a(z) = C_a\exp\Big\{a\int_0^z (e^{-s}-1)s^{-1}ds\Big\},$$

where

$$C_a = L_a(0) = \int_0^\infty \tau_a(t)\,dt = F_a(0),$$

the constant we wish to evaluate. Substituting (2.19) into (2.18), we have an explicit formula for $\tau_a(u)$,

$$(2.20)\qquad \tau_a(u) = C_a(2\pi i)^{-1}\int_{\sigma-i\infty}^{\sigma+i\infty}\exp\Big\{zu + a\int_0^z (e^{-s}-1)s^{-1}ds\Big\}dz.$$

If γ denotes Euler's constant, then

$$\gamma = -\int_0^1 (e^{-s}-1)s^{-1}ds - \int_1^\infty e^{-s}s^{-1}ds.$$

Thus, as the integrals are analytic,

$$\exp\Big\{a\gamma + a\int_0^z (e^{-s}-1)s^{-1}ds\Big\} = \exp\Big\{a\int_1^z (e^{-s}-1)s^{-1}ds - a\int_1^\infty e^{-s}s^{-1}ds\Big\}$$

$$= \exp\Big\{-a\log z - a\int_z^\infty e^{-s}s^{-1}ds\Big\}$$

$$= z^{-a}\exp\Big\{-a\int_z^\infty e^{-s}s^{-1}ds\Big\} = z^{-a}+K(z),$$

where $K(z)$ is defined by the relation

$$\exp\Big\{-a\int_z^\infty e^{-s}s^{-1}ds\Big\} = 1 - z^aK(z) = 1+O(|z^{-1}e^{-z}|).$$

Hence,

$$K(z) = O(|z^{-1-a}e^{-z}|).$$

Placing these results in (2.20) yields

$$(2.21)\qquad \tau_a(u) = C_ae^{-a\gamma}(2\pi i)^{-1}\int_{\sigma-i\infty}^{\sigma+i\infty} z^{-a}e^{zu}dz + C_ae^{-a}(2\pi i)^{-1}\int_{\sigma-i\infty}^{\sigma+i\infty} e^{zu}K(z)\,dz.$$

If $0 < u \leqslant 1$, we claim

$$\int_{\sigma-i\infty}^{\sigma+i\infty} e^{zu}K(z)\,dz = 0.$$

As, if $\sigma > 1$, and we let C_T be the arc of the circle $|z| = T$ from $\sigma - iT$ to $\sigma + iT$, $\mathrm{Re}(z) > 0$, then

$$\Big|\int_{C_T} e^{zu}K(z)\,dz\Big| \ll T|e^{-z(1-u)}|T^{-a-1} \ll T^{-a}\to 0.$$

This proves the claim as $e^{zu}K(z)$ is analytic for $\mathrm{Re}(z) > 0$. Hence, by (2.2) for $0 < u \leqslant 1$,

$$\tau_a(u) = C_ae^{-a\gamma}(2\pi i)^{-1}\int_{\sigma-i\infty}^{\sigma+i\infty} z^{-a}e^{zu}dz.$$

But

$$(2\pi i)^{-1}\int_{\sigma-i\infty}^{\sigma+i\infty} z^{-a}e^{zu}dz = (2\pi i)^{-1}u^{a-1}\int_{u\sigma-i\infty}^{u\sigma+i\infty} e^{-z}z^{-a}dz = u^{a-1}\Gamma(a)^{-1}.$$

But $\tau_a(u) = u^{a-1}$ for $0 < u \leqslant 1$, giving

THEOREM 2.3. $F_a(0) = \int_0^\infty \tau_a(t)\,dt = \Gamma(a)e^{a\gamma}$.

Utilizing (2.20) we could derive the asymptotic behavior of $\tau_a(u)$ when a is fixed and $u\to\infty$, by the method of steepest descent. One could then generalize the result of de Bruijn for $\tau_1(u)$. For the proof, see de Bruijn [5]. However, we wish a result which is uniform in a.

Let us now return to (2.20). We note that if $|\mathrm{Re}(z)|$ is bounded and $|z|\to\infty$, then $\mathrm{Re}\big(-\int_0^z (e^{-s}-1)s^{-1}ds\big) > c\log|z|$. Hence, by moving the line of integration and changing z into $-z$ we may rewrite (2.20) in the form

$$(2.22)\qquad \tau_a(u) = (2\pi i)^{-1}F_a(0)\int_{\sigma-i\infty}^{\sigma+i\infty}\exp\Big\{-uz+a\int_0^z (e^s-1)s^{-1}ds\Big\}dz$$

for $\sigma\geqslant 0$, $u>0$.

Hence, for $r>-a$, $\sigma>0$,

$$(2.23)\qquad (2\pi i)F_a(0)^{-1}\int_0^\infty \tau_a(u)u^{a+r}du$$

$$=\int_0^\infty u^{a+r}du\int\exp\Big\{-uz+a\int_0^z (e^s-1)s^{-1}ds\Big\}dz$$

$$=\int\exp\Big\{a\int_0^z (e^s-1)s^{-1}ds\Big\}\int_0^\infty u^{a+r}\exp\{-az\}du$$

$$=\Gamma(a+r+1)\int_{\sigma-i\infty}^{\sigma+i\infty}\exp\Big\{a\int_0^z (e^s-1)s^{-1}ds\Big\}z^{-a-r-1}dz.$$

The interchanging of the integrals in the above equation is easily justified. This identity now enables us to prove the following important theorem about the moments of $\tau_a(u)$.

THEOREM 2.4. *If r is fixed,*

$$c_1=\int_0^{\log 2}(e^s-1)s^{-1}ds-\log\log 2,$$

then as $a\to\infty$

$$F_a(0)^{-1}\int_0^\infty \tau_a(u)u^{a+r}du\sim c_2a^{-1/2}\Gamma(a+r+1)\exp\{ac_1\}$$

for some constant c_2.

Proof. Let $g(z)=\int_{\sigma_1}^z (e^s-2)s^{-1}ds$, $\sigma_1=\log 2$. Then $g'(\sigma_1)=0$, $g''(\sigma_1)=2^{-1}\sigma_1>0$. Also

$$\mathrm{Re}\{-g(\sigma_1+it)\}=2\int_0^{|t|}\frac{\sin y+(1-\cos y)y}{\sigma_1^2+y^2}dy.$$

Hence, if $|t|>\delta$,

$$\mathrm{Re}\{-g(\sigma_1+it)\}>c\delta,$$

and if $|t|>2$,

$$\mathrm{Re}\{-g(\sigma_1+it)\}>c\log|t|$$

for some positive constant c.

