ACTA ARITHMETICA
V (1958)

On the probability that n and $g(n)$ are relatively prime*

by

P. Erdös (London, Ontario) and G. G. Lorentz (Syracuse)

1. Introduction. It is a well-known theorem of Čebyšev that the probability of the relation $(n, m) = 1$ is $6\pi^{-2}$. One can expect this still to remain true if $m = g(n)$ is a function of n, provided that $g(n)$ does not preserve arithmetic properties of n. In this paper we consider the case when $g(x)$ is the integral part of a smooth function $f(x)$, which increases slower than x. More exactly, let $Q(x)$ be the number of $n \leqslant x$ with the property $(n, g(n)) = 1$. The probability that n and $g(n)$ are relatively prime is then by definition the limit $\lim\limits_{x\to\infty}\{Q(x)/x\}$. Our main result is that if $f(x)$ satisfies some mild smoothness assumptions, has the property (A) $f(x) = o(x/\log\log x)$ and satisfies condition (B) of § 2, then the probability in question exists and is equal to $6\pi^{-2}$. Condition (B) means roughly that $f(x)$ increases faster than the function $\log x \log_4 x$. In § 3 we show that condition (B) is the best possible. Condition (A) may be perhaps relaxed; but it cannot be replaced by $f(x) = O(x/\log\log\log x)$. We also consider the average number of divisors of $(n, g(n))$. This is the limit $\lim\limits_{x\to\infty}\{S(x)/x\}$, where $S(x)$ is the sum of the numbers of divisors of all numbers $(n, g(n))$, $n \leqslant x$. We assume throughout that $f(x)$ is a monotone increasing positive function with a piecewise continuous derivative; $F(y)$ will denote the inverse of $f(x)$. By φ, μ, σ, d we denote the standard number-theoretic functions, by $\log_2 x, \log_3 x, \ldots$ the iterated logarithms of x.

We begin with some elementary identities. Let $Q_k(x)$ be the number of integers $n \leqslant x$ such that n and $g(n)$ have no common factors $\leqslant k$. If $S(x, d)$ is the number of $n \leqslant x$ with $d|(n, g(n))$, then

$$\sum_{d|k!} \mu(d) S(x, d) = \sum_{d|k!} \mu(d) \sum_{d|(n,g(n))} 1 = \sum_{n\leqslant x} \sum_{d|(k!,n,g(n))} \mu(d).$$

* This work was supported in part by the National Science Foundation Grant, NSF G-1975 held by one of the authors at Wayne State University.

By the properties of the function μ, the inner sum is 1 if $(k!, n, g(n)) = 1$, i. e., if $(n, g(n))$ has no divisors greater than 1 and not exceeding k, and otherwise is 0. Hence

$$Q_k(x) = \sum_{d|k!} \mu(d) S(x, d). \tag{1}$$

In particular, if $k = n$, then, since $S(x, d) = 0$ for $d > g(x)$, we obtain

$$Q(x) = \sum_{d=1}^{g(x)} \mu(d) S(x, d). \tag{2}$$

There are similar but obvious formulas for $S(x)$ and $S_k(x)$ — the sum of the numbers of divisors, not exceeding k, of all numbers $(n, g(n))$ with $n \leqslant x$, namely

$$S_k(x) = \sum_{d=1}^{k} S(x, d), \tag{3}$$

$$S(x) = \sum_{d=1}^{g(x)} S(x, d). \tag{4}$$

A function $f(x)$ will be called *homogeneously equidistributed* modulo 1 (or shortly h. e.) if for each integer d,

$$h(x) = \frac{1}{d} f(dx)$$

is equidistributed modulo 1. This means that for each subinterval I of $(0, 1)$, the density of n's for which $h(n) - [h(n)]$ belongs to I, is equal to the length of I.

