On functions connected with prime divisors of an integer

P. Erdős
A Magyar Tudományos Akademia
Matematikai Kutató Intézet
Reáltanoda u. 13-15, Pf 127
H 1364 BUDAPEST, HUNGARY

J. L. Nicolas
Département de Mathématiques
Université de Limoges
123 Av. Albert Thomas
F-87060 Limoges cedex
FRANCE

Let \( n \) be an integer. We write its standard factorization into primes

\[
n = q_1 q_2 \ldots q_k
\]

with \( q_1 < q_2 < \ldots < q_k \).

We define:

\[
\begin{align*}
f(n) &= \sum_{i=1}^{k-1} \frac{1}{q_i q_{i+1}} ; \\
F(m) &= \sum_{i=1}^{k-1} \left(1 - \frac{1}{q_i q_{i+1}} \right). \\
h(n) &= \sum_{i=1}^{k-1} \frac{1}{q_i q_{i+1}} ; \\
h(n) &= \sum_{1 \leq i < j \leq k} \frac{1}{q_j - q_i}
\end{align*}
\]

and \( \sigma(n) = k \). When \( k = 1 \), the above empty sums are 0. Moreover, we say that \( n \) is a champion for the function \( f \) (or an \( f \)-champion) if

\[
m < n \Rightarrow f(m) < f(n).
\]

In [Erd 2], it was shown that \( n(x) = \prod_{p \leq x} p \) was a \( f \)-champion for \( x \)

large enough, but was not a \( F \)-champion for all \( x \) large enough. We shall consider here the following problem. Is \( n(x) \) a \( h \)-champion?

a \( h \)-champion?

In [Erd 3] and [De K], function \( h \) is studied. It is shown that

\[
\frac{\log n(x)}{\log \log n(x)} < h(n(x)) < \frac{\log n(x) \log \log \log n(x)}{(\log \log n(x))^2}, \quad (1)
\]

For all \( n \), we have:

381

\[ h(n) \leq \vartheta(n) \ll \frac{\log n}{\log \log n}. \]

Let \( t_1 = 3, t_2 = 5, t_3 = 7, t_4 = 11, t_5 = 13, \ldots \) be the sequence of twin primes, and let us assume that this sequence is infinite and that \( t_k \ll k \log^2 k \). Then for the sequence \( n_k = t_1 t_2 \ldots t_k \), it is not difficult to see that

\[ h(n_k) \ll \frac{\log n_k}{\log \log n_k}. \]

With (1), this relation shows that, for \( x \) large enough, \( n(x) \) is not a \( h \)-champion. But we have assumed a strong hypothesis about twin primes. Without any conjecture, we shall prove:

**Theorem 1.** Let \( n(x) = \prod p \). For \( x \) large enough, \( n(x) \) is not a \( p \times x \)-champion, i.e. there exists \( m < n(x) \) with \( h(m) > h(n(x)) \).

**Proof.** It follows from Maier’s result (cf [Mai]) that there exists an absolute constant \( D > 1 \), such that for all \( k \) and for \( x \) large enough, there exist between \( x^{1/D} \) and \( x \), \( k \) consecutive primes \( p_1, \ldots, p_k \) and a constant depending on \( k \), say \( \epsilon(k) \), with the property:

\[ p_{i+1} - p_i \geq \epsilon_k (\log x) \varphi(x), \quad 1 \leq i \leq k - 1, \]

where \( \varphi(x) \) is a function going to infinity with \( x \).

We apply this result with \( k = 2D + 3 \). Moreover between \( x \) and \( 2x \), there certainly exist 2 prime \( q_1 \) and \( q_2 \) such that the difference \( q_2 - q_1 \leq \frac{11}{10} \log x \). We consider

\[ m = \frac{n(x), q_1 q_2}{p_2 \cdots p_{2D+2}} \leq \frac{4x^2}{x^{(2D+1)/D}} n(x). \]

Thus \( m \) is smaller than \( n(x) \) for \( x \) large enough. Further:

\[ h(m) \geq h(n(x)) + \frac{1}{q_2-q_1} - \frac{2D+2}{\sum_{i=1}^{2D+2} p_i+1 - p_i} \]

\[ \geq h(n(x)) + \frac{10}{\prod \log x} - \frac{(2D+2)}{\epsilon_k \log x \varphi(x)} \]
which is bigger than $h(n(x))$ for $x$ large enough.