Therefore, if $\delta=a^{-1/2}$,

$$(2.24)\qquad \int_{\delta<|t|<2}\exp\{ag(\sigma_1+it)\}(\sigma_1+it)^{-r-1}dt=O(\exp\{-a^{1/2}\})$$

and

$$\int_{|t|>2}\exp\{ag(\sigma_1+it)\}(\sigma_1+it)^{-r-1}dt=O(2^{-a}).$$

Finally,

$$(2.25)\qquad \int_{-\delta}^{\delta}\exp\{ag(\sigma_1+it)\}(\sigma_1+it)^{-r-1}dt$$

$$=\sigma_1^{-r-1}\int_{-\delta}^{\delta}\exp\{-ag''(\sigma_1)t^2\}(1+O(at^3)+O(\delta))dt$$

$$=\sigma_1^{-r-1}\int_{-\infty}^{\infty}\exp\{-ag''(\sigma_1)t^2\}dt+O(a^{-1})$$

$$=\sqrt{\pi}\,\sigma_1^{-r}(2\sigma_1a)^{-1/2}+O(a^{-1}).$$

To prove Theorem 2.4, we note that by (2.24) and (2.25)

$$(2.26)\qquad \frac{1}{2\pi i}\int_{\sigma_1-i\infty}^{\sigma_1+i\infty}\exp\Big\{a\int_0^z (e^s-1)s^{-1}ds\Big\}z^{-a-r-1}dz$$

$$=\frac{1}{2\pi}\exp\Big\{a\int_0^{\sigma_1}(e^s-1)s^{-1}ds-a\int_1^{\sigma_1}s^{-1}ds\Big\}\int_{-\infty}^{\infty}\exp\{ag(\sigma_1+it)\}(\sigma_1+it)^{-r-1}dt$$

$$=c_2\sigma_1^{-r}a^{-1/2}\exp\{ac_1\}(1+O(a^{-1/2})),$$

where

$$c_2=\frac{1}{2\pi}(\pi/2\sigma_1)^{1/2},\qquad c_1=\int_0^{\sigma_1}(e^s-1)s^{-1}-\log\log 2.$$

Combining (2.26) with (2.23) yields Theorem 2.4.

COROLLARY. *If* $K = \exp\{c_1 - 1\}$, c_1 *defined in Theorem* 2.4, *then there exist constants* c_3 *and* c_4 *such that*

$$\int_0^\infty \tau_a(u) u^{a-1} du \sim c_3 F_a(0) a^{a-1} K^a,$$

$$\int_0^\infty F_a(u) u^{a-1} du \sim c_4 F_a(0) a^{a-1} K^a.$$

Proof.

$$\int_0^\infty F_a(u) u^{a-1} du = a^{-1} \int_0^\infty \tau_a(u) u^a du.$$

Hence, our corollary immediately follows from Theorem 2.4 by using Sterling's formula.

THEOREM 2.5. [1] $\zeta_a \sim aK$ *as* $a \to \infty$ *where* $K = 1.22\ldots$

Proof. If d is a constant, $d < K$, $x = da$, then

$$G_a(x) = ax^{-a} \int_x^\infty u^{a-1} F_a(u-\tfrac{1}{2}) \{F_a(0) - F_a(u-\tfrac{1}{2})\}^{-1} du$$

$$\geqslant ax^{-a} F_a(0)^{-1} \int_x^\infty u^{a-1} \tau_a(u) du$$

$$> ax^{-a} F_a(0)^{-1} \Big(\int_0^\infty u^{a-1} \tau_a(u) du - x^{a-1} \int_0^x \tau_a(u) du \Big)$$

$$\geqslant a(K/d)^a - a \to \infty, \quad \text{as} \quad a \to \infty.$$

Conversely, let $d > K$, $x = da$. If $u > x$, then

$$F_a(0) - F_a(u-\tfrac{1}{2}) = \int_0^{u-1/2} \tau_a(t) dt \leqslant x^{-a+1} \int_0^\infty \tau_a(t) t^{a-1} dt = o\big(F_a(0)\big) \text{ as } a \to \infty.$$

Hence,

$$G_a(x) = ax^{-a} \int_x^\infty u^{a-1} F_a(u-\tfrac{1}{2}) \big(F_a(0) - F_a(u-\tfrac{1}{2})\big)^{-1} du$$

$$\leqslant ax^{-a} F_a(0)^{-1} \int_x^\infty u^{a-1} F_a(u-\tfrac{1}{2}) du$$

$$\leqslant ax^{-a} F_a(0)^{-1} \int_0^\infty u^{a-1} F_a(u) du \leqslant a(K/d)^a \to 0 \quad \text{as} \quad a \to \infty.$$

The above inequalities complete the proof of our theorem.

[1] Theorem 2.5 was first proved by Dr. H. C. Rumsey by quite a different method.

III

§ 1. We shall now evaluate the sums introduced in Chapter I (e.g. $\sum f(n)^{-1}$ where $n \leqslant x$ and all prime factors of $n \epsilon T$). Let us set $a_n n^{-1} = f(n)^{-1}$.

We restate Assumption (B) in terms of $\{a_p\}$. Thus, to each N let $\{a_p\}$ be a set of non negative real numbers satisfying

$$(3.1) \qquad \sum_{p \leqslant x} a_p < C_1 X (\log X)^{-1}, \quad X \leqslant \log N,$$

$$(3.2) \qquad \sum_{p \leqslant x} (a_p - a) < C_2 X (\log X)^{-2}, \quad X > \log N,$$

for some constants C_1, C_2. There exists a positive δ independent of N such that

$$(3.3) \qquad a_p < (1-\delta) p,$$

$$(3.4) \qquad a_p = a, \quad p > N^\lambda, \ \lambda > 0.$$

We now define a_n multiplicatively from a_p. (3.4) is introduced to artificially define a_p for $p > N^\lambda$. If $p < N^\lambda$ but $p \notin T$, we let $a_p = 0$.

In the following, $K, C_1, C_2, \ldots$ are positive constants independent of N and X. We define

$$(3.5) \qquad \psi(X_1, X_2) = \sum_n a_n n^{-1}, \quad n < X_1, \ p(n) < X_2,$$

where $p(n)$ denotes the maximum prime divisor of n.