THEOREM 1. *If $f(x)$ is homogeneously equidistributed, then*

$$\overline{\lim_{x\to\infty}} \frac{Q(x)}{x} \leqslant 6\pi^{-2}, \quad \underline{\lim_{x\to\infty}} \frac{S(x)}{x} \geqslant \frac{1}{6}\pi^2. \tag{5}$$

Proof. It follows from the definition of $S(x, d)$ that this is the number of integers k with $kd \leqslant x$ and $d|g(kd)$; or the number of $k \leqslant xd^{-1}$ so that

$$\frac{1}{d} f(kd) - \left[\frac{1}{d} f(kd)\right]$$

is in the interval $(0, 1/d)$. Since $f(x)$ is h. e., $\lim_{x\to\infty} [S(x, d)/x] = d^{-2}$. Taking now into consideration the relations (1), (3) and the inequalities $Q_k(x) \geqslant Q(x)$, $S_k(x) \leqslant S(x)$, we obtain (5), since

$$\sum_{d=1}^{\infty} d^{-2} \mu(d) = 6\pi^{-2}, \quad \sum_{d=1}^{\infty} d^{-2} = \frac{1}{6}\pi^2.$$

All known simple criteria for $f(x)$ to be equidistributed modulo 1 (by Weyl, Pólya-Szegö, see Koksma [2], p. 88) guarantee also that $af(bx)$ is equidistributed for arbitrary positive constants a, b. The simplest set of conditions is

$$f(x) = o(x) \quad \text{for} \quad x \to \infty, \tag{A_1}$$

$$xf'(x) \to \infty \quad \text{for} \quad x \to \infty, \tag{B_1}$$

and the additional hypothesis that $f'(x)$ decreases. We shall mention here that the last assumption and (B_1) can be replaced by

$$\int_0^y |F''(u)|\, du = o(F(y)) \tag{C_1}$$

($F(y)$ is assumed here to have a piecewise continuous second derivative). If $f'(x)$ decreases, the last integral is equal to $1/f'(x) + \text{const}$ with $x = F(y)$, and hence (C_1) is implied by (B_1). Further natural conditions which in the presence of (B_1) imply (C_1) are

$$\lim_{u\to\infty} \{F''(u)/F'(u)\} = 0 \quad \text{or} \quad \int_0^y |F''(u)|\, du = O(F'(y)).$$

To establish our statement it is sufficient to show that $f(n)$ is equidistributed mod 1 if it satisfies (A_1) and (C_1). Let $I = (\alpha, \alpha+\delta) \subset (0, 1)$, then the number of n's for which $[f(n)] = k$ and $f(n) - [f(n)] \in I$, is $\Delta F_k + O(1)$, where $\Delta F_k = F(s_k) - F(t_k)$, $s_k = k + \alpha + \delta$, $t_k = k + \alpha$, except if $k + \alpha + \delta > f(n) = y$, when $s_k = y$. Because of (A_1), the total number of $n \leqslant x$ with $f(x) - [f(n)] \in I$ is

$$N = \sum_{k+\alpha \leqslant f(x)} \Delta F_k + o(x).$$

Now

$$\left| \frac{\Delta F_k}{\delta} - \{F(s_k) - F(s_{k-1})\} \right| = |F'(\eta_2) - F'(\eta_1)| \leqslant \int_{k-1}^{s_k} |F''(u)|\, du,$$

hence

$$\frac{1}{x} N = \frac{\delta}{x} \sum_{k+\alpha \leqslant f(x)} \{F(s_k) - F(s_{k+1}\} + O\left(\frac{1}{x} \int_0^{f(x)} |F''(u)|\, du\right) = \delta + o(1),$$

by (C_1).

2. The Main Theorem. Our main result is the following:

THEOREM 2. *Let $f(x)$ be* h. e. *and let*

(A) $$f(x) = o(x/\log_2 x),$$

(B) $$\frac{xf'(x)}{\log_3 f(x)} \to \infty,$$

(C) $f'(y) \leqslant Mf'(x)$ *for some constant M for all $y \geqslant x > 0$.*

Then

(6) $$\lim_{x\to\infty} \frac{Q(x)}{x} = \frac{6}{\pi^2}.$$

Proof. Let $Q_k(x)$ be defined as in § 1. Then by (1),

$$\lim_{x\to\infty} \frac{Q_k(x)}{x} = \sum_{d|k!} \frac{\mu(d)}{d^2} = \frac{6}{\pi^2} + \delta_k,$$

where $\delta_k \to 0$ for $k \to \infty$. To prove the theorem it is therefore sufficient to show that

