On the probability that \( n \) and \( f(n) \) are relatively prime II

by

R. R. Hall (York)

Let \( f(n) \) be an additive function and set

\[
T(x) = \sum_{\substack{n \leq x \\ (n, f(n)) = 1}} 1.
\]

Our ultimate object is to find the weakest conditions on \( f \) which ensure that

\[
T(x) \sim \frac{6}{\pi^2} x.
\]

In the preceding paper [1] we showed that in the particular case

(1)

\[
f(n) = \sum_{p \mid n} p,
\]

we have

\[
T(x) = \frac{6}{\pi^2} x + O\left( \frac{x}{(\log x)^{16}(\log \log x)^{10}} \right)
\]

where we use the familiar notation \( \log \) for iterated logarithms. Our immediate object is to extend this result, and we are able to replace the \( p \) in (1) by a class of functions of \( p \) which include the polynomials as a special case.

The integer valued function \( g(n) \) will be called a pseudo-polynomial if

\[
g(n + k) = g(n) \mod k
\]

for all \( n \) and \( k \). Every polynomial with integer coefficients is a pseudo-polynomial, but not all pseudo-polynomials are polynomials, and I am grateful to Dr. Woodall at Nottingham University for constructing an example, which will be described later. We have

**Theorem 1.** Let \( g(n) \) be a pseudo-polynomial. For each prime \( p \) define

\[
B(p) = \max_{0 < b < p - 1} \sum_{\substack{n = 1 \\ g(n) = b \mod p}}^p 1
\]

and suppose that \( g \) satisfies the following conditions:
(i) For each square-free \( q \) there exists an \( a \), prime to \( q \), for which
\[
g(a) \equiv 0 \mod q.
\]

(ii) The series
\[
\sum_{p} \frac{1}{p} \left( \frac{B(p)}{p} \right)^{1/2}
\]
is convergent; and

(iii) \( \log(1 + |g(n)|) = O(n^{4/\log 3} m) \).

Suppose that
\[
f(n) = \sum_{p \leq n} g(p).
\]

Then there exists an absolute constant \( C \), independent of \( g \) such that
\[
T(n) = \sum_{p \leq n} \frac{1}{p} = \frac{6}{\pi^2} n + O\left( n \sum_{p \leq n} \frac{1}{p^{1/2}} + \frac{n}{\sqrt{\log n}} \right).
\]

Two questions naturally present themselves:

(a) Does every polynomial satisfy these conditions?

(b) Is there a pseudo-polynomial, which satisfies the conditions and is not a polynomial?

The answer to (a) is no, even if we restrict ourselves to polynomials whose coefficients have highest common factor 1. For example, \( g(n) = n^2 + 3n + 2 \) does not satisfy condition (i) for \( g = 6 \). However, with a slight modification we are more successful:

**Theorem 2.** Let \( g^*(n) \) be any polynomial with integer coefficients. Then there exists a constant \( m = m(g^*) \) depending on \( g^* \), such that the new polynomial defined by
\[
g(n) = g^*(n) + m(g^*)
\]
satisfies the conditions of Theorem 1.

I am unable to provide the answer to question (b). However, it will be shown that the Woodall pseudo-polynomial can be constructed to satisfy the first two conditions.

I am grateful to Professor Erdős for finding the proof of Lemma 2 during his visit to Nottingham in 1969.

**Proofs of the Theorems.** We give proof of Theorem 2 first, as it is shorter.

Suppose that \( g^* \) is of degree \( d \); thus for any choice of \( m \),
\[
g(n) = O(n^d)
\]
and for every \( p \),
\[
B(p) \leq d.
\]

Thus conditions (ii) and (iii) are satisfied, and in fact are very weak for polynomials.