THEOREM 3.1. *If*

$$J_a(u) = \int_0^u \tau_a(t) dt,$$

then

$$\psi(X^u, X) = A_a(N) \Gamma(a)^{-1} J_a(u) (\log X)^a + O\big((\log\log N)^K D_a(X) (\log\log X)^{4a}\big),$$

where

$$A_a(N) = \prod_p (1 - a_p p^{-1})^{-1} (1 - p^{-1})^{-a}$$

and

$$D_a(X) = \begin{cases} (\log X)^{a-1} & \text{for} \quad a > 1, \\ \log\log X & \text{for} \quad a = 1, \\ 1 & \text{for} \quad 0 < a < 1. \end{cases}$$

Also

$$(\log\log N)^{-K} \leqslant A_a(N) \leqslant (\log\log N)^K.$$

Proof. The O terms appearing above are independent of u, X, N and K depends only upon C_1 and C_2. For $0 < u \leqslant 1$, then by definition, $\psi(X^u, X) = \psi(X^u, X^u)$, and $J_a(u) = a^{-1}u^a$. Thus, it is sufficient to prove our result for $0 < u \leqslant 1$ only when $u = 1$, by replacing X by X^u. Let $s = \sigma + it$ be a complex variable, $\mathrm{Re}(s) = \sigma > 1$; then

$$\varphi_N(s) = \sum_{n=1}^{\infty} a_n n^{-s} = \prod_p (1-a_p p^{-s})^{-1}. \tag{3.6}$$

We note *via* (3.4) that $a_n < C_3^{\nu(n)}$, so the Dirichlet series in (3.6) converges for $\mathrm{Re}(s) > 1$. Also by (3.4) our infinite product converges in this range, and thus the equality in (3.6).

$\zeta(s)$ denotes the Riemann Zeta function. By (3.3), $1 - a_p p^{-s} \neq 0$ for $\sigma \geqslant 1$, so for $\sigma > 1$,

$$\log \varphi_N(s)\zeta(s)^{-a} = -\sum_p \big((\log(1-a_p p^{-s}) - a\log(1-p^{-s})\big) \tag{3.7}$$

$$= \sum_{e=1}^{\infty}\sum_p (a_p^e - a)e^{-1}p^{-es}.$$

Again by (3.3) for $\sigma \geqslant 1$,

$$\Big|\sum_{e=2}^{\infty}\sum_p (a_p - a)e^{-1}p^{-s}\Big| \leqslant \sum_p a_p^2 p^{-2}(1-p^{-1})^{-1} + ap^{-2}(1-p^{-1}) - 1 \leqslant C_4.$$

(Note: By (3.1) and (3.4), $\sum_{p<x} a_p^2 \leqslant \big(\sum_{p<x} a_p\big)^2 = O\big(X^2(\log X)^{-2}\big)$.)

By partial summation using (3.1) and (3.4), $\sigma \geqslant 1$, then $\sum_p (a_p - a)p^{-\sigma - it}$ converges and

$$\Big|\sum_p (a_p^e - a)p^{-\sigma - it}\Big| = O\big(\log\log\log N + \log(|t|+1)\big). \tag{3.8}$$

Hence, for some constant K

$$(\log\log N)^{-K} \ll A_a(N) \ll (\log\log N)^K. \tag{3.9}$$

It is well known that $\zeta(s)$ is analytic for $s \neq 1$,

$$\zeta(s) = (s-1)^{-1} + O(1), \qquad |s-1| \leqslant 1,$$
$$\log|\zeta(s)| = O\big(\log(|t|+1)\big), \qquad \sigma \geqslant 1,\ |t| \geqslant 1.$$

Hence, by (3.7) and (3.8) for $\sigma \geqslant 1$

$$\varphi_N(s) = A_a(N)(s-1)^{-a} + O\big((\log\log N)^K |s-1|^{1-a}\big), \qquad |s-1| \leqslant 1;$$
$$\varphi_N(s) = O\big((\log\log N)^K |t|^K\big), \qquad |t| \geqslant 1.$$

We now quote the following result, typical of various Tauberian Theorems.

LEMMA 3.1. *If for* $\mathrm{Re}(s) > 1$, $\varphi(s) = \sum_{n=1}^{\infty} b_n n^{-s}$, $b_n \geqslant 0$, $\varphi(s)$ *is analytic*, $\varphi(s) = (s-1)^{-a} + O(\Omega|s-1|^{1-a})$ *for* $|s-1| \leqslant 1$, $|\varphi(s)| = O(\Omega|s|^K)$ *for* $|t| > 1$, $\sigma \geqslant 1$, *then*

$$\sum_{n=1}^{x} b_n n^{-1} = \Gamma(a+1)^{-1}(\log X)^a + O\big(\Omega P_a(X)\big).$$

Proof. The proof only differs slightly from the usual Tauberian proofs associated with the prime number theorem it will not be given (see Chapter 3 of Titchmarsh [18]).

Applying Lemma 3.1 to $\varphi_N(s)$ where $\Omega = (\log\log N)^K$ proves Theorem 3.1 for $u \leqslant 1$.

We now complete the proof for $u > 1$. We shall give the complete proof only for $a > 1$, the case $a \leqslant 1$ varies only in minor detail.

Assume that for a given positive integer r, and for all X_1, X_2, and $u = (\log X_1)/\log X_2$ where $r - 1 < u \leqslant r$, then

$$|\psi(X_1, X_2) - A_a(N)(\log X_2)^a\Gamma(a)^{-1}J_a(u)| < d_r(\log X_1)^{4a}(\log X_2)^{-1-3a}, \tag{3.10}$$

where $d_1 = C_5(\log\log N)^K$, $d_{r+1} = 2d_r$ for small r, and $d_{r+1} = d_r$ for r sufficiently large.

We have seen *via* Lemma 3.1, that (3.10) holds when $r = 1$. We shall proceed by induction on r but first need certain formulae.

Let $X > \log N$, $r - 1 < u \leqslant r$, and q runs over all primes between X and $X^{1+1/u}$. Then by (3.2)

$$\sum_q a_q q^{-1}(\log q)^a J_a\big((\log X^{u+1}/q)/\log q\big) \tag{3.11}$$

$$= a\int_X^{X^{1+1/u}} J_a\big((\log X^{u+1})/(\log t) - 1\big)(\log t)^{a-1}t^{-1}dt + O\big((\log X)^{a-1}\big)$$

$$= a(u+1)^a(\log X)^a\int_u^{u+1} J_a(t-1)t^{-a-1}dt + O\big((\log X)^{a-1}\big),$$

$$\sum_q a_q q^{-1}(\log X^{u+1}/q)^{4a}(\log q)^{-1-3a} \leqslant 2au^{a-1}(\log X)^{a-1}. \tag{3.12}$$

By the definition of $\tau_a(u)$ (see (2.3)),

$$\frac{d}{du}\big(J_a(u)\big) = \tau_a(u) = au^{-1}\int_{u-1}^{u}\tau_a(t)\,dt = au^{-1}\big(J_a(u) - J_a(u-1)\big),$$

hence,

$$\frac{d}{du}\left(u^{-a}J_a(u)\right) = -au^{-a-1}J_a(u-1),$$

or

$$(3.13)\qquad J_a(u+1) = (1+u^{-1})^a J_a(u) - a(u+1)^a \int_u^{u+1} J_a(t)t^{-a-1}dt.$$