(7) $$\overline{\lim_{x\to\infty}} \frac{R_k(x)}{x}$$

is arbitrarily small if k is sufficiently large. Here $R_k(x) = Q_k(x) - Q(x)$ is the number of $n \leqslant x$ such that for some prime p with $k < p \leqslant g(x)$ we have $p|n$, $p|g(n)$. It follows that

(8) $$R_k(x) \leqslant \sum_{k<p\leqslant g(x)} S(x, p).$$

We consider the contribution to the sum (8) of the part of the curve $y = g(x)$ given by $g(n) = m$; these n satisfy $F(m) \leqslant n < F(m+1)$. We put $k_m = F(m+1) - F(m)$, except when $m+1 > x$, in which case we put $k_m = F(x) - F(m)$. The contribution to $S(x, p)$ is zero if $p \nmid m$, otherwise it does not exceed

$$\frac{1}{p}[F(m+1) - F(m)] + 1 = \frac{k_m}{p} + 1.$$

Hence

(9) $$\begin{aligned} R_k(x) &\leqslant \sum_{k<p\leqslant g(x)} \sum_{\substack{m=1 \\ p|m}}^{g(x)} \left(\frac{k_m}{p} + 1\right) \\ &= \sum_{k<p\leqslant g(x)} \sum_{l\leqslant g(x)/p} \left(\frac{k_{lp}}{p} + 1\right) \\ &\leqslant \sum_{k<p\leqslant g(x)} \frac{g(x)}{p} + \sum_{k<p\leqslant g(x)} \sum_{l\leqslant l_0-1} \frac{k_{lp}}{p} + \sum_{k<p\leqslant g(x)} \frac{k_{l_0 p}}{p} \\ &= \Sigma_1 + \Sigma_2 + \Sigma_3, \end{aligned}$$

say, where $l_0 = [g(x)/p]$. For $x \to \infty$ we have by (A)

$$\Sigma_1 \leqslant g(x) \sum_{p\leqslant x} \frac{1}{p} = O\big(g(x)\log_2 x\big) = o(x).$$

For $l > m$ we have with properly chosen $\xi, \xi_1, \xi \leqslant \xi_1$,

$$k_l = \frac{1}{f'(\xi_1)}, \qquad k_m = \frac{1}{f'(\xi)},$$

hence by (C),

(10) $$k_m \leqslant Mk_l, \qquad l > m.$$

Therefore,

$$k_p + k_{2p} + \ldots + k_{(l_0-1)p} \leqslant \frac{M}{p}(k_p + k_{p+1} + \ldots + k_{l_0 p - 1}) \leqslant \frac{Mx}{p},$$

so that for an arbitrary $\varepsilon > 0$,

$$\Sigma_2' \leqslant \sum_{p\geqslant k} \frac{Mx}{p^2} < \varepsilon x,$$

if k is sufficiently large.

The sum Σ_3 we split into two parts Σ_3', Σ_3'', the first sum being extended over all p for which

(11) $$l_0 p < g(x) + 1 - A\log_2 g(x),$$

and where $A = M\varepsilon^{-1}$, and the second corresponding to p for which the opposite inequality holds. In the first case by (10) and (C),

$$\begin{aligned} k_{l_0 p} &\leqslant \frac{M}{g(x) - l_0 p + 1}(k_{l_0 p} + k_{l_0 p+1} + \ldots + k_{g(x)}) \\ &\leqslant \frac{Mx}{g(x) - l_0 p + 1} < \frac{Mx}{A\log_2 g(x)} = \frac{\varepsilon x}{\log_2 g(x)}, \end{aligned}$$

hence for large x,

$$\Sigma_3' \leqslant \frac{\varepsilon x}{\log_2 g(x)} \sum_{p\leqslant g(x)} \frac{1}{p} \leqslant 2\varepsilon x.$$