The number of solutions of
\[
g(n) = g^*(n) + m = 0 \mod q
\]
is at most \( \sigma(q) \), whatever the choice of \( m \), since \( q \) is square-free. Since
\[
n = \frac{\log q}{\log \log q}
\]
\[r(q) \leq \frac{\log q}{\log \log q}
\]
it follows that for each \( d \) there exists a constant \( Q = Q(d) \) such that for \( q > Q \),
\[
\sigma(q) < q^{\sigma(q)}
\]
and hence that every polynomial of degree \( d \) satisfies condition (i) except perhaps for some values of \( q \) less than \( Q(d) \).

We can choose \( m = m(g^*) \) such that
\[
g(1) = g^*(1) + m \equiv 0 \mod p
\]
for every prime \( p < Q \), by the Chinese remainder theorem. It follows that for \( q < Q \), there is at least one \( a \), namely \( a = 1 \), such that
\[
g(a) \equiv 0 \mod q,
\]
and for \( q > Q \) the conclusion follows from (2). This completes the proof.

**Proof of Theorem 1.** We only give those details of the proof which differ materially from the proof contained in [1].

**Lemma 1.** For \( p \leq x \) and all \( \alpha \),
\[
\sum_{f(m) = a \mod p} \mu(m) \ll \alpha \left( \frac{B(p)}{p} + \frac{\log p}{\log x} \right).
\]

This is proved as in [1]; as before our next step is to replace this estimate over square-free \( m \) by a similar one for all \( m \). The following lemma replaces Lemma 3 of the previous paper, the proof being due to Professor Erdős.

In the next paper of this series we prove rather more: for each fixed \( r \) we have
\[
\sum_{m \leq x} Q^r(x, m) \ll \omega
\]
and this enables us to use Hölder's inequality in place of the Cauchy–Schwarz inequality in the application. Therefore the exponent 1/2 of \( B(p)/p \) in Theorem 1 could be improved to any fixed number < 1.
Lemma 2. Let \( Q(x, m) \) denote the number of integers \( n \leq x \) whose square-free kernel, that is

\[
\prod_{p \mid n} p
\]
is equal to \( m \). Then

\[
\sum_{n \leq x} Q^2(x, m) \ll x.
\]

Proof. We have

\[
\sum_{n \leq x} Q^2(x, m) = \sum_{k=1}^{\infty} k^3 \sum_{Q(x, m) = k} 1 \leq \sum_{k=1}^{\infty} k^3 \sum_{Q(x, m) > k} 1
\]
so that it is sufficient to show that for each \( k \) the number of \( m \)'s for which \( Q(x, m) \geq k \) does not exceed \( \Delta x/k^4 \) for some constant \( \Delta \) independent of \( k \) and \( x \). For the \( m \)'s not exceeding \( x/k^4 \) we make the simple estimation

\[
\sum_{Q(x, m) \leq x/k^4} 1 \leq x/k^4.
\]

Next, let \( m > x/k^4 \) and suppose \( m \) has \( s \) distinct prime factors not exceeding \( k^4 \). If \( n \) has square-free kernel \( m \) and \( n \leq x \),

\[
m = n_1 p_1^{a_1} \cdots p_s^{a_s}, \quad a_i \geq 0
\]
and we are looking for the number of solutions of the inequality

\[
a_1 \log p_1 + a_2 \log p_2 + \ldots + a_s \log p_s \leq \log \frac{x}{m}, \quad a_i \geq 0;
\]
which does not exceed the number of solutions of

\[
(a_1 + a_2 + \ldots + a_s) \log 2 \leq 4 \log k.
\]

Let \( V_r(y) \) be the number of solutions of the inequality

\[
\beta_1 + \beta_2 + \ldots + \beta_s \leq y, \quad \beta_i \geq 0.
\]
Plainly

\[
V_r(y) = \sum_{\nu=0}^{\lfloor y \rfloor} V_{r-1}(y - \beta_1) \leq \int \nu^{r-1} V_{r-1}(t) \, dt,
\]

\( V \) being monotonic, and since \( V_1(y) \leq y + 1 \) it follows by induction that

\[
V_r(y) \leq \frac{(y + r)^r}{r!}.
\]

