

On the difference between consecutive numbers prime to n : II

By CHRISTOPHER HOOLEY (Durham)

1. Introduction

We return to the topic of the distribution of the numbers $a_1, \dots, a_{\varphi(n)}$, which as in the previous paper¹⁾ of the same title are defined to be (in ascending order) the $\varphi(n)$ numbers not exceeding n that are prime to n . Previously¹⁾ it was shewn that the intervals

$$\Delta_i = a_{i+1} - a_i$$

satisfied the inequality

$$(A) \quad \sum_{i=1}^{\varphi(n)-1} \Delta_i^\alpha = O \left\{ n \left(\frac{n}{\varphi(n)} \right)^{\alpha-1} \right\}$$

for²⁾ $1 \leq \alpha < 2$. Here we shall study the intervals in a different aspect by considering how for each n the values of Δ_i are distributed statistically. Our main conclusion, which is embodied in Theorem 1, is that the ratio

$$\frac{\Delta_i}{n/\varphi(n)}$$

is distributed approximately as a gamma variable with parameter 1, provided n is to be regarded as tending to infinity through a sequence for which $n/\varphi(n)$ also tends to infinity. The distribution of $\Delta_i \varphi(n)/n$ for each n is thus essentially independent of n when $n/\varphi(n) \rightarrow \infty$, as had been conjectured earlier by P. ERDŐS³⁾ for the special case where n was a product $2 \cdot 3 \cdot \dots \cdot p$ of consecutive primes.

A consequence of Theorem 1 is that the inequality (A) can be replaced by an asymptotic formula when $n/\varphi(n) \rightarrow \infty$. This formula is stated without proof in Theorem 2.

¹⁾ On the difference of consecutive numbers prime to n , *Acta Arith.* **8**, (1963), 295–299.

²⁾ This inequality can be seen to be also valid for $0 \leq \alpha < 1$ by the use of Hölder's inequality.

³⁾ P. ERDŐS, Some unsolved problems, *Magyar Tud. Akad. Kutató Int. Közl.* **6** (1961), 221–254.

2. Notation

The following notation will be adopted throughout.

The letters $d, d_1, l_i, \lambda_i, r, s, t$ are positive integers, where r satisfies the condition $2 \leq r \leq s$; b_i, B_i, m are non-negative integers; p is a (positive) prime number.

The integer n is to be regarded as tending to infinity through a sequence of values for which $n/q(n)$ also tends to infinity, all (appropriate) inequalities that are valid for sufficiently large $n/q(n)$ being assumed to hold. The variable c belongs to a range $\mathfrak{R}$ of values that is bounded at each end by positive constants.

The constants implied by the O notation need depend at most on $\mathfrak{R}$ and s (initial dependence on r being replaced by that on s in view of the inequality $2 \leq r \leq s$). The notation $k = o(|h|)$ denotes a relation of the form

$$k = \varepsilon(n) |h|,$$

where $\varepsilon(n) \rightarrow 0$ as $n \rightarrow \infty$ through the appropriate sequence, the passage to the limit being

(i) in Section 4 uniform with respect to $l_1, l_2, \dots, l_{r-1}$ provided $0 < l_1 < l_2 < \dots < l_{r-1} < y$

(ii) before equation (23) not necessarily uniform with respect to $\mathfrak{R}$ and s (but being uniform with respect to c in a given $\mathfrak{R}$)

(iii) in equation (23) not necessarily uniform with respect to $\mathfrak{R}$ (s being no longer present)

(iv) in Theorem 2 not necessarily uniform with respect to α .

The letters A, B indicate functions (of the appropriate variables) that are bounded by absolute constants; D is an absolute constant.

3. Formula for $f_n(c)$

The ultimate aim is to determine an asymptotic formula for $f_n(c)$, which we define to be the number of differences

$$A_i = a_{i+1} - a_i, \quad 0 < i < q(n),$$

for which

$$A_i < y = \frac{cn}{q(n)}.$$

We begin by proving a preliminary formula for $f_n(c)$.