Unfortunately we were not able to prove the same theorem than theorem 1 for the function $h$. To get the same result we need 2 very strong conjectures:

(H1) \[ \forall \epsilon > 0 , \forall \eta > 0 , \exists x_0 \text{ such that for } x \geq x_0 \text{ and } y \geq x^\epsilon , \]
\[ (1 - \eta) \frac{y}{\log x} \leq \pi(x) - \pi(x - y) \leq (1 + \eta) \frac{y}{\log x} . \]

(H2) There exists a fixed $\beta < 1/100$ such that, for $x$ large enough, it is always possible to find between $x$ and $x + x^\beta$, four primes $q_1, q_2 = q_1 + 2, q_3 = q_1 + 6, q_4 = q_1 + 8$.

Hypothesis (H1) has been partially proved by Hoheisel for a fixed $\epsilon < 1$. The Riemann hypothesis implies (H1) for all $\epsilon > 1/2$. We shall prove:

Theorem 2. Under the assumption of (H1) and (H2), for $x$ large enough, $n(x) = \prod_{p \leq x} p$ is never a $h$-champion number.

To prove theorem 2, we need 3 lemmas.

Lemma 1. There is an absolute constant $K$ such that for all $x, y, d \in \mathbb{Z}$, $2 \leq y \leq x$,
\[ \sum_{q \text{ prime}} \frac{1}{\log^2 \frac{y}{p}} \prod_{p | d} (1 + 1/p). \]
\[ \sum_{x - y < q \leq x} \frac{1}{|q - d|} \text{ is prime} \]

Moreover
\[ \sum_{1 \leq d \leq x} \frac{1}{d} \prod_{p | d} (1 + 1/p) \leq K' \log x. \]

Proof. The first part is a classical application of sieve's method, (cf[Hall], Cor. 2.4.1, or [Sie] for an effective value of $K$). For the second fact, let us call $w(d) = \frac{1}{d} \prod_{p | d} (1 + 1/p)$. It is a multiplicative function, and,
\[ \sum_{d \leq x} w(d) \leq \prod_{p \leq x} (1 + w(p) + \ldots + w(p^k) + \ldots) \]
We complete the proof by using Mertens formula (cf [Harl]) to estimate the first product and observing that the second product is convergent.

**Lemma 2.** Let \( 0 < \alpha < \beta < 1 \) be fixed real numbers. We define

\[
U(x, \alpha, \beta) = \sum_{x-x^\alpha < p \leq x^\beta} \frac{1}{x-p}.
\]

Under the assumption of hypothesis (H1), we have for \( x \) going to infinity:

\[
U(x, \alpha, \beta) = \beta - \alpha + o(1).
\]

**Proof.** We apply (H1) with \( \varepsilon = \alpha, \gamma, x \) and \( y = x - p \). We get for \( p \leq x-x^\alpha \), and \( x \) large enough:

\[
\frac{(1-\gamma)(x-p)}{\log x} \leq \pi(x) - \pi(p) \leq (1+\gamma) \frac{x-p}{\log x}
\]

and

\[
\frac{1}{\log x(\pi(x)-\pi(p))} \leq \frac{1+\gamma}{x-p} \leq \frac{1}{\log x} \frac{1}{(\pi(x)-\pi(p))}.
\]

Further, we apply (H1) with \( \varepsilon = \alpha, \gamma, x, y = x^\alpha \):

\[
\frac{1-\gamma}{\log x} x^\alpha \leq \pi(x) - \pi(x-x^\alpha) \leq \frac{1+\gamma}{\log x} x^\alpha.
\]

The same inequality holds with \( \beta \) instead of \( \alpha \).

Then we have

\[
U(x, \alpha, \beta) \leq \frac{1+\gamma}{\log x} \sum_{x-x^\alpha < p \leq x^\beta} \frac{1}{\pi(x)-\pi(p)}
\]

\[
= \frac{1+\gamma}{\log x} \sum_{x-x^\alpha < p \leq x^\beta} \frac{1}{\pi(x)-\pi(x-x^\alpha)} \sum_{j \leq \pi(x)-\pi(x-x^\alpha)} \frac{1}{x^j}.
\]
\[ \leq \frac{1 + \eta}{\log x} \sum_{j} 1/j, \]

where \( j \) runs between \( \frac{(1 - \eta) x^\alpha}{\log x} \) and \( \frac{(1 + \eta) x^\beta}{\log x} \).