Thus if \( m > x/k^4 \),

\[
Q(x, m) \leq \frac{(c \log k + s)^s}{s!}, \quad c = \frac{4}{\log 2},
\]

where \( s \) is the number of prime factors of \( m \) not exceeding \( k^4 \). If \( Q(x, m) \geq k \), setting \( s = u \log k \) and noting that \( e^s \geq (s/e)^s \), we deduce that

\[
\left( \frac{\log k + u \log k}{\frac{u}{e} \log k} \right)^{u \log k} \geq k = e^{o_k}
\]
and so

\[
\left( e \left( 1 + \frac{c}{u} \right) \right)^u \geq e.
\]

Hence \( u \geq c' \), an absolute constant which could be derived from the value of \( c \). Hence \( m \) must have at least \( c' \log k \) distinct prime factors not exceeding \( k^4 \), and the number of such \( m \)'s does not exceed

\[
\sum_{p_1 < k^4} \sum_{p_2 < k^4} \ldots \sum_{p_s < k^4} \sum_{m \in \mathbb{Z}} 1 \leq \frac{x}{x} \left( \sum_{p < k^4} \frac{1}{p} \right)^s \leq x \left( \frac{e}{s} \sum_{p < k^4} \frac{1}{p} \right)^s
\]
where \( s \) is the least integer not less than \( c' \log k \). Now there exists an absolute constant \( c'' \) such that

\[
\sum_{p < k^4} \frac{1}{p} \leq \log \log k + c''
\]
and a constant \( k_0 \) such that for \( k \geq k_0 \),

\[
e(\log \log k + c'') \leq \frac{c' \log k}{e^{c''}}
\]
and for these \( k \) the sum above does not exceed \( x/k^4 \). For \( k > k_0 \) it does not exceed \( e^{-c''} x \leq B x/k^4 \) where \( c'' \) and \( B = c'' k_0^4 \) are again absolute constants. Putting these results together we find that the number of \( m \)'s for which \( Q(x, m) \geq k \) does not exceed

\[
\frac{x}{k^4} + \max \left( 1, B \right) \frac{x}{k^4} \leq \frac{A x}{k^4}
\]
which completes the proof.

Lemma 3. For all \( p \leq \sqrt{x} \) and all \( a \),

\[
\sum_{\text{m|n = a mod p}} 1 \leq a \left( \sqrt{\frac{B(p)}{p}} + \sqrt{\frac{\log p}{\log x}} \right)
\]
Proof. Denoting the sum on the left by \( S \) we have,
\[
S = \sum_{\substack{m = 1 \atop m \equiv a \mod p}}^\infty |\mu(m)|Q(x, m)
\]
and so by the Cauchy–Schwarz inequality,
\[
S^2 \leq \left( \sum_{n \leq x} Q^2(x, m) \right)^{\frac{1}{2}} \left( \sum_{n \leq x} |\mu(m)|^2 \right)^{\frac{1}{2}} \leq x^{2\frac{1}{2}} \left( \frac{B(p)}{p} + \frac{\log p}{\log x} \right)
\]
by the last two lemmas. The result follows.

**Lemma 4.** Under the conditions on \( g \) given in the theorem, for each \( q \) we have
\[
\sum_{\substack{a \equiv q \mod 2 \atop f(a) = 0 \mod 2}} 1 = \frac{x}{q} + O\left( \frac{x \exp(C_1 q \log q)}{\log x} \right)
\]
where \( C_1 \) is an absolute constant, independent of \( q \).