It is convenient to extend the definition of the numbers a_i by defining, generally, for any positive integer i (whether less than $q(n)$ or otherwise), a_i to be the i^{th} positive integer prime to n . We then define $g_n(c)$ like $f_n(c)$, except that the condition $0 < i < q(n)$ is to be replaced by $0 < i \leq q(n)$. Since in fact $a_{q(n)+1} = n+1$ and $y > 2$, we see that

$$(1) \quad f_n(c) = g_n(c) - 1.$$

To obtain an expression for $g_n(c)$ we define

$$N_r(n, y) = N_r(y) = N_r$$

to be the number of sets of r numbers $a_{i_1}, a_{i_2}, \dots, a_{i_r}$ satisfying the conditions

$$a_{i_1} < a_{i_2} < \dots < a_{i_r}; \quad a_{i_r} - a_{i_1} < y; \quad 0 < i_1 \leq \varphi(n).$$

Then, by an application of a well known exclusion principle, as used for example in Brun's sieve method, we have

$$g_n(c) \equiv N_2 - N_3 + \dots + (-1)^t N_t,$$

if t be even, and

$$g_n(c) \equiv N_2 - N_3 + \dots + (-1)^t N_t,$$

if t be odd, from which we deduce by using (1)

$$(2) \quad f_n(c) = N_2 - N_3 + \dots + (-1)^{s-1} N_{s-1} + AN_s - 1.$$

4. Estimation of N_r ; first stage

For any set of integers $l_1, l_2, \dots, l_{r-1}$ such that $0 < l_1 < l_2 < \dots < l_{r-1}$ we denote by $F(l_1, l_2, \dots, l_{r-1})$ the number of values of m in the range $0 \leq m < n$ for which the numbers $m, m+l_1, \dots, m+l_{r-1}$ are all prime to n . Then

$$(3) \quad N_r = \sum_{0 < l_1 < \dots < l_{r-1} < y} F(l_1, \dots, l_{r-1}).$$

Next we evaluate $F(l_1, \dots, l_{r-1})$ by the sieve of Eratosthenes. Since the set $m, m+l_1, \dots, m+l_{r-1}$ contributes to $F(l_1, \dots, l_{r-1})$ if and only if

$$(m\{m+l_1\}\dots\{m+l_{r-1}\}, n) = 1,$$

we have

$$\begin{aligned} F(l_1, \dots, l_{r-1}) &= \sum_{0 \leq m < n} \sum_{\substack{d | m(m+l_1)\dots(m+l_{r-1}) \\ d | n}} \mu(d) = \\ &= \sum_{d | n} \mu(d) \sum_{\substack{m(m+l_1)\dots(m+l_{r-1}) \equiv 0 \pmod{d} \\ 0 \leq m < n}} 1 = n \sum_{d | n} \frac{\mu(d)}{d} v(d; l_1, \dots, l_{r-1}), \end{aligned}$$

where, for any set⁴⁾ of integers $b_1, \dots, b_{r-1}$, $v(d; b_1, \dots, b_{r-1})$ is the number of roots of the congruence

$$m(m+b_1)\dots(m+b_{r-1}) \equiv 0 \pmod{d}.$$

Since $v(d; b_1, \dots, b_{r-1})$ is a multiplicative function of d , that is to say, if $(d', d'') = 1$, then $v(d'; b_1, \dots, b_{r-1})v(d''; b_1, \dots, b_{r-1}) = v(d'd''; b_1, \dots, b_{r-1})$, we conclude that

$$(4) \quad F(l_1, \dots, l_{r-1}) = n \prod_{p | n} \left(1 - \frac{v(p; l_1, \dots, l_{r-1})}{p} \right) = n \Pi_1,$$

say.

It should be remarked here that for any prime number p $v(p; b_1, \dots, b_{r-1})$ admits of the alternative interpretation as the number of distinct residues $(\bmod p)$ in the system $0, b_1, \dots, b_{r-1}$.

⁴⁾ The definition of $v(d; b_1, \dots, b_{r-1})$ will be required later for the case in which the restrictions $b_i \neq b_j$ and $b_i \neq 0$ do not apply.

The formula obtained above for $F(l_1, \dots, l_{r-1})$ is not suitable in its present form for the estimation of N_r . It is first necessary to derive from it a modified and approximate formula by writing the product Π_1 as

$$(5) \quad \Pi_1 = \prod_{p \equiv Y} \prod_{p > Y} = \Pi_2 \Pi_3,$$

say where⁵⁾

$$Y = \frac{\log y}{(\log \log y)^{\frac{1}{2}}}.$$

We consider Π_3 by writing it as

$$(6) \quad \Pi_3 = \prod_{(A)} \prod_{(B)} = \Pi_4 \Pi_5,$$

say, where (A) indicates that the product is over the (appropriate⁶⁾) primes p for which $v(p; l_1, \dots, l_{r-1}) < r$, and (B) indicates that the product is over the (appropriate) primes p for which $v(p; l_1, \dots, l_{r-1}) = r$.