We deduce:

\[ U(x, a, \beta) \leq (1 + \eta) (\beta - a + o(1)). \]

In the same way we can obtain the lower bound

\[ U(x, a, \beta) \geq (1 - \eta) (\beta - a + o(1)), \]

and choosing \( \eta \) as small as we want completes the proof of lemma 2.

**Lemma 3.** For \( q \) prime, and real \( x \), we define:

\[ V(q) = \sum_{p \leq x} \frac{1}{q - p} \quad \text{and} \quad W(q, x) = \sum_{q \leq x} \frac{1}{p - q}. \]

Then we have under the assumption of (H1):

\[ \lim_{x \to \infty} V(q) \geq 1, \]

and for \( 0 < a < 1 \),

\[ \sum_{x - x^a < q \leq x} V(q) + W(q, x) \leq (1 + a + o(1)) \frac{x^a}{\log x}. \]

**Proof.** With the notation of lemma 2, we get:

\[ V(q) \geq U(q, a, 1) \]

for all \( a > 0 \), and thus \( \lim_{x \to \infty} V(q) \geq 1 \). We observe that replacing hypothesis (H1) by Pachet's theorem will give

\[ \lim_{x \to \infty} V(q) > 0. \]

We have now to prove (4). We choose \( \epsilon > 0 \), and \( \epsilon < a \). Then, we have:

\[ \sum_{x - x^a < q \leq x} U(q, \epsilon, 1) + \sum_{x - x^a < q \leq x} \frac{1}{d} \sum_{d \mid q} \epsilon^d \]

**Lemma 2** tells us that first sum is
\[ (\pi(x) - \pi(x-x^\theta)) (1-\epsilon+o(1)) \]

which, by (H1) is smaller than \((1+\epsilon) \frac{x^\theta}{\log x} (1+o(1))\). Applying lemma 1 to the second sum shows that it is bounded above by

\[ K \sum_{d \leq x^{\epsilon \delta}} \frac{x^\theta}{d^2 \log^2 x} \prod_{p \mid d} (1 + \frac{1}{p}) \leq KK' \frac{e^\theta}{\log x} \cdot \]

And, since we can choose \(\epsilon\) as small as we want, this completes the proof of

\[ \sum_{x-x^\theta < q \leq x} V(q) = (1+o(1)) \frac{x^\theta}{\log x}. \]

It remains to evaluate

\[ \sum_{x-x^\theta < q \leq x} W(q, x) = \sum_{x-x^\theta < q \leq x} \left( \sum \frac{1}{p-q} + \sum \frac{1}{p-q} \right) \]

\[ \leq \sum_{x-x^\theta < q \leq x} \left( \sum \frac{1}{p} \right) + \left( \sum \sum \frac{1}{p} \right). \]

We treat the first sum by lemma 1 as above. The second sum is smaller than

\[ \sum_{x-x^\theta < p \leq x} U(p, \epsilon, \delta) \]

by observing that \(p-p^\theta \leq x-x^\theta\) and \(p^\epsilon \leq p-p^\epsilon\). This sum is, as above, smaller than \(e^\theta \frac{x^\theta}{\log x} (1+\epsilon+o(1))\), which ends the proof of lemma 3.

**Proof of theorem 2.**

We first choose \(\delta = 1/100\). Let \(T = \pi(x) - \pi(x-x^\theta)\) and \(N\) the number of primes \(q\) verifying \(x-x^\theta < q \leq x\) and \(V(q) + W(q, x) \geq 1+2a\). It follows from lemma 2 that

\[ N(1+2a) + (1+o(1))(T-N) \leq (1+\delta+o(1))T \]

which implies
\[ N \leq (1/2 + o(1))T \]

and then it is possible to find 5 primes \( p_i \), \( 1 \leq i \leq 5 \) between \( x - x^\sigma \) and \( x \) and such that

\[ W(p_1) + W(p_1, x) \leq 1 + 2\sigma. \]

Since \( V(p_1) \geq 1 + o(1) \), this implies \( W(p_1, x) \leq 2\sigma + o(1) \).