We now note

$$(3.14)\qquad \psi(X^{u+1}, X) = \psi(X^{u+1}, X^{1+1/u}) - \sum_n a_n n^{-1};$$

where $n < X^{u+1}$, $X < p(n) < X^{1+1/u}$,

$$= \psi(X^{u+1}, X^{1+1/u} - \sum_q a_q q^{-1} \sum_n a_n n^{-1};$$

where $n < X^{u+1}/q$, $p(n) \leqslant q$

$$= \psi(X^{u+1}, X^{1+1/u}) - \sum_q a_q q^{-1} \psi(X^{u+1}/q, q).$$

The right-hand side of (3.14) falls with the hypothesis of (3.10) (e.g. $(\log X^{n+1}/q)/\log q < u$ as $q > X$); hence, we can apply (3.10). Using (3.11), (3.12), and (3.13) we have by (3.14)

$$|\psi(X^{u+1}, X) - A_a(N)(\log X)^a \Gamma(a)^{-1} J_a(u+1)|$$

$$\leqslant d_r(\log X)^{a-1}\{u^{4a}(1+u^{-1})^{a-1} + 2au^{a-1}+1\}$$

$$< 2d_r(\log X)^{a-1}(u+1)^{4a}$$

and $\leqslant d_r(\log X)^{a-1}(u+1)^{4a}$ for r sufficiently large. This completes the proof of (3.10). To complete the proof of Theorem 3.1 we must achieve a sharper inequality for u large with respect to X.

For all u, $u > u_1 = 4a\log\log X$, then by (2.8)

$$0 < J_a(u) - J_a(u_1) < \int_{u_1}^{\infty} \tau_a(t)\,dt = O\left((\log X)^{-2a}\right),$$

$$0 < \psi(X^u, X) - \psi(X^{u_1}, X) < \prod_{p<x}(1-a_p p^{-1})^{-1} - \psi(X^{u_1}, X)$$

$$< A_a(N)(\log X)^a\left(e^{a\gamma} - \Gamma(a)^{-1}J_a(u_1)\right)$$

$$A_a(N)(\log X)^a\left(e^{a\gamma} - \Gamma(a)J_a(u)\right) + O\left((\log\log N)^K(\log X)^{a-1}\right)$$

$$< O\left((\log\log N)^K(\log X)^{a-1}\right).$$

As $J_a(\infty) = \Gamma(a)e^{a\gamma}$ by Theorem 3.2. Hence

$$|\psi(X^u, X) - A_a(N)(\log X)^a\Gamma(a)^{-1}J_a(u)| = O\left((\log\log N)^K u_1^{4a}(\log X)^{a-1}\right)$$

$$= O\left((\log\log N)^K D_a(x)(\log\log x)^{4a}\right).$$

Theorem 3.1 is only applicable when X is sufficiently large with respect to $\log N$, or else the error term is larger than the main term. We shall now correct this omission.

LEMMA 3.2. *For some* $C_6 > 0$,

$$\left|\varphi(X^u, X) - \prod_{p<x}(1-a_p p^{-1})^{-1}\right| = O\left(e^{-C_6 u}(\log\log N)^K(\log X)^a\right).$$

Proof. Let $H_X(\sigma) = \prod_{p<x}(1-a_p p^{-\sigma})^{-1}$ where $\sigma \geqslant 1-\varepsilon$, $\varepsilon = (\log X)^{-1}\times \times\log(1+\frac{1}{2}\delta)$, the δ is defined by (3.3). Then by (3.3), $1-a_p p^{-\sigma} > \frac{1}{2}\delta$. Also $|\log H_X(\sigma)| = O(\log\log N)$, the proof being identical to that to prove (3.8). Hence, as $a_n \geqslant 0$,

$$\sum_{n>x^u} a_n n^{-1+\varepsilon} < H_X(1-\varepsilon), \qquad p(n) < X,$$

or

$$\sum_{n>x^u} a_n n^{-1} < X^{-\varepsilon u}H_X(1-\varepsilon) = O\left(X^{-\varepsilon u}(\log\log N)^K(\log X)^a\right)$$

which is equivalent to Lemma 3.2.

We recall that we defined $F_a(u) = \int_u^\infty \tau_a(t)\,dt$, so

$$J_a(u) + F_a(u) = \Gamma(a)e^{a\gamma}.$$

LEMMA 3.3. *If* $n \leqslant X^u$, $p(n) \leqslant X$, *then*

$$\prod_{p>x}(1-a_p p^{-1}) - \psi(X^u, X) = A_a(N)\Gamma(a)^{-1}F_a(u)(\log X)^a$$

$$+ O\left((\log\log N)^K D_a(X)\log\log X\right).$$

Proof. By (3.8),

$$(3.15)\qquad \prod_{p<x}(1-a_p p^{-1})^{-1} = A_a(N)\prod_{p<x}(1-p^{-1})^{-a} + O\left((\log\log N)^K(\log X)^{a-1}\right)$$

$$= A_a(N)e^{a\gamma}(\log X)^a + O\left((\log\log N)^K(\log X)^{a-1}\right)$$

by Merten's Theorem on $\prod_{p<x}(1-p^{-1})^{-a}$. Our lemma then follows immediately by Theorem 3.1.

§ 2. We now return to the formulae of Chapter I and state the final results.

T_{N^λ} was a set of primes less than N^λ, $\lambda > 0$; Q was the set of all positive integers all of whose prime factors were in T_{N^λ}. Now

$$\sum_n f(n)^{-1}, \qquad n < N^{\beta/2}, n \in Q,$$

equals

$$\sum_n a_n n^{-1} = \psi(N^{\beta/2}, N^\lambda).$$

(i.e., if $p \notin T_{N^\lambda}$, we have let $a_p = 0$).

Thus, Theorem 1.1 combined with Theorem 3.1 immediately yields Theorem 1, namely if

$$B_a(N) = \Gamma(a)\prod_p \big(1-f(p)^{-1}\big)^{-1}(1-p^{-1})^{-a}, \quad p \in T_{N^\lambda},$$

then

$$M(S_N, T_{N^\lambda}) \leqslant B_a(N) J_a(\tfrac{1}{2}\beta_1\lambda^{-1})^{-1} N(\log N)^{-a} + \lambda^{-a} O\big(N(\log N)^{-a-1/2}\big)$$
$$\leqslant B_a(N) J_a(\tfrac{1}{2}\beta\lambda^{-1})^{-1} N(\log N)\lambda^{-a}\big(1+o(1)\big).$$

The reason we can replace β_1 by β is that $J_a(u)$ is continuous for $u>0$. In the following, we will also write β for β_1.