To investigate Π_4 let q be the number of primes p for which $p > Y$ and $v(p; l_1, \dots, l_{r-1}) < r$ but not subject to any other restrictions. Then since any such prime divides the non-vanishing product

$$l_1 l_2 \dots l_{r-1} \prod_{i < j \leq r-1} (l_j - l_i),$$

we have

$$Y^q < l_1 l_2 \dots l_{r-1} \prod_{i < j \leq r-1} (l_j - l_i) < y^{r^2},$$

and therefore

$$q = O\left(\frac{\log y}{\log Y}\right).$$

Next, by this and the definition of (A), we have⁷⁾

$$\begin{aligned} 1 &\leq \frac{\Pi_4}{\prod_{(A)} \left(1 - \frac{r}{p}\right)} \leq \prod_{(A)} \left(1 - \frac{r}{p}\right)^{-1} \leq \left(1 - \frac{r}{Y}\right)^{-q} = \\ &= 1 + O\left(\frac{\log y}{Y \log Y}\right) = 1 + O\left(\frac{1}{(\log \log y)^{\frac{1}{2}}}\right) = 1 + o(1). \end{aligned}$$

Therefore, by this and (6), we have

$$(7) \quad \Pi_3 = \{1 + o(1)\} \prod_{(A)} \left(1 - \frac{r}{p}\right) \prod_{(B)} \left(1 - \frac{r}{p}\right) = \{1 + o(1)\} \prod_{\substack{p|n \\ p > Y}} \left(1 - \frac{r}{p}\right).$$

⁵⁾ The precise value of Y being not important here, we choose one that is convenient.

⁶⁾ i. e. those primes p for which $p|n$, $p > Y$, and $v(p; l_1, \dots, l_{r-1}) < r$.

⁷⁾ The relationship between the orders of magnitude of y , Y , and q makes the subsequent estimation of $\left(1 - \frac{r}{Y}\right)^{-q}$ valid.

Finally, since

$$\left(1 - \frac{r}{p}\right) \left(1 - \frac{1}{p}\right)^{-r} = 1 + O\left(\frac{1}{p^2}\right)$$

for $p > Y$, we have

$$\begin{aligned} \prod_{\substack{p|n \\ p > Y}} \left(1 - \frac{r}{p}\right) &= \prod_{\substack{p|n \\ p > Y}} \left(1 - \frac{1}{p}\right)^r \prod_{\substack{p|n \\ p > Y}} \left\{1 + O\left(\frac{1}{p^2}\right)\right\} = \left\{1 + O\left(\frac{1}{Y}\right)\right\} \left\{ \prod_{\substack{p|n \\ p > Y}} \left(1 - \frac{1}{p}\right) \right\}^r = \\ &= \left\{1 + O\left(\frac{1}{Y}\right)\right\} \psi_Y^r(n), \end{aligned}$$

say, and therefore from this and (7),

$$(8) \quad \Pi_3 = \{1 + o(1)\} \psi_Y^r(n).$$

The combination of (5) and (8) gives the required formula for $F(l_1, \dots, l_{r-1})$. Writing now

$$\Pi_2 = G(l_1, \dots, l_{r-1})$$

to indicate the dependence of Π_2 on $l_1, \dots, l_{r-1}$, we have, by (3), (4), (5), (8), and note (i) in Section 2,

$$\begin{aligned} (9) \quad N_r &= n \{1 + o(1)\} \psi_Y^r(n) \sum_{0 < l_1 < \dots < l_{r-1} < y} G(l_1, \dots, l_{r-1}) \\ &= n \{1 + o(1)\} \psi_Y^r(n) \Sigma_{r,y}, \text{ say.} \end{aligned}$$

In the next section we proceed to the estimation of $\Sigma_{r,y}$.