We set \( n = \prod_{p \leq x} p \) and \( m = \frac{n}{p_1 p_2 p_3 p_4 p_5} \).

We have:

\[
\hat{h}(n) = \hat{h}(m) + \sum_{i=1}^{5} \left( V(p_i) + W(p_i, x) \right) + \sum_{1 \leq i < j \leq 5} \frac{1}{p_i p_j}
\]

\[
\leq \hat{h}(m) + \sum_{i=1}^{5} \left( V(p_i) + 2W(p_i, x) \right)
\]

(5) \[ \hat{h}(n) \leq \hat{h}(m) + 5 + 20 \sigma + o(1). \]

Further, we use hypothesis (H2) to get four primes \( q_1, \ldots, q_4 \) such that \( x + x^\sigma \leq q_1 \leq x + x^{2\sigma} \) and \( q_2 = q_1 + 2 \), \( q_3 = q_1 + 6 \), \( q_4 = q_1 + 8 \). We set

\[ n' = m q_2 q_3 q_4. \]

Then

\[
\hat{h}(n') = \hat{h}(m) + \sum_{i=1}^{4} \left( \sum_{p \leq x} \frac{1}{q_i - p} \right) - \sum_{15 \leq i \leq 15} \frac{1}{q_i p} + \frac{41}{24}
\]

\[
\leq \hat{h}(m) + 4 \sum_{p \leq x} \frac{1}{q_i - p} + \frac{41}{24} + o(1)
\]

\[
\geq \hat{h}(m) + 4U(x^{2\sigma}, 2\sigma, 1) + \frac{41}{24} + o(1)
\]

\[
= \hat{h}(m) + 4(1 - 2\sigma) + \frac{41}{24} + o(1).
\]

With (5), we obtain:

\[ \hat{h}(n') \geq \hat{h}(n) + \frac{17}{24} - 28\sigma + o(1) \geq \hat{h}(n) \]

for \( x \) large enough. And since
n' \leq n \left(\frac{x+x^{2a}}{x-x^{a}}\right)^{\frac{4}{5}} < n

n cannot be a champion number for \( \hat{h} \).
Let \( x = 41, n = \prod_{p \leq 41} p \). J. Selfridge has observed that

\( \hat{h}(37^4) > \hat{h}(n) \).

But it seems much more difficult to find the smallest \( x \) such that

\( \prod_{p \leq x} p \) is not a champion for \( \hat{h} \).

We shall end this paper with some remarks and problems. It is
well known that the maximal order of \( \omega(n) \) is \( \log n \log \log n \) \( 1+o(1) \). In
\( \text{[Erd} 1] \), it is proved that

\[ \text{Card} \{ n \leq x; \omega(n) \geq c \frac{\log x}{\log \log x} \} \geq x^{1-c+o(1)} \]

for \( 0 < c < 1 \). In \( \text{[Erd} 2] \), it is proved that the maximal order of \( F(n) \)
is \( (1+o(1)) \sqrt{\log n} \). It is interesting to study:

\[ \Omega_c(x) = \text{Card} \{ n \leq x; F(n) \geq c \sqrt{\log x} \} \]

for \( 0 < c < 1 \). For small \( c \), it is easy to get a lower bound for \( \Omega_c(x) \).

We define \( k \) as the largest integer such that

\[ 2^{k(k+1)/2} \leq x \]

and for \( \sqrt{x} \leq i \leq k \), we consider a random prime \( p_i \) belonging to

\( \sigma^1, 2^1 \), where \( \sigma \) is a fixed real number, \( \frac{1}{2} \leq \sigma < 1 \). We set

\[ n = \prod_{\sqrt{x} \leq i \leq k} p_i. \]

Clearly \( n \leq x \) and

\[ F(n) > (k - \sqrt{x} - 2)(1 - \frac{1}{2\sigma}) \geq \frac{2}{\log 2} \frac{1 - \frac{1}{2\sigma} + o(1)}{\sqrt{\log x}}. \]

How many such \( n \)'s do we have?
\[ \prod_{\sqrt{k} \leq n \leq k} (\pi(2^n) - \pi(2^{n-1})) \geq \prod_{\sqrt{k} \leq n \leq k} \gamma \left( \frac{1 - \frac{1}{2^k}}{1} \right)^{2^i} \]

where \( \gamma \) is a fixed constant. An estimation of this last product shows that for \( c < \sqrt{\log 2} \left( 1 - \frac{1}{2^c} \right) \), we have

\[ \psi_c(x) \gtrsim x \exp \left( - \frac{1 + o(1)}{\sqrt{2 \log 2}} \right) \sqrt{\log x \log \log x}. \]

It is possible to improve the above reasoning, and for instance to get a lower bound for \( \psi_c(x) \) for all \( c, 0 < c < 1 \), by using the techniques of \cite{erdos1959distribution}.