**Proof.** We follow Lemmas 6 and 7 of [1]. Setting
\[
F_q(s, l | q) = \sum_{n=1}^{\infty} \frac{1}{n^s} \exp\{2\pi i f(nq) - f(q)l|q\},
\]
we find that
\[
\frac{1}{q} \sum_{a=1}^{q} e^{2\pi i f(a)q/q} F_q(s, l | q) = \sum_{\substack{a \equiv q \mod 2 \atop f(a) = 0 \mod 2}} n^{-s}.
\]
Since \( F_q(s, l) = \zeta(s) \), the result will follow if we can show that for those \( l < q \), \( F \) is regular and not too large in some region to the left of the line \( \Re s = 1 \). Now
\[
F_q(s, l | q) = F^*_q(s, l | q) \prod_x (\zeta(x, \chi))^{\eta(x, l | q(0))}
\]
where \( F^*_q \) is regular and bounded by \( q^l \) for \( \Re s > \frac{1}{2} \). It involves the prime factors of \( q \) itself. Here
\[
\tau_q(z, l) = \sum_{a=1}^{q} \chi(a) e^{2\pi i a l/q}.
\]
The first half of the proof is identical to the old Lemma 7. However, we then used the fact that for \( (l, q) = 1 \),
\[
\sum_{a=1}^{q} \tau_q(a) e^{2\pi i l a/q} = \mu(q);
\]
in fact, all that is required is that its real part is bounded away from \( \varphi(q) \), that is, that no \( \tau_q \) has a simple pole at \( s = 1 \). Now in the present case,
\[
1 - R \frac{\tau_q(\alpha, l)}{\varphi(q)} \frac{1}{\varphi(q)} \sum_{a=1}^{q} 2 \sin^2 \frac{\pi g(a)}{q} \geq 1 \quad ((l, q) = 1)
\]
under the condition of the theorem that \( g(a) \neq 0 \mod q \) for some \( a \) prime to \( q \). The rest of the proof follows as before.

**Lemma 5.** We have that
\[
\sum_{\substack{a \equiv q \mod 2 \atop f(a) = 0 \mod 2}} \sum_{\substack{m=1 \atop m \equiv \alpha \mod p}}^\infty 1 = O\left( \frac{x \log x}{\log \log x} \right)
\]
provided
\[
\log H \geq 2A(\log x)(\log x).
\]

**Proof.** Either \( f(mp) = f(m) \) or \( f(mp) + f(p) \) according to whether \( p \mid m \) or not. Now \( g(p) = g(0) \mod p \) so the summation condition is that \( p \mid f(m) \) or \( f(m) + g(0) \); we allow either possibility the sum will be increased. We invert the order of summation and estimate the number of prime factors of \( f(m) \) and \( f(m) + g(0) \). The above sum does not exceed
\[
\sum_{\substack{m \equiv \alpha \mod H \atop f(m) \neq f(m) + g(0)}} \sum_{\substack{m \equiv \alpha \mod H \atop f(m) \neq f(m) + g(0)}} \pi\left( \frac{x}{m} \right) + 2 \sum_{\substack{m \equiv \alpha \mod H \atop f(m) \neq f(m) + g(0)}} \frac{\log |f(m)| + |g(0)|}{\log H}
\]
say. Now
\[
f(m) = \sum_{\substack{m \equiv \alpha \mod H \atop f(m) \neq f(m) + g(0)}} g(p) = O\left( \frac{\log m}{\log \log m} \max_{n \leq m} |g(n)| \right)
\]
and so for \( f(m) \neq 0 \) or \( -g(0) \) and \( m \leq x \),
\[
\log |f(m)| + |g(0)| = O\left( x \log x \right).
\]
It follows that
\[
S_1 \ll \frac{2}{H \log H} = O\left( \frac{x \log x}{\log \log x} \right).
\]
We split \( S_1 \) into two terms, \( S'_1 \) and \( S'_1 \) according as \( f(m) = 0 \) or \( f(m) = -g(0) \), and it is sufficient to treat \( S'_1 \) the other case being similar.
For any prime \( \omega \) we have

\[
S_1 \ll \sum_{\substack{w \leq H \leq \infty \atop f(\omega) \equiv f(\omega, \text{mod} \ \omega)}} \frac{x}{w} \ll \frac{x}{\log H} \sum_{\substack{w \leq H \leq \infty \atop f(\omega) \equiv f(\omega, \text{mod} \ \omega)}} \frac{1}{w}.
\]