5. Transformation of $\Sigma_{r,y}$

In order to examine $\Sigma_{r,y}$ let d_1 indicate, generally, either 1 or square-free numbers composed entirely of prime factors p such that $p \leq Y$. A bound for numbers of this type will be needed, and is easily obtained. In fact

$$\begin{aligned} d_1 &\leq (Y)^{\pi(Y)} < Y^{\left(\frac{2Y}{\log Y}\right)} = e^{2Y} \\ &= e^{\frac{2 \log y}{(\log \log y)^{\frac{1}{2}}}} = y^{\frac{2}{(\log \log y)^{\frac{1}{2}}}}. \end{aligned}$$

Hence

$$(10) \quad d_1 < Z, \quad \text{where} \quad Z = y^{\frac{2}{(\log \log y)^{\frac{1}{2}}}} < y.$$

Now

$$G(l_1, \dots, l_{r-1}) = \sum_{d_1|n} \frac{\mu(d_1)}{d_1} v(d_1; l_1, \dots, l_{r-1}).$$

Therefore

$$(11) \quad \Sigma_{r,y} = \sum_{0 < l_1 < \dots < l_{r-1} < y} \sum_{d_1 | n} \frac{\mu(d_1)}{d_1} v(d_1; l_1, \dots, l_{r-1}) = \sum_{d_1 | n} \frac{\mu(d_1)}{d_1} S(d_1, y),$$

where

$$S(d_1, y) = \sum_{0 \leq b_1, \dots, b_{r-1} < d_1} v(d_1; b_1, \dots, b_{r-1}) \sum_{\substack{0 < \lambda_1 < \dots < \lambda_{r-1} < y \\ \lambda_i \equiv b_i \pmod{d_1}}} 1,$$

the variables of summation in the inner sum being denoted by λ_i instead of l_i , since we no longer wish them in the sequel to be necessarily restricted by the condition imposed at the beginning of Section 4.

We estimate $S(d_1, y)$ by adopting an artifice which may seem unnatural, but which obviates a complicated lattice point calculation. We consider the contribution $S'(d_1, y)$, say, to $S(d_1, y)$ due to all sets $b_1, \dots, b_{r-1}$ in the outer sum which correspond to a fixed selection of $r-1$ numbers $\beta_1, \dots, \beta_{r-1}$ (not necessarily all distinct). Letting Σ' indicate summation over all arrangements $b_1, \dots, b_{r-1}$ of $\beta_1, \dots, \beta_{r-1}$, we have that this contribution is given by

$$(12) \quad v(d_1; \beta_1, \dots, \beta_{r-1}) \sum'_{b_1, \dots, b_{r-1}} \sum_{\substack{0 < \lambda_1 < \dots < \lambda_{r-1} < y \\ \lambda_i \equiv b_i \pmod{d_1}}} 1,$$

and then from the form of this expression we see that the condition $0 < \lambda_1 < \dots < \lambda_{r-1} < y$ in the inner sum may be replaced by $0 < \lambda_{j_1} < \dots < \lambda_{j_{r-1}} < y$, where $j_1, \dots, j_{r-1}$ is a permutation of $1, \dots, r-1$. Therefore (12) may be replaced by

$$(13) \quad \frac{1}{(r-1)!} v(d_1; \beta_1, \dots, \beta_{r-1}) \sum'_{b_1, \dots, b_{r-1}} \sum_{\substack{\lambda_i \equiv b_i \pmod{d_1} \\ 0 < \lambda_i < y \\ \lambda_i \neq \lambda_j \text{ if } i \neq j}} 1.$$

A simple argument shews that the inner sum in (13) is

$$\left(\frac{y}{d_1} + O(1) \right)^{r-1},$$

which is

$$\frac{y^{r-1}}{d_1^{r-1}} + O\left(\frac{y^{r-2}}{d_1^{r-2}}\right)$$

by (10); from this we have that (13) is

$$\left\{ \frac{1}{(r-1)!} \frac{y^{r-1}}{d_1^{r-1}} + O\left(\frac{y^{r-2}}{d_1^{r-2}}\right) \right\} v(d_1; \beta_1, \dots, \beta_{r-1}) \sum'_{b_1, \dots, b_{r-1}} 1.$$