As observed by G. Tenenbaum, an upper bound of the same form, but with a different constant, can be obtained: Since \( F(n) \leq \psi(n) \), we have:

\[ \psi_c(x) \lesssim \text{card}\{n \leq x; \psi(n) \geq c\sqrt{\log x}\} \]

\[ \lesssim x^{-c\sqrt{\log x}} \left( \sum_{n \leq x} \psi(n) \right) \]

for all \( z \geq 1 \). The above sum can be evaluated by convolution method, and we get

\[ \psi_c(x) \lesssim x^{-c\sqrt{\log x}} x(\log x)^{z-1}. \]

Choosing \( z = (c\sqrt{\log x})/\log \log x \) gives:

\[ \psi_c(x) \lesssim x \exp(- (c/2 + o(1)) \sqrt{\log x \log \log x}). \]

It is possible to improve slightly the constant \( c/2 \) in the above expression. Using optimization results of \cite{erdos1959distribution} show that if \( \psi(n) \leq c\sqrt{\log n} \), with \( 0 < c < 2 \), then \( F(n) \leq \lambda(c) c^{\frac{1}{2}} \sqrt{\log n} (1 + o(1)) \), where

\[ \lambda(c) = 1 - \frac{1}{2} \exp \left( \frac{2(1 - c^2/4)}{c^2} \right) < 1. \]

So, \( (6) \) is valid with \( \psi_{\lambda(c)}(x) \) instead of \( \psi_c \) on the left hand side.

Let us denote by \( r(n) \) the number of divisors of \( n \), we write the divisors
and we define
\[
g(n) = \sum_{i=1}^{r(n)-1} \frac{d_i}{d_{i+1}} ; \quad G(n) = \sum_{i=1}^{r(n)-1} \left(1 - \frac{d_i}{d_{i+1}}\right)\]
\[
H(n) = \sum_{i=1}^{r(n)-1} \frac{1}{d_{i+1}-d_i} ; \quad H(n) = \sum_{1 \leq i < j \leq r(n)} \frac{1}{d_j - d_i}.
\]
From the obvious inequality
\[
1 - \frac{d_i}{d_{i+1}} \leq \log \left(\frac{d_{i+1}}{d_i}\right)
\]
we easily deduce
\[
(7) \quad r(n) - 1 - \log n \leq g(n) \leq r(n) - 1.
\]
In [Nic], \((r+f)\)-champion numbers were considered when \(f\) is a slowly increasing function. By the same method, it is not difficult to prove that a \(r\)-champion number large enough is a \(g\)-champion, and that if \(n\) is a \(g\)-champion, it is largely composite (i.e., \(m \leq n \Rightarrow r(m) \leq r(n)\)).

In fact, the calculation of \(r\)-champions and \(g\)-champions shows that they exactly coincide from the very beginning up to 6 millions. We do not see how to prove that they coincide up to infinity.

The calculation of \(G\)-champions up to 6 millions shows that all \(r\)-champions are \(G\) champions, and that largely composite numbers look like \(G\)-champions with a few exceptions. For instance 672 is a \(G\)-champion and is not largely composite, and 630 and 650 are largely composite but not \(G\)-champions. We do not see at all how to prove something about that. In fact, (7) tells us that \(G(n) = r(n) - 1 - g(n) \leq \log n\), which is very small comparatively to high values of \(r(n)\).

Computing \(H(n)\) gives 14 values of \(n\), the largest of which is 5040, for which \(H(n) > r(n)\). We conjecture that for \(n > 5040\), we have \(H(n) < r(n)\).

More information about these functions can be found in [Bal], [Erd 5], [Ten], [Vose].

We thank very much G. Tenenbaum, and the referees for several valuable suggestions.

References