In the second case, $g(x) + 1 - A\log_2 g(x) \leqslant l_0 p \leqslant g(x)$, hence p divides one of the consecutive numbers $g(x) + 1 - [A\log_2 g(x)], \ldots, g(x)$, hence also their product N. Clearly,

$$N \leqslant f(x)^{A\log_2 f(x)}.$$

We use the relation (¹)

$$\sum_{p|n} \frac{1}{p} \leqslant C \log_3 n$$

and obtain

$$\Sigma_3'' \leqslant \max_{m \leqslant g(x)} k_m \sum_{p|N} \frac{1}{p} \leqslant \max_{\xi \leqslant x} \frac{1}{f'(\xi)} C \log_3 f(x)^{4\log_2 f(x)}$$

$$\leqslant C_1 \max_{\xi \leqslant x} \frac{1}{f'(\xi)} \log_3 f(x) \leqslant \varepsilon x$$

for large x, by (C) and (B). Substituting our estimates into (9), we obtain that (7) does not exceed 5ε for large x.

THEOREM 3. *Let* $f(x)$ *be* h. e. *and satisfy* (C), *moreover*

(A′) $$f(x) = o(x/\log x),$$

(B′) $$xf'(x)/\log_2 f(x) \to \infty.$$

Then the average order of the number of divisors of $(n, g(n))$ *is* $\frac{1}{6}\pi^2$:

$$\lim_{x\to\infty} \frac{S(x)}{x} = \frac{1}{6}\pi^2.$$

Instead of (8) we have now

$$S(x) - S_k(x) = \sum_{k<n\leqslant g(x)} S(x, n),$$

where n runs through all integers, prime or not. The proof is similar to that of theorem 2, but simpler.

3. Counterexamples. To show that condition (B) is the best possible in Theorem 2, we shall use the following fact. There is an absolute constant

(¹) This result is well known, but since we do not know who first proved it we give a short proof. It follows from the prime number theorem (or from a more elementary result) that

$$\prod_{p<2\log x} p > n.$$

Therefore by a simple argument

$$\sum_{p|n} \frac{1}{p} < \sum_{p<2\log x} \frac{1}{p} < c \log_3 x.$$

C such that for each $\varepsilon_1 > 0$, there is an $\varepsilon_2 > 0$ and infinitely many values of n with the property

(12) $$\frac{\varphi(m)}{m} < \varepsilon_1 \text{ for all } m \text{ with } n \leqslant m \leqslant n + \varepsilon_2 \log_3 n.$$

See [1], p. 129, where $\sigma(m)/m > 2$ is shown to be possible for $n \leqslant m \leqslant n + C_1 \log_3 n$. The same proof establishes $\sigma(m)/m > 1/\varepsilon_1$ in intervals $n \leqslant m \leqslant n + \varepsilon_2 \log_3 n$, and the known connections between φ and σ give (12).

THEOREM 4. *Let* $f(x)$ *be increasing and let*

(B″) $$\frac{xf'(x)}{\log_3 f(x)} \leqslant M.$$

Then

(13) $$\lim_{x\to\infty} \frac{Q(x)}{x} < \frac{6}{\pi^2}.$$

Proof. From (B″) we obtain by integration $f(x) \leqslant \log^2 x$ for all large x. It follows also that $f'(x) \to 0$, hence that $g(x)$ takes all large integral values. From (2), using the argument and notations of § 2 we have, if $d(n)$ is the number of divisors on n,

(14) $$Q(x) = \sum_{d=1}^{g(x)} \mu(d) S(x, d) = \sum_{d=1}^{g(x)} \mu(d) \sum_{d|m} \left\{\frac{k_m}{d} + O(1)\right\}$$

$$= \sum_{m=1}^{g(x)} \sum_{d|m} \frac{\mu(d)}{d} k_m + \sum_{m=1}^{g(x)} O(d(m))$$

$$= \sum_{m=1}^{g(x)} k_m \frac{\varphi(m)}{m} + O\big(g(x) \log g(x)\big)$$

$$= \sum_{m=1}^{g(x)} k_m \frac{\varphi(m)}{m} + O(\log^3 x).$$

We take x such that $g(x) = n$ is one of the n for which (12) holds. Let $x_1 = (1+\delta/M)x$, $\delta > 0$, $n_1 = g(x_1)$. Then we have by (B″) for some $x < \xi < x_1$,