By Theorem 1.2, for $q \in T_{N^\lambda}$, $n < (N^\beta/q)^{1/2}$, $p(n) \leqslant q$, we have

$$(3.16)\quad M(S_N, T_{N^\lambda}) \geqslant N\Big\{1-\sum_q f(q)^{-1}\Big(\sum_n f(n)^{-1}\Big)^{-1}\Big\} + O\big(N(\log N)^{-a-1}\big)$$
$$\geqslant N\Big\{1-\sum_q a_q q^{-1}\psi\big((N^\beta/q)^{1/2}, q\big)^{-1}\Big\} + O\big(N(\log N)^{-a-1}\big).$$

To evaluate the right-hand side of (3.16) we use the identity; for $q < N^\lambda$

$$(3.17)\quad \prod_q (1-a_q q^{-1}) = 1-\sum_q a_q q^{-1}\prod_{p<q}(1-a_p p^{-1}).$$

If $T = \exp\big((\log N)(\log\log N)^{-1}\big)$, and if we let $X^u = (N^\beta/q)^{1/2}$, $X = q$ in Lemma 3.2, we have

$$(3.18)\quad \sum_{q<T} a_q q^{-1}\Big\{\psi\big((N^\beta/q)^{1/2}, q\big)^{-1} - \prod_{p<q}(1-a_p p^{-1})\Big\}$$
$$= O\Big(\sum_{q<T} a_q q^{-1}(\log q)^a \exp\{-C_6(\log\log N)\}\Big) = O\big((\log N)^{-a}\big).$$

If $v = (\log N^\beta/q)$, then by Theorem 3.1 and Lemma 3.3, for $q > T$,

$$(3.19)\quad \psi\big((N^\beta/q)^{1/2}, q\big)^{-1} - \prod_{p<q}(1-a_p p^{-1})$$
$$= A_a(N)e^{-a\gamma}F_a\big(\tfrac{1}{2}(v-1)\big)J_a\big(\tfrac{1}{2}(v-1)\big)^{-1}(\log q)^{-a} + O\big((\log q)^{-a-1/2}\big).$$

If we multiply (3.19) by $a_q q^{-1}$ and sum over all q, $T < q < N^\lambda$, we have by partial summation, for $S = \frac{1}{2}(\log N^\beta)/(\log T)$,

$$(3.20)\quad \sum_{T<q} a_q q^{-1}\Big\{\psi\big((N^\beta/q)^{1/2}, q\big)^{-1} - \prod_{p<q}(1-a_p p^{-1})\Big\}$$
$$= a\big(A_a(N)e^{a\gamma}\big)^{-1}(\log N^\beta)^{-a}\int_{\beta\lambda^{-1}}^{S} F_a\big(\tfrac{1}{2}(v-1)\big)J_a\big(\tfrac{1}{2}(v-1)\big)^{-1}v^{a-1}\,dv + O\big((\log N)^{-a-1/3}\big)$$
$$= a\big(A_a(N)e^{a\gamma}\big)^{-1}(\tfrac{1}{2}\beta)^{-a}(\log N)^{-a}\int_{\frac{1}{2}\beta\lambda^{-1}}^{\infty} F_a(u-\tfrac{1}{2})J_a(u-\tfrac{1}{2})^{-1}u^{a-1}\,du + o\big((\log N)^{-a}\big)$$
$$= a\big(A_a(N)e^{a\gamma}\big)^{-1}\lambda^{-a}G_a(\tfrac{1}{2}\beta\lambda^{-1})(\log N)^a\big(1+o(1)\big),$$

by the definition (2.11).

By (3.15), for $q < N^\lambda$ we then have

$$\prod_q (1-a_q q^{-1}) = \big(A_a(N)e^{a\gamma}\big)^{-1}\lambda^{-a}(\log N)^{-a}\big(1+o(1)\big).$$

Using the last equation, and placing (3.17), (3.18) and (3.20) into (3.16) we have shown

$$M(S_N, T_{N^\lambda}) \geqslant \big(A_a(N)e^{a\gamma}\big)^{-1}\lambda^{-a}N(\log N)^{-a}\big(1-G_a(\tfrac{1}{2}\beta\lambda^{-1})\big)\big(1+o(1)\big).$$

This completes the proof of Theorem 1 recalling that $B_a(N) = \big(\Gamma(a)A_a(N)\big)^{-1}$.

IV. Applications

§ 1. Let a be a positive integer; $d_1, d_2, \ldots, d_a$ distinct integers which do not form a complete set of residues for any prime;

$$K(\chi) = \prod_{j=1}^{a}(\chi + d_j),$$

$$S_N = \{K(n) \mid n = 1, 2, \ldots, N\}, \qquad T_{N^\lambda} = \{p \mid p = N^\lambda\}.$$

If a_p denotes the number of distinct d_j (mod p), then for $m \in S_N$,

$$\sum_{p\mid m} 1 = N(a_p p^{-1}) + R_p, \qquad |R_p| \leqslant a.$$

Set $f(d)^{-1} = \big(\prod_{p\mid d} a_p p^{-1}\big)$; then

$$\sum_{d\mid m} 1 = Nf(d)^{-1} + R_d, \qquad |R_d| \leqslant a^{\nu(d)}.$$

If $p \nmid \prod_{j\neq k}(d_j-d_k)$, then $f(p)^{-1} = ap^{-1}$. Thus, $\beta = 1$ and

$$M(S_N, T_{N^\lambda}) \leqslant B_a\lambda^{-a}N(\log N)^{-a}J_a(\tfrac{1}{2}\lambda^{-1})^{-1}\big(1+o(1)\big), \tag{4.1}$$

$$M(S_N, T_{N^\lambda}) \geqslant \big(\Gamma(a)e^{a\gamma}\big)^{-1}B_a\lambda^{-a}N(\log N)^{-a}\big(1-G_a(\tfrac{1}{2}\lambda^{-1})\big)\big(1+o(1)\big). \tag{4.2}$$

We note that B_a can be taken to be independent of N, and if $\lambda^{-1}\leqslant 2$, then by the definition of $J_a(u)$,

$$M(S_N, T_{N^\lambda}) \leqslant a2^aB_aN(\log N)^{-a}\big(1+o(1)\big). \tag{4.3}$$

ζ_a was defined by $G_a(\zeta_a) = 1$. Thus, if $\lambda_1^{-1} > 2\zeta_a$, then $M(S_N, T_{N^{\lambda_1}}) > 0$. We have thus shown that there exist infinitely many n for which the least prime factor of $K(n)$ is $> n^{(2\zeta_a)^{-1}}$. Hence, $K(n)$ does not have more than $a2\zeta_a$ prime factors. By Theorem 2.1, $\zeta_a < (1.25)a$ for a sufficiently large.

However, if we are concerned with the problem of finding n for which $K(n)$ has a small number of prime divisors, not how large we can make the least prime divisors, we can strengthen this result considerably. Let $\nu(m)$ denote the number of distinct prime divisors of m.