Therefore

$$(14) \quad S(d_1, y) = \left\{ \frac{1}{(r-1)!} \frac{y^{r-1}}{d_1^{r-1}} + O\left(\frac{y^{r-2}}{d_1^{r-2}}\right) \right\} \sum_{0 \leq b_1, \dots, b_{r-1} < d_1} v(d_1; b_1, \dots, b_{r-1}) = \\ = \left\{ \frac{1}{(r-1)!} \frac{y^{r-1}}{d_1^{r-1}} + O\left(\frac{y^{r-2}}{d_1^{r-2}}\right) \right\} M_r(d_1),$$

say. Returning to $\Sigma_{r,y}$ we have, by (11) and (14),

$$\begin{aligned}
 (15) \quad \Sigma_{r,y} &= \sum_{d_1|n} \left\{ \frac{1}{(r-1)!} \frac{y^{r-1}}{d_1^{r-1}} + O\left(\frac{y^{r-2}}{d_1^{r-2}}\right) \right\} \frac{\mu(d_1)}{d_1} M_r(d_1) = \\
 &= \frac{y^{r-1}}{(r-1)!} \sum_{d_1|n} \frac{\mu(d_1) M_r(d_1)}{d_1^r} + O\left(y^{r-2} \sum_{d_1|n} \frac{|\mu(d_1)| M_r(d_1)}{d_1^{r-1}}\right) = \\
 &= \frac{y^{r-1}}{(r-1)!} \Sigma_1 + O(y^{r-2} \Sigma_2),
 \end{aligned}$$

say.

6. The multiplicativity of $M_r(d)$

The next step in the estimation of $\Sigma_{r,y}$ depends on the property that $M_r(d)$ is a multiplicative function of d .

If $(d', d'') = 1$, then from the definition of $M_r(d)$

$$(16) \quad M_r(d') M_r(d'') = \sum_{\substack{0 \leq b'_1, \dots, b'_{r-1} < d' \\ 0 \leq b''_1, \dots, b''_{r-1} < d''}} v(d'; b'_1, \dots, b'_{r-1}) v(d''; b''_1, \dots, b''_{r-1}).$$

For each set $b'_1, \dots, b'_{r-1}, b''_1, \dots, b''_{r-1}$ occurring in the above sum, define $B_1, \dots, B_{r-1}$ by the conditions

$$B_i = \begin{cases} b'_i \pmod{d'} \\ b''_i \pmod{d''} \end{cases} \quad ; \quad 0 \leq B_i < d' d''.$$

The summand in (16) becomes

$$\begin{aligned}
 v(d'; B_1, \dots, B_{r-1}) v(d''; B_1, \dots, B_{r-1}) &= \\
 &= v(d' d''; B_1, \dots, B_{r-1}),
 \end{aligned}$$

since $v(d; B_1, \dots, B_{r-1})$ is a multiplicative function of d ; while, as $b'_1, \dots, b'_{r-1}, b''_1, \dots, b''_{r-1}$ vary, each set of $B_1, \dots, B_{r-1}$ satisfying

$$0 \leq B_1, \dots, B_{r-1} < d' d''$$

is obtained exactly once. Therefore

$$M_r(d') M_r(d'') = \sum_{0 \leq B_1, \dots, B_{r-1} < d' d''} v(d' d''; B_1, \dots, B_{r-1}) = M_r(d' d''),$$

as asserted at the beginning of the paragraph.

7. Estimation of Σ_2

For the assessment of Σ_2 we have

$$\Sigma_2 \leq Z \sum_{d_1|n} \frac{|\mu(d_1)| M_r(d_1)}{d_1^r},$$

by (10), and then, by the multiplicativity of $M_r(d)$,

$$\Sigma_2 \cong Z \prod_{\substack{p|n \\ p \equiv Y}} \left(1 + \frac{M_r(p)}{p^r} \right).$$

Next, since

$$M_r(p) \leq rp^{r-1},$$

we have

$$\begin{aligned} \Sigma_2 &\cong Z \prod_{\substack{p|n \\ p \equiv Y}} \left(1 + \frac{r}{p} \right) \cong Z \prod_{\substack{p|n \\ p \equiv Y}} \left(1 + \frac{1}{p} \right)^r = O \left\{ Z \prod_{\substack{p|n \\ p \equiv Y}} \left(1 + \frac{1}{p} \right)^{2r} \prod_{\substack{p|n \\ p \equiv Y}} \left(1 - \frac{1}{p} \right)^r \right\} = \\ &= O \left\{ Z \prod_{\substack{p|n \\ p \equiv Y}} \left(1 + \frac{1}{p} \right)^{2r} \prod_{\substack{p|n \\ p \equiv Y}} \left(1 - \frac{1}{p} \right)^r \right\} = O \left(Z \log^{2r} Y \prod_{\substack{p|n \\ p \equiv Y}} \left(1 - \frac{1}{p} \right)^r \right), \end{aligned}$$

by the Mertens formula. Therefore, by (10),

$$(17) \quad \Sigma_2 = o \left\{ y \prod_{\substack{p|n \\ p \equiv Y}} \left(1 - \frac{1}{p} \right)^r \right\}.$$

