Functiones et Approximatio
XXXIII (2005), 57–72

INTEGERS REPRESENTABLE AS THE SUM OF POWERS OF THEIR PRIME FACTORS

JEAN-MARIE DE KONINCK[1] & FLORIAN LUCA[2]

Abstract: Given an integer $\alpha \geqslant 2$, let S_α be the set of those positive integers n, with at least two distinct prime factors, which can be written as $n = \sum_{p|n} p^\alpha$. We obtain general results concerning the nature of the sets S_α and we also identify all those $n \in S_3$ which have exactly three prime factors. We then consider the set T (resp. T_0) of those positive integers n, with at least two distinct prime factors, which can be written as $n = \sum_{p|n} p^{\alpha_p}$, where the exponents $\alpha_p \geqslant 1$ (resp. $\alpha_p \geqslant 0$) are allowed to vary with each prime factor p. We examine the size of $T(x)$ (resp. $T_0(x)$), the number of positive integers $n \leqslant x$ belonging to T (resp. T_0).
Keywords: Prime factorization

1. Introduction

Identifying all those positive integers n such that

$$n = \sum_{p|n} p^\alpha \tag{1}$$

for some integer $\alpha \geqslant 2$ is certainly a difficult problem. Since prime powers p^α (with $\alpha \geqslant 2$) trivially satisfy (1), we shall examine the set S_α, namely the set of those positive integers n satisfying (1) but which have at least two distinct prime factors.

We first obtain general results concerning the nature of the sets S_α. We then identify all those $n \in S_3$ which have exactly 3 prime factors. We further consider the more general equation

$$n = \sum_{p|n} p^{\alpha_p}, \tag{2}$$

2001 Mathematics Subject Classification: 11A41, 11A25

[1] Research supported in part by a grant from NSERC.
[2] Research supported in part by projects SEP-CONACYT 37259-E and 37260-E.

where the exponents α_p are allowed to vary with each prime factor p. Clearly all prime powers have such a representation (2). So let us define T (resp. T_0) as the set of all positive integers n having a representation (2) with each $\alpha_p \geqslant 1$ (resp. $\alpha_p \geqslant 0$) but with at least two distinct prime divisors. We obtain a non trivial upper bound for the number $T_0(x)$ of positive integers $n \leqslant x$ belonging to T_0.

Finally, we give a heuristic argument yielding lower and upper estimates for $T(x)$, the number of positive integers $n \leqslant x$ belonging to T.

2. General observations

For each integer $n \geqslant 2$, let $\omega(n)$ stand for the number of distinct prime factors of n and let $P(n)$ stand for the largest prime factor of n. We first make the following observations. Given $\alpha \geqslant 2$ and $n \in S_\alpha$, we have:

(i) $P(n) < n^{1/\alpha}$.

(ii) Letting $r = \omega(n)$, then $r \geqslant 3$ and r is odd; this is easily established by considering separately the cases "n odd" and "n even".

(iii) If α is even, then $\omega(n)$ cannot be a multiple of 3; one can see this by considering separately the cases "$3|n$" and "$3 \nmid n$".

(iv) If $\omega(n) = \alpha$, then n cannot be squarefree, since otherwise, comparing the arithmetic mean with the geometric mean of the prime factors of n, we get

$$n = q_1 q_2 \dots q_\alpha = q_1^\alpha + q_2^\alpha + \dots + q_\alpha^\alpha \geqslant \alpha q_1 q_2 \dots q_\alpha = \alpha n,$$

a contradiction, since $\alpha \geqslant 2$.

(v) If $n \in S_2$, then, in view of (ii) and (iii), $r := \omega(n)$ is odd, $r \geqslant 5$; moreover:
  * if $r = 5$, then $n \equiv 5$ or $8 \pmod{24}$,
  * if $r = 7$, then $n \equiv 7,\ 10,\ 15$ or $18 \pmod{24}$,
  * otherwise $r \geqslant 11$.

(vi) A computer search shows that S_3 contains at least 6 elements, namely:

$$\begin{aligned}
378 &= 2 \cdot 3^3 \cdot 7 = 2^3 + 3^3 + 7^3,\\
2548 &= 2^2 \cdot 7^2 \cdot 13 = 2^3 + 7^3 + 13^3,\\
2\,836\,295 &= 5 \cdot 7 \cdot 11 \cdot 53 \cdot 139 = 5^3 + 7^3 + 11^3 + 53^3 + 139^3,\\
4\,473\,671\,462 &= 2 \cdot 13 \cdot 179 \cdot 593 \cdot 1621 = 2^3 + 13^3 + 179^3 + 593^3 + 1621^3,\\
23\,040\,925\,705 &= 5 \cdot 7 \cdot 167 \cdot 1453 \cdot 2713 = 5^3 + 7^3 + 167^3 + 1453^3 + 2713^3,\\
21\,467\,102\,506\,955 &= 5 \cdot 7^3 \cdot 313 \cdot 1439 \cdot 27791 = 5^3 + 7^3 + 313^3 + 1439^3 + 27791^3.
\end{aligned}$$