THEOREM 4.1. *If a is sufficiently large, there exists infinitely many n for which $\nu\big(K(n)\big) < a(\log a+2)$.*

Proof. Let $m\in S_N$ all of whose prime factors are $> N^{1/4\zeta_a}$. We then define "weights" a_p such that if $\nu(m)$ is too large $\sum_{p|m}a_p \geqslant 1$.

If $S_N(p)$ denotes the subset of S_N which are divisible by p, we prove

$$\sum_p a_pM\big(S_N(p), N^{1/4\zeta_a}\big) < M(S_N, N^{1/4\zeta_a}). \tag{4.4}$$

To prove (4.4), define

$$r = \tfrac{1}{2}a(\log 2\zeta_a)-1+\tfrac{1}{2}\zeta_a^{-1}, \qquad a_p = \{\tfrac{1}{2}-(\log p)(\log N)^{-1}\}r^{-1}$$

for the primes between $N^{1/4\zeta_a}$ and $N^{1/2}$. Note that if all prime factors of m are greater than $N^{1/4\zeta_a}$ and $\nu(m)\geqslant 2r+2a$, then

$$\sum_{p|m}a_p \geqslant \big(\tfrac{1}{2}\nu(m)-a\big)r \geqslant 1.$$

The set T consists of all primes, and we let N^λ stand for T_{N^λ}. By the upper bound

$$\begin{aligned}\sum_p a_pM\big(S_N(p), N^{1/4\zeta_a}\big)\{B(4\zeta_a)^aN(\log N)^{-a}\}^{-1} &\leqslant \sum_p a_pa_p^{-1}J_a\big(2\zeta_a\big(1-(\log p)(\log N)^{-1}\big)\big)^{-1}\\ &\leqslant a/r\int_{1/4\zeta_a}^{1/2}(\tfrac{1}{2}-t)t^{-1}J_a\big(2\zeta_a(1-t)\big)^{-1}dt\\ &\leqslant a/r\big(\tfrac{1}{2}(\log 2\zeta_a)-(\tfrac{1}{2}-1/4\zeta_a)\big)J_k(\infty)^{-1}\\ &\leqslant J_k(\infty)^{-1}\big(1+o(1)\big)\end{aligned} \tag{4.5}$$

by the choice of r. In the above we use that $J_k(u) = J_k(\infty)\big(1+o(1)\big)$ for $u>\zeta_a$. Now

$$\begin{aligned}M\big(S(N), N^{1/4\zeta_a}\big)\{B(4\zeta_a)^aN(\log N)^{-a}\}^{-1}\\ \geqslant k(4\zeta_a)^{-a}\int_{\zeta_a}^{4\zeta_a}J_a(u-\tfrac{1}{2})^{-1}u^{a-1}du = J_a(\infty)^{-1}\big(1+o(1)\big).\end{aligned} \tag{4.6}$$

(4.5) coupled with (4.6) and our results about the asymptotic value of ζ_a proves Theorem 4.1.

§ 2. We now consider a special case of the previous example. Let

$$S_N = \{n(n+2)|\ n\leqslant N\},$$

and T the set of all primes.

We readily see that $f(p)^{-1} = 2/p$ for $p>2$ and $=\frac{1}{2}$ for $p=2$. Hence,

$$B_2(N) = 2\prod_{p>2}(1-2/p)(1-1/p)^{-2} = 2\prod_{p>2}\big(1-(p-1)^{-2}\big) \tag{4.7}$$

and

$$\begin{aligned}M(S_N, N^{1/6})\{B_26^{-2}N(\log N)^{-2}\}^{-1} &\geqslant J_2(\infty)^{-1}\{1-G_2(3)\}\big(1+o(1)\big)\\ &\geqslant 2(3)^{-2}\int_{\zeta_2}^{3}J_2(u-\tfrac{1}{2})^{-1}u\,du > 2(3)^{-2}\int_{2.212}^{3}J_2(u-\tfrac{1}{2})^{-1}u\,du > .25.\end{aligned} \tag{4.8}$$

Define

$$a_p = \tfrac{1}{2}-(\log p)(\log N)^{-1}, \qquad N^{1/6}<p<N^{1/2}.$$

Then

$$\begin{aligned}\sum_p a_pM\big(S_N(p), N^{1/6}\big)\{B_26^2N(\log N)^{-2}\}^{-1} &\leqslant 2\sum_p p^{-1}a_pJ_2\big(\tfrac{1}{2}(\log N/p)(\log N^{1/6})^{-1}\big)^{-1}\big(1+o(1)\big)\\ &= \int_{1/6}^{1/2}(1-2u)u^{-1}J_2\big(3(1-u)\big)^{-1}du+o(1),\end{aligned} \tag{4.9}$$

by partial integration,

$$< .23.$$

The last inequalities were derived numerically with the aid of Table 2.

If all the prime factors of m, $m = n(n+2)$, are $> N^{1/6}$, then

$$\sum_{p|m}a_p \geqslant 1 \tag{4.10}$$

if $\nu(n)\geqslant 4$, $\nu(n+2)\geqslant 4$, or $\nu(n) = \nu(n+2) = 3$.

Equations (4.8) and (4.9) combined with (4.10) prove the following theorem:

THEOREM 4.2. *There exist infinitely many n such that $\nu(n) \leqslant 2$ and $\nu(n+2) \leqslant 3$, or $\nu(n) \leqslant 3$ and $\nu(n+2) \leqslant 2$.*

(For results of a similar nature, see Rademacher [11] and Vinogradov [19].)

If we had let $S_N = \{n(N-n) \mid n \leqslant N\}$, for N even, the same method immediately implies; if N is sufficiently large, $N = r+s$ where $\nu(r) \leqslant 3$ and $\nu(s) \leqslant 3$, or $\nu(r) \leqslant 2$ and $\nu(s) \leqslant 3$.

§ 3. Let $\pi(N_1) = N$ (then $\pi(N_1)$ denotes the number of primes $\leqslant N_1$), and let

$$S_N = \{q+2 \mid q \text{ a prime } \leqslant N_1\}.$$

On the E. R. H. (Extended Riemann Hypothesis), if d is odd, then (see Ankeny [1])

$$\sum_q 1 = N\varphi(d)^{-1} + O(N^{1/2}\log d), \quad q < N_1, \quad d \mid q+2.$$

Thus, we may apply the sieve with $a = 1$, $\beta = \frac{1}{2}$, and

$$B_1 = 2\prod_{p>2}\big(1-(p-1)^{-1}\big)(1-p^{-1})^{-1} = B_2.$$

By the definition of $J_1(u)$ and ζ_1,

$$\text{(4.11)} \quad M(S_N, N^{1/6})\{6B_1N(\log N)^{-1}\}^{-1} \geqslant J_1(\infty)^{-1}\big(1-G_1(6/4)\big) = \tfrac{2}{3}\int_{\zeta_1}^{3/2} J_1(u-\tfrac{1}{2})^{-1}du + o(1) = \tfrac{2}{3}\int_{\zeta_1}^{3/2}(u-\tfrac{1}{2})^{-1}du + o(1) > .418.$$