$$n_1 - n \leqslant 1 + f'(\xi)(x_1 - x) \leqslant 1 + \frac{\delta}{M} \xi f'(\xi)$$

$$\leqslant 1 + \delta \log_3 f(\xi) \leqslant 1 + \delta \log_3 n_1,$$

hence $n_1 - n \leqslant \text{const}\cdot\delta\log_3 n$. By (14) and (12),

$$Q(x_1)-Q(x) = \sum_{n<m\leqslant n_1} k_m \frac{\varphi(m)}{m} + O(\log^3 x) < \varepsilon(x_1-x)+O(\log^3 x),$$

for an arbitrary $\varepsilon > 0$, if δ is sufficiently small. This gives

$$\frac{Q(x_1)}{x_1} = \frac{Q(x)}{x}\cdot\frac{x}{x_1}+\varepsilon\frac{x_1-x}{x_1}+o(1),$$

and if α denotes the constant $\alpha = (1+\delta/M)^{-1} < 1$, we obtain by Theorem 1,

$$\underline{\lim}\frac{Q(x_1)}{x_1} \leqslant \frac{6}{\pi^2}\alpha+\varepsilon(1-\alpha) < \frac{6}{\pi^2}.$$

A simple computation shows that $f(x) = c\log x\log_4 x$ satisfies (B″) as stated in the introduction.

In the same way we can prove $\underline{\lim}[Q(x)/x] = 0$, if instead of (B″) we have $xf'(x)/\log_3 f(x) \to 0$.

Similar statements hold for the condition (B′) of Theorem 3. If $f(x)$ is increasing and

$$\text{(B‴)} \qquad xf'(x)/\log_2 f(x) \leqslant M,$$

then

$$\text{(15)} \qquad \overline{\lim}\frac{S(x)}{x} > \frac{\pi^2}{6};$$

and if even $xf'(x)/\log_2 f(x) \to 0$, then $\overline{\lim}\{S(x)/x\} = +\infty$.

To prove for example (15), we note that there are arbitrary large n with $\sigma(n)/n \geqslant C\log_2 n$; if n has this property, we put $f(x) = n$ and $x_1 = x+M^{-1}x/\log_2 f(x)$; then $x_1/x \to 1$ and

$$f(x_1)-f(x) = f'(\xi)(x_1-x) \leqslant \frac{1}{M}f'(\xi)\xi/\log_2 f(\xi) \leqslant 1$$

by (B‴). Hence $k_n \geqslant x_1 - x$. As in (14) we obtain

$$S(x_1)-S(x) = k_n\frac{\sigma(n)}{n}+O(\log^3 x) \geqslant CM^{-1}x+O(\log^3 x),$$

therefore by Theorem 1,

$$\overline{\lim}\{S(x_1)/x_1\} \geqslant \underline{\lim}\{S(x)/x\}+CM^{-1} > \tfrac{1}{6}\pi^2.$$

THEOREM 5. *There exists a function $f(x)$ with the properties*

$$\text{(A″)} \qquad f(x) = O(x/\log_3 x),$$

$$\text{(C″)} \qquad f(x) \text{ is concave and } f'(x) \to 0,$$

such that $\underline{\lim}[Q(x)/x] < 6\pi^{-2}$.

Proof. Let $\varepsilon_1 > 0$ be arbitrary; we select $\delta = \varepsilon_2$ according to (12) and put $l = [\delta\log_3 n]$. For some of the integers n of type (12) we put

$$\text{(16)} \qquad f(x) = \frac{1}{l}(x-n) \quad \text{for} \quad N_n = nl+n \leqslant x < 2N_n.$$

We choose a sequence of n's satisfying (12) in such a way that the intervals $(N_n, 2N_n)$ are disjoint; the function $f(x)$ is obtained by linear interpolation outside of the intervals $(N_n, 2N_n)$. It is easy to check that $f(x)$ is concave and satisfies (A″), (C″).