8. Estimation of Σ_1

We first express Σ_1 as a product by the formula

$$(18) \quad \Sigma_1 = \prod_{\substack{p|n \\ p \equiv Y}} \left(1 - \frac{M_r(p)}{p^r} \right),$$

and then evaluate $M_r(p)$. Recalling that $v(p; b_1, \dots, b_{r-1})$ is the number of distinct residues (mod p) in the system $0, b_1, \dots, b_{r-1}$, let $\Psi_r(p, k)$ be the number of such systems for which $v(p; b_1, \dots, b_{r-1}) = k$ and $0 \leq b_i < p$. Then, since

$$\Psi_r(p, k) = \frac{1}{p} \Phi_r(p, k),$$

where $\Phi_r(p, k)$ is the number of systems $b_1, b_2, \dots, b_r$ ($b_i < p$) in which there are exactly k distinct residues (mod p), we have

$$(19) \quad M_r(p) = \frac{1}{p} \sum_{k=1}^p k \Phi_r(p, k),$$

there being in fact no contribution to $M_r(p)$ from the terms in the above sum for which $k > r$.

Next

$$\Phi_r(p, k) = \sum_{x_1, \dots, x_k} \sum_{b_1, \dots, b_r} 1 = \sum_{x_1, \dots, x_k} \Sigma_{x_1, \dots, x_k},$$

say, where the outer sum is over all combinations $\alpha_1, \dots, \alpha_k$ of $0, 1, \dots, p-1$ taken k at a time, and the inner sum is over all arrangements $b_1, \dots, b_r$ of $\alpha_1, \dots, \alpha_k$ in which each value of α_i must occur at least once but possibly more than once. Since the inner sum depends only on k and r , we may write

$$\Sigma_{\alpha_1, \dots, \alpha_k} = u_r(k),$$

and then

$$(20) \quad \Phi_r(p, k) = \binom{p}{k} u_r(k).$$

Let $v_r(k)$ be defined like $u_r(k)$ as the number of arrangements $b_1, \dots, b_r$ of $\alpha_1, \dots, \alpha_k$, but *without* the restriction that each value of α_i must occur at least once as a b_j . Then, applying Legendre's exclusion principle, we infer that

$$u_r(k) = \sum_{t=0}^{k-1} (-1)^t \binom{k}{t} v_r(k-t).$$

But $v_r(k) = k^r$; therefore

$$u_r(k) = \sum_{t=0}^{k-1} (-1)^t \binom{k}{t} (k-t)^r,$$

and so $u_r(k)$ is the coefficient of x^r in the expansion of

$$r!(e^x - 1)^k$$

in ascending powers of x . Hence, by (19) and (20), $M_r(p)$ is the coefficient of x^r in the expansion of

$$\frac{r!}{p} \sum_{k=1}^p k \binom{p}{k} (e^x - 1)^k,$$

there being in fact no contribution to this sum from terms for which $k > r$. Since, for $|z| < 1$,

$$\sum_{k=1}^p k \binom{p}{k} z^k = pz(1+z)^{p-1},$$

we see that $M_r(p)$ is the coefficient of x^r in

$$r! e^{(p-1)x} (e^x - 1) = r! (e^{px} - e^{(p-1)x}),$$

and thus deduce that

$$M_r(p) = p^r - (p-1)^r.$$

Substituting this result in (18) we evaluate Σ_1 explicitly by the formula

$$(21) \quad \Sigma_1 = \prod_{\substack{p|n \\ p \equiv Y}} \left(\frac{p-1}{p} \right)^r = \prod_{\substack{p|n \\ p \equiv Y}} \left(1 - \frac{1}{p} \right)^r.$$