Let $a_p = \frac{1}{2}$ for $N^{1/6} < p < N^{1/3}$; then

$$\text{(4.12)} \quad \sum_p a_p M\big(S_N(p), N^{1/6}\big)\{6B_1N(\log N)^{-1}\}^{-1} \leqslant \tfrac{1}{2}\sum_p p^{-1}J_1\Big(\tfrac{3}{2}\big(1-(\log p)(\log N)^{-1}\big)\Big)^{-1} + o(1) = \tfrac{1}{2}\int_{1/6}^{1/3} J_1\big(\tfrac{3}{2}(1-u)\big)^{-1}u^{-1}du + o(1) < \tfrac{1}{2}\int_{1/6}^{1/3} u^{-1}du < .347.$$

Hence,

$$\sum_p a_p M\big(S_N(p), N^{1/6}\big) < M(S_N, N^{1/6}).$$

If $q+2$ has all of its prime factors $> N^{1/6}$, and if $\nu(q+2) \geqslant 4$, then

$$\sum_{p \mid q+2} a_p \geqslant \tfrac{3}{2} > 1.$$

Thus, we have proved the following theorem:

THEOREM 4.3. *Under the E. R. H., there exist infinitely many primes q such that $q+2$ has at most 3 prime factors.*

We have actually shown there exist primes q for which $q+2$ has at most one prime factor $< N^{1/3}$.

In an almost identical manner we could prove there exist infinitely many primes q such that $\nu\big(\frac{1}{2}(q-1)\big) \leqslant 3$ under the E. R. H.

Also, without any hypothesis, we could prove there exist infinitely many n such that $\nu(n^2+1) \leqslant 3$.

§ 4. Let $2 = p_1 < p_2 < \ldots$ be the set of primes. What can we prove about the differences $p_{j+1}-p_j$? If the twin prime theorem were true, $p_{j+1}-p_j$ would equal 2 infinitely often; but what can be proved? Let

$$c_1 = \liminf_j\, (p_{j+1}-p_j)(\log p_j)^{-1}.$$

Erdös proved $c_1 < 1$ and Rankin sharpened this to $c_1 \leqslant 1$. (See Erdös [6], and Rankin [15].)

We shall now prove,

THEOREM 4.4. $c_1 \leqslant 15/16$.

Proof. Let N be large, $q_1 < q_2 < \ldots < q_s = N$ be the primes between $N(\log N)^{-1}$ and N. So, $\zeta = N(\log N)^{-1} + O\big(N(\log N)^{-2}\big)$. Denote by $H(d)$ the number of j such that $q_{j+1}-q_j = d$. Then

$$\text{(4.11)} \quad \sum_d H(d) = S, \quad \sum_d dH(d) = N + O\big(N(\log N)^{-1}\big).$$

On the other hand, $H(d)$ is less than the number of $n \leqslant N$ such that n and $n+d$ are both primes. Hence, by (4.1),

$$\text{(4.12)} \quad H(d) < 8B_2N(\log N)^{-2}\psi(d)\big(1+o(1)\big)$$

where

$$\psi(d) = \prod_p (1-p^{-1})(1-2p^{-1})^{-1}, \quad p \mid d, \quad p > 2.$$

LEMMA 4.1. *If d runs over all even numbers $< X$,*

$$\sum_d \psi(d) = B_2^{-1}X + O(\sqrt{X}), \quad \sum_d d\psi(d) = \tfrac{1}{2}B_2^{-1}X^2 + O(X^{3/2}).$$

Proof. Let t be odd and square free. Define $\omega(t)$ multiplicatively by $\omega(p) = (p-2)^{-1}$. Then $\psi(d) = \sum_{t \mid d}\omega(t)$. Also $\omega(t) = O(t)^{-1/2}$. Hence,

$$\sum_{d<x}\psi(d) = \sum_{t<x}\omega(t)\sum_d 1, \quad t \mid d,\ d < x,$$
$$= \tfrac{1}{2}\sum_{t<\sqrt{x}}\omega(t)t^{-1} + O(\sqrt{X}) = \tfrac{1}{2}\sum_t \omega(t)t^{-1} + O(\sqrt{X})$$
$$= \tfrac{1}{2}\prod_p\big(1+(p-2)^{-1}p^{-1}\big) + O(\sqrt{X}),$$

which proves the first equality of our lemma. The second equality follows in the same manner.

Let c be a positive constant. Then

$$8B_2 \sum_d{}_1 \psi(d) N(\log N)^{-2} = N(\log N)^{-1} + O\big(N(\log N)^{-2}\big),$$

$$c\log N < d < (c+\tfrac{1}{8})\log N$$

and

$$(4.13)\quad 8B_2 \sum_d{}_1 \psi(d)\, d N(\log N)^{-2} = 4\big((c+\tfrac{1}{8})^2 - c^2\big)N + O\big(N(\log N)^{-1}\big).$$

Assume $H(d) = 0$ for $d \leqslant c\log N$. Then, as

$$H(d) < 8B_2\psi(d)N(\log N)^{-2}\big(1+o(1)\big),$$

$$\sum_d H(d)d = N + O\big(N(\log N)^{-1}\big)$$

$$\geqslant 8B_2\Big(\sum_d{}_1 \psi(d)\,d\Big)\big(N(\log N)^{-2}\big)\big(1+o(1)\big)$$

$$\geqslant 4\big((c+\tfrac{1}{8})^2 - c^2\big)N\big(1+o(1)\big).$$

Hence, $c \leqslant 15/16$, thus proving Lemma 4.1.

Using more complicated methods we can sharpen the bound slightly. Under the E. R. H., we can improve our result by 3 to $c_1 \leqslant 7/8$. However, if we combine the sieve method with the "circle" method we could prove $c_1 \leqslant \frac{1}{2}$. (See Rankin [15].)

TABLE 1*

a	x	$G_a(x)$
1	1.032	1.0027117
1	1.034	.99624306
1.5	1.612	1.002664300
1.5	1.614	.997648370
2.0	2.210	1.0018957
2.0	2.212	.99767852
2.5	2.816	1.003276300
2.5	2.8180	.999535960
3.0	3.4280	1.003249000
3.0	3.4300	.999838950
3.5	4.044	1.002180100
3.5	4.046	.999014950

* The authors would like to thank Dr. J. Muscat for the computations in Table 1.