Now

\[
\sum_{\substack{w \leq H \leq \infty \atop f(\omega) \equiv f(\omega, \text{mod} \ \omega)}} \frac{1}{w} = \int_{1}^{\infty} \left( \sum_{\substack{w \leq H \leq \infty \atop f(\omega) \equiv f(\omega, \text{mod} \ \omega)}} \frac{1}{y^2} \right) dy + \frac{1}{w} \sum_{\substack{w \leq H \leq \infty \atop f(\omega) \equiv f(\omega, \text{mod} \ \omega)}} 1
\]

\[
\ll \int_{1}^{\infty} \left( \sum_{\substack{w \leq H \leq \infty \atop f(\omega) \equiv f(\omega, \text{mod} \ \omega)}} \frac{1}{y^2} \right) dy + \int_{1}^{\infty} \left( \sqrt{\frac{E(\omega)}{\omega}} + \sqrt{\frac{\log \omega}{\log y}} \right) dy + \sqrt{\frac{E(\omega)}{\omega}} + \sqrt{\frac{\log \omega}{\log x}}
\]

if \( \omega \ll \sqrt{x} \). Since the series

\[
\sum_{\substack{w \leq H \leq \infty \atop f(\omega) \equiv f(\omega, \text{mod} \ \omega)}} \frac{1}{w} \left( \frac{B(p)}{p} \right)^{1/2}
\]

is convergent, its partial sums are bounded and for any \( K \) there exists an \( \omega \ll K \) such that

\[
\left( \frac{B(\omega)}{\omega} \right)^{1/2} = O\left( \frac{1}{\log \log K} \right).
\]

Hence for all \( K \ll \sqrt{x} \) we have

\[
S_1 = O\left( \frac{x}{\log H} \left( \frac{\log x}{\log \log K} + \sqrt{\frac{\log K}{(\log \log x)^2}} \right) \right)
\]

and we select

\[
\log K = \frac{\log x}{(\log \log x)^2}.
\]

Since \( \log H \geq (2 \log \log x) / \log x \) we obtain our result.

Proof of the Theorem. Set

\[
P(x) = \prod_{p \leq x} p.
\]

Then for all \( x \),

\[
T(x) = \sum_{w \leq H} \mu(w) + \theta \sum_{w \leq H} \sum_{p \leq x} \frac{1}{p}
\]

where \( |\theta| \leq 1 \).

And therefore

\[
T(x) = \sum_{w \leq H} \mu(w) \sum_{\substack{n \equiv 0 \mod w \atop \text{gcd}(w, \text{gcd}(n, w)) = 1, x}} 1 + \theta \sum_{w \leq H} \sum_{\substack{n \equiv 0 \mod w \atop \text{gcd}(w, \text{gcd}(n, w)) = 1, x}} 1
\]

\[
= \sum_{w \leq H} \frac{\mu(w)}{w^2} + \sum_{w \leq H} \mu(w) \left\{ \sum_{\substack{n \equiv 0 \mod w \atop \text{gcd}(w, \text{gcd}(n, w)) = 1}} 1 - \frac{x}{w^2} \right\}
\]

\[
+ \theta \sum_{w \leq H} \sum_{\substack{n \equiv 0 \mod w \atop \text{gcd}(w, \text{gcd}(n, w)) = 1}} 1 + \theta \sum_{w \leq H} \sum_{\substack{n \equiv 0 \mod w \atop \text{gcd}(w, \text{gcd}(n, w)) = 1}} 1
\]

\[
= \frac{6}{\sqrt{x}} O\left( \frac{x}{\log \log x} \right) + O\left( \sum_{w \leq H} \exp \left\{ C_1 \sqrt{q \log q - \frac{\log \log x}{q^2}} \right\} \right)
\]

\[
+ O\left( \sum_{\substack{w \leq H \leq \infty \atop f(\omega) \equiv f(\omega, \text{mod} \ \omega)}} \frac{1}{w} \left( \frac{B(p)}{p} \right)^{1/2} \right) + O\left( \sqrt{\frac{\log H}{\log x}} + \left( \frac{x \log x}{\log \log x} \right) \right).
\]