9. Estimation of N_r ; final stage

We estimate N_r by collecting together the results already obtained. We have, by (15), (17) and (21),

$$\Sigma_{r,y} = \frac{y^{r-1}}{(r-1)!} (1+o(1)) \prod_{\substack{p|n \\ p \leq Y}} \left(1 - \frac{1}{p}\right)^r,$$

and then, by this and (9),

$$\begin{aligned} (22) \quad N_r &= n\{1+o(1)\} \frac{y^{r-1}}{(r-1)!} \prod_{\substack{p|n \\ p \leq Y}} \left(1 - \frac{1}{p}\right)^r \cdot \psi_Y^r(n) = \\ &= n\{1+o(1)\} \frac{y^{r-1}}{(r-1)!} \prod_{p|n} \left(1 - \frac{1}{p}\right)^r = \frac{c^{r-1}}{(r-1)!} \varphi(n)\{1+o(1)\}. \end{aligned}$$

10. The distribution of the intervals

The theorem on $f_n(c)$ is an easy deduction from (2) and (22). We have, for any given integer s ,

$$\begin{aligned} f_n(c) &= \sum_{r=2}^{s-1} (-1)^r N_r + AN_s - 1 = \varphi(n) \left(- \sum_{r=2}^{s-1} \frac{(-c)^{r-1}}{(r-1)!} \right) + o\{\varphi(n)\} + \\ &\quad + A\{1+o(1)\} \varphi(n) \frac{c^{s-1}}{(s-1)!} - 1. \end{aligned}$$

Now, for $s > s_0(\mathfrak{H})$,

$$\left| \sum_{m=s}^{\infty} \frac{(-c)^{m-1}}{(m-1)!} \right| \leq \frac{c^{s-1}}{(s-1)!};$$

wherefore

$$f_n(c) = \varphi(n) \left(1 - e^{-c} + B \frac{c^{s-1}}{(s-1)!} \right) + o\{\varphi(n)\} - 1,$$

and thus

$$\left| \frac{f_n(c)}{\varphi(n)} - (1 - e^{-c}) \right| \leq D \frac{c^{s-1}}{(s-1)!} + o(1)$$

for any given $s > s_0(\mathfrak{H})$, where it is to be noted that the left-hand side does not depend on s . Letting $n \rightarrow \infty$ through an sequence of values for which $n/\varphi(n) \rightarrow \infty$, we deduce

$$\overline{\lim}_{n \rightarrow \infty} \left| \frac{f_n(c)}{\varphi(n)} - (1 - e^{-c}) \right| \leq D \frac{c^{s-1}}{(s-1)!}$$

for any fixed $s > s_0(\mathfrak{H})$. Therefore, now taking s to be arbitrarily large, we have

$$\overline{\lim}_{n \rightarrow \infty} \left| \frac{f_n(c)}{\varphi(n)} - (1 - e^{-c}) \right| = 0,$$

or, what is equivalent,

$$(23) \quad f_n(c) = \varphi(n)\{1 + o(1)\}(1 - e^{-c}).$$

We thus have the following theorem.

Theorem 1. *Let $a_1, a_2, \dots, a_{\varphi(n)}$ be, in ascending order of magnitude, the $\varphi(n)$ integers not exceeding n that are prime to n ; let Δ_i be the length of the interval between a_i and a_{i+1} ; and let $f_n(c)$ be the number of intervals for which*

$$\Delta_i < cn/\varphi(n),$$

where c lies in any fixed range $\Re$ that is bounded at either end by positive constants. Then, as $n \rightarrow \infty$ through a sequence of values for which $n/\varphi(n) \rightarrow \infty$, we have

$$f_n(c) = \varphi(n)\{1 + o(1)\}(1 - e^{-c})$$

uniformly in $\Re$.

As stated in the introduction, our other theorem can be deduced from this by using the methods of the earlier paper ⁸⁾.

Theorem 2. *We have for $0 \leq \alpha < 2$*

$$\sum_{i=1}^{\varphi(n)-1} (a_{i+1} - a_i)^\alpha = \{1 + o(1)\} \Gamma(\alpha + 1) n \left(\frac{n}{\varphi(n)} \right)^{\alpha-1},$$

as $n \rightarrow \infty$ through a sequence of values for which $n/\varphi(n) \rightarrow \infty$.

(Received December 5, 1963.)

⁸⁾ See 1).