TABLE 2

u	$J_2(u)$	u	$J_2(u)$
1.	.5000	1.8	1.4755
1.05	.5512	1.85	1.5395
1.1	.6047	1.9	1.6029
1.15	.6602	1.95	1.6656
1.2	.7175	2.0	1.7274
1.25	.7763	2.05	1.7737
1.3	.8366	2.1	1.8199
1.35	.8981	2.15	1.8768
1.4	.9605	2.2	1.9362
1.45	1.0238	2.25	1.9868
1.5	1.0877	2.3	2.0398
1.55	1.1521	2.35	2.0911
1.6	1.2168	2.4	2.1412
1.65	1.2816	2.45	2.1894
1.7	1.3465	2.5	2.2374
1.75	1.4112		

If $x > \zeta_a$, we note that

$$1 - G_a(x) = 1 - ax^{-a}\int_x^\infty F_a(u-\tfrac{1}{2})J_a(u-\tfrac{1}{2})^{-1}u^{a-1}\,du$$

$$= 1 - (\zeta_a x^{-1})^a G_a(\zeta_a) + ax^{-a}\int_a^x F_a(u-\tfrac{1}{2})J_a(u-\tfrac{1}{2})^{-1}u^{a-1}\,du$$

$$= a\Gamma(a)e^{a\gamma}x^{-a}\int_{\zeta_a}^x J_a(u-\tfrac{1}{2})^{-1}u^{a-1}\,du.$$

Also, by definition,

$$J_1(u) = \begin{cases} u, & 0 \leqslant u \leqslant 1, \\ 2u - 1 - u\log u, & 1 \leqslant u \leqslant 2. \end{cases}$$

$$J_2(u) = \begin{cases} \tfrac{1}{2}u^2, & 0 \leqslant u \leqslant 1, \\ 2u^2 - 2u + \tfrac{1}{2} - u^2\log u, & 1 \leqslant u \leqslant 2. \end{cases}$$

References

[1] N. C. Ankeny, *The least quadratic non-residue*, Annals of Math. 55 (1952), pp. 65-72.

[2] A. Buchstab, *Neue Verbesserungen in der Methode des Eratosthenischen Siebes*, Rec. Math. N. S. 4 (1938), pp. 375-387.

[3] — *On those numbers in an arithmetic progression all prime factors of which are small in magnitude*, Doklady Akad. Nauk. SSSR (N. S.) 67 (1949), pp. 5-9.

[4] N. G. de Bruijn, *On the numbers of uncancelled elements in the sieve of Eratosphenes*, Nederl. Akad. Wetensch. Proc. 52 (1950), pp. 803-812.

[5] — *The asymptotic behavior of a function occuring in the theory of primes*, Ind. Math. Soc., 15, pp. 25-32.

[6] P. Erdös, *The difference of consecutive primes*, Duke Math. J. 6 (1940), pp. 438-441.

[7] — *On some applications of Brun's method*, Acta. Sci. Math. Szeged. 13 (1949), pp. 57-63.

[8] — *Problems and results on the difference of consecutive primes*, Publicationes Mathematicae 1 (1949), pp. 33-37.

[9] T. Estermann, *Eine neue Darstellung und neue Anwendungen der Viggo Brunschen Methode*, J. Reine Angew. Math. 168 (1932), p. 106.

[10] R. D. James, *Recent progress in the Goldbach problem*, Bull. Amer. Math. Soc. 49 (1943), p. 422.

[11] H. Rademacher, *Beiträge zur Viggo-Brunschen Methode in der Zahlentheorie*, Hamburg Abh. 3 (1924), p. 12.

[12] V. Ramaswami, *On the number of positive integers $\leqslant x$ and free of prime divisors $> x^c$*, Bull. Amer. Math. Soc. 55 (1949), p. 1122.

[13] — *On the number of positive integers $\leqslant x$ and free of prime factors $> x^c$ and a problem of S. S. Pillai*, Duke Math. J. 16 (1949), pp. 99-109.

[14] R. A. Rankin, *The differences between consecutive prime numbers III*, J. London Math. Soc. 22 (1947), pp. 226-230.

[15] — *The differences between consecutive prime numbers IV*, Proc. of the Cambridge Phil. Soc. 36 (1940), pp. 255-266.

[16] L. Schnirelmann, *Über additiven Eigenschaften von Zahlen*, Math. Ann. 107 (1933), pp. 649-690.

[17] Atle Selberg, *The general sieve method and its place in prime number theory*, Proc. Int. Cong. Math. 1 (1950), pp. 286-293.

[18] E. C. Titchmarsh, *The theory of the Riemann zeta-function*, Oxford Univ. Press, 1951.

[19] A. I. Vinogradov, *Applications of $\zeta(s)$ to the sieve of Eratostenes*, Mat. Sb. N. S. 41 (83) (1957), pp. 49-80, correction pp. 415-416.

[20] Y. Wang, *On the representation of a large even integer as a sum of a prime and a product of at most 4 primes*, Acta Math. Sinica 6 (1956), pp. 565-582.

MATHEMATICS DEPARTMENT
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Reçu par la Rédaction le 31. 5. 1963

ACTA ARITHMETICA
X (1964)

On a conjecture of Erdös in additive number theory

by

R. L. Graham (Murray Hill, N. J.)

1. Introduction. Let t and a be real numbers and let $S_t(a)$ denote the sequence $(s_1, s_2, \ldots)$ defined by $s_n = [ta^n]$ (where [ ] denotes the greatest integer function). It was conjectured by Erdös several years ago that if $t > 0$ and $1 < a < 2$ then every sufficiently large integer n can be expressed as $n = \sum_{k=1}^{\infty} \varepsilon_k s_k$ where $\varepsilon_k = 0$ or 1 and all but a finite number of the ε_k are 0. In general, a sequence of integers which has this property is said to be *complete* and if every positive integer is so expressible then the sequence is said to be *entirely complete*. While the additive structure of $S_t(a)$ is far from being completely understood at present, it is the object of this paper to shed some light on this question. In particular, the set T of all points (t, a) of the unit square $S = \{(t, a)\colon 0 < t < 1, 1 < a < 2\}$ for which $S_t(a)$ is complete will be determined. It will be seen T has an area of approximately 0.85.

2. Preliminary remarks. If $A = (a_1, a_2, \ldots)$ is a sequence of integers then $P(A)$ is defined to be the set of all sums of the form $\sum_{k=1}^{\infty} \varepsilon_k a_k$ where $\varepsilon_k = 0$ or 1 and all but a finite number of the ε_k are 0. In this paper, we adopt the convention that a sum of the form $\sum_{k=a}^{b}$ is 0 for $b < a$. We now give several results which will be needed later.

THEOREM 1. (J. Folkman.) *Let $A = (a_1, a_2, \ldots)$ be a sequence of positive integers such that:*

1. $a_n + a_{n+1} \leqslant a_{n+2}$ *for* $n \geqslant 1$.
2. *There exist $m \geqslant 0$ and $r \geqslant 0$ such that $m \notin P(A)$ and*

$$\sum_{k=1}^{r} a_k < m < a_{r+2}.$$

Then A is not complete.