Moreover, $g(x) = n+s$ for $x = nl+n+sl+t$, $0 \leqslant s \leqslant n$, $0 \leqslant t < l$. Hence the numbers $(m, g(m))$ for $N_n \leqslant m < 2N_n$ are exactly the numbers

$$\text{(17)} \qquad (nl+n+sl+t,\ n+s) = (n+t,\ n+s); \qquad t = 0, 1, \ldots, l-1, \quad s = 0, 1, \ldots, n.$$

Fixing t, we see that the number of $s = 0, 1, \ldots, n$ with $(n+t, n+s) = 1$ is at most $2\varepsilon_1(n+t)$, since $\varphi(n+t)/(n+t) < \varepsilon_1$ by (12). Therefore,

$$Q(2N_n)-Q(N_n) \leqslant 2\varepsilon_1(n+l)l+1,$$
$$\underline{\lim}\{Q(2N_n)/2N_n\} \leqslant \tfrac{1}{2}\overline{\lim}\{Q(N_n)/N_n\}+\varepsilon_1 < 6\pi^{-2},$$

which proves our assertion.

Similarly, there are functions $f(x)$ satisfying (C″) with $f(x) = O(x/\log_2 x)$ for which (15) holds. We take in (16), $l = [\delta\log_2 n]$ and n such that $\sigma(n)/n > C\log_2 n$. Then the sum of the number of divisors of the numbers (17) is greater than

$$\sum_{s=1}^{n} d\big((n, n+s)\big) = \sum_{d|n}\frac{n}{d} = \sigma(n) \geqslant Cn\log_2 n > C_1 N_n$$

with large C_1. Therefore

$$S(2N_n)-S(N_n) \geqslant C_1 N_n,$$

and (15) follows.

At present we can not decide whether condition (A) of Theorem 2 can be weakened to $o(x/\log_3 x)$.

References

[1] P. Erdös, *Note on consecutive abundant numbers*, J. London Math. Soc., 10 (1935), p. 128-131.

[2] J. F. Koksma, *Diophantische Approximationen*, Ergebnisse der Mathematik, Bd. 4 Heft 4, Berlin 1936.

[3] J. Lambek and L. Moser, *On integers n relatively prime to f(n)*, Canadian J. of Math., 7 (1955), p. 155-158.

[4] G. L. Watson, *On integers n relatively prime to [an]*, Canadian J. of Math. 5 (1953), p. 451-455.

THE UNIVERSIY OF WESTERN ONTARIO
and
SYRACUSE UNIVERSITY

Reçu par la Rédaction le 15. 3. 1958

ACTA ARITHMETICA
V (1958)

On a question of additive number theory

by

P. Erdös (Saskatoon) and P. Scherk (Saskatoon)

1. Let $A = \{a\}$, $B = \{b\}, \ldots$ denote sets of non-negative integers containing the number zero;

$$\sum_1^k A_\lambda = \Big\{\sum_1^k a_\lambda\Big\} \quad (a_\lambda \in A_\lambda,\ \lambda = 1, 2, \ldots, k).$$

Thus $\sum A_\lambda$ consists of all the numbers $a_1 + a_2 + \ldots + a_k$ where each a_λ lies in the corresponding A_λ. For a given integer n let $[A]$ denote the number of positive elements of A up to and including n. $\bar{A}$ denotes the set of the integers $\leqslant n$ which do not belong to A.

It is well known and easy to see that $n \notin A+B$ implies $[A]+[B] \leqslant n-1$. The corresponding problem for three or more sets does not lead to anything new. For then

$$n \notin \sum_1^k A_\lambda \tag{1}$$

implies $n \notin A_\lambda + A_\mu$ and thus $[A_\lambda]+[A_\mu] \leqslant n-1$; $1 \leqslant \lambda < \mu \leqslant k$. Adding these $\frac{1}{2}k(k-1)$ inequalities we readily obtain

$$\sum_1^k [A_\lambda] \leqslant \tfrac{1}{2}k(n-1). \tag{2}$$

That (2) cannot be improved can be seen by taking $A_1 = A_2 \ldots = A_k =$ $=$ set of integers between $[\frac{1}{2}n]+1$ and $n-1$ together with 0.

This question becomes more interesting if we require n to be the smallest number not in $\sum A_\lambda$. For $k = 3$ and $n < 15$ one can show[1] that

$$[A_1]+[A_2]+[A_3] \leqslant n-1.$$

[1] Written communication from Professor H. B. Mann.