There exists an absolute constant \( C_1 \) such that every

\[
q \ll \sqrt{e} x.
\]

Thus for \( x = C \log \log x \) and \( \log H = (2 \log \log x) / \log x \) we have

\[
T(x) = \frac{6}{\sqrt{x}} O\left( \frac{x}{\log \log x} \right) + O\left( \sum_{\substack{w \leq H \leq \infty \atop f(\omega) \equiv f(\omega, \text{mod} \ \omega)}} \frac{1}{w} \left( \frac{B(p)}{p} \right)^{1/2} \right).
\]

This completes the proof.

The Woodall pseudo-polynomial. The pseudo-polynomials form a ring, of which the ring \( Z[x] \) of polynomials with integer coefficients is a sub-ring. There are several interesting questions we can ask about the algebraic structure of this ring, for example, whether it is an integral domain; all we are going to show now is that there is an infinite class of pseudo-polynomials which are not polynomials.

Choose (integer) values for \( g(0) \) and \( g(1) \) arbitrarily. We may then select \( g(2) = g(0) \mod 2 \) so that it is not the value of the linear function of \( n \) determined by \( g(0) \) and \( g(1) \).

Next, select \( g(3) = g(0) \mod 3 \) and \( = g(1) \mod 2 \) so that it is not the value of the quadratic function determined by \( g(0), g(1) \) and \( g(2) \). Proceeding indefinitely, we obtain a pseudo-polynomial which is not a polynomial. Thus \( Z[x] \) is a proper sub-ring of the pseudo-polynomials, and a coset of \( Z[x] \) (regarded additively or multiplicatively) gives an infinite class; alternatively, each pair of values of \( g(0) \) and \( g(1) \) gives a different pseudo-polynomial.
Remarks. At each stage of the construction, we have to solve a congruence
\[ g(n) = t \mod N \]
where \( N \) is the lowest common multiple of the integers not exceeding \( n \). We may select at least one of the first two solutions of this congruence, so that
\[ g(n) \leq \sigma^4 n \]
for some fixed \( A \). But this is not good enough for condition (iii).
Condition (i) is easily arranged by setting \( g(1) = 1 \).
Condition (ii) is more difficult. Nothing in the construction implies that the numbers \( g(0), g(1), g(2), \ldots, g(p-1) \) are well distributed mod \( p \); in fact \( B(p) \) could be \( p \). We can make \( g \) satisfy (ii) by selecting \( g(n) \) to satisfy congruences to moduli \( p > n \), but so far as I can see at the expense of dropping condition (iii). Suppose that for \( n < p \leq t(n) \) (some increasing function of \( n \)) we set
\[ g(n) = t_p(n) \mod p \]
where \( t_p(n) \) is one of the most deficient residue classes mod \( p \) so far. Then for all \( p \),
\[ B(p) \leq t^{-1}(p) \]
that is, the number of \( n \) for which \( g(n) \) is not corrected mod \( p \). Roughly we want
\[ t^{-1}(p) \leq (\log \log p) \alpha \]
for some \( \alpha > 2 \), so that we shall satisfy conditions (i) and (ii) if for example
\[ t(n) = n (\log \log n)^{\alpha} \]
This however, could make \( \log (1 + |g(n)|) \) too large. The conclusion is that there are infinitely many pseudo-polynomials satisfying the first two conditions, which are not polynomials.

I do not know of any number-theoretic function which presents itself naturally and is a pseudo-polynomial. The chances are that it would satisfy our conditions, and this is one way that the problem could be solved.

Reference


Received on 14.4.1970

\[ (31) \]