(vii) If $n \in S_4$, then $\omega(n) = 7$ or $\omega(n) \geqslant 11$. To show this, first let $r = \omega(n)$. We know from (ii) that $r \geqslant 3$ and odd; but from (iii), it follows that $r \neq 3$; hence, $r \geqslant 5$. But $r \neq 5$; indeed, if $r = 5$, then first assume that $5|n$; in this case, since $p^4 \equiv 1 \pmod 5$ for all primes $p \neq 5$,

$$n = 625 + q_2^4 + q_3^4 + q_4^4 + q_5^4 \equiv 0 + 4 = 4 \pmod 5,$$

which contradicts $5|n$; on the other hand, if n is not a multiple of 5, then $n \equiv 5 \pmod 5$, again a contradiction. Hence, $r \geqslant 7$. Finally, in view of (iii), $r \neq 9$. Hence, we may conclude that $r = 7$ or $r \geqslant 11$.

(vii) It is not known if T is an infinite set. However, if there exist infinitely many primes p of the form $p = \dfrac{2^k + 3^\ell}{5}$, then $\#T = +\infty$, the reason being that in this case, we have $2 \cdot 3 \cdot p = 2^k + 3^\ell + p$.

(viii) Using a parity argument, it is clear that any number $n \in T$ has an odd number of distinct prime divisors. One can check that the smallest element of T is 30; in fact, 30 has two representations of type (2), namely

$$30 = 2 \cdot 3 \cdot 5 = 2 + 3 + 5^2 = 2^4 + 3^2 + 5.$$

Letting $T(x) := \#\{n \leqslant x : n \in T\}$, a computer search shows that $T(100) = 6$, $T(10^3) = 42$, $T(10^4) = 109$, $T(10^5) = 321$ and $T(10^6) = 973$. On the other hand, the smallest odd element of T is 915, in which case we have

$$915 = 3 \cdot 5 \cdot 61 = 3^6 + 5^3 + 61.$$

3. Identifying those $n \in S_3$ with $\omega(n) = 3$

Theorem 1. *If $n \in S_3$ and $\omega(n) = 3$, then $n = 2 \cdot 3^3 \cdot 7$ or $n = 2^2 \cdot 7^2 \cdot 13$.*

Proof. We prove this in 9 steps.

1. Write $x < y < z$ for the three distinct prime factors of n. Note that the given relation forces $z|y^3 + x^3$, so that $z|y + x$ or $z|y^2 - yx + x^2$, and similarly $y|z + x$, or $y|z^2 - zx + x^2$, and $x|z + y$, or $x|z^2 - zy + y^2$.

2. Assume $z|y+x$. Since $y + x < 2y < 2z$, this is possible only when $z = y + x$. If $x > 2$, then $y+x$ is even, and so it cannot be an odd prime. Thus, $x = 2$, $z = y+2$, but then

$$x^3 + y^3 + z^3 = 8 + y^3 + (y + 2)^3 \equiv 16 \pmod y,$$

which is impossible. Thus, $z \nmid y + x$, and $z|y^2 - yx + x^2$. Since $z > 3$, we also conclude that $z \equiv 1 \pmod 3$, because the relation $y^2 - yx + x^2 \equiv 0 \pmod z$ implies that $(2y - x)^2 \equiv -3x^2 \pmod z$, which means that $\left(\dfrac{-3}{z}\right) = 1$, which is equivalent to the fact that $z \equiv 1 \pmod 3$. Here, and in what follows, for an odd prime p and an integer a we use $\left(\dfrac{a}{p}\right)$ for the Legendre symbol of a in respect to p.

3. Assume that $z^2|n$. In this case, we then get $z^2|y^3 + x^3$, and by the previous arguments, it follows that $z^2|y^2 - yx + x^2$. This is impossible because $y^2 - yx + x^2 = y^2 - x(y - x) < y^2 < z^2$. Thus, $z\|n$.

4. Assume that $y|z+x$. Write $z := \lambda y - x$, with some positive integer λ. Clearly $\lambda \geqslant 2$. We then get $x \equiv \lambda y \pmod z$. Since we also have $y^2 - yx + x^2 \equiv 0 \pmod z$, we get $y^2 - y(\lambda y) + (\lambda y)^2 \equiv 0 \pmod z$. Thus, $z|y^2(1-\lambda+\lambda^2)$, and therefore $z|1-\lambda+\lambda^2$. If $\lambda = 2$, we get $z|1-2+2^2=3$, which is impossible. If $\lambda = 3$, we get $z|1-3+3^2=7$. Thus, $z=7$, and therefore $7 = 3y - x$. Since y is odd, we get $x = 2$ and therefore $y = 3$, which does give the solution

$$2^3 + 3^3 + 7^3 = 2 \cdot 3^3 \cdot 7$$

mentioned in the statement of our theorem.

Assume now that $\lambda \geqslant 4$. Then,

$$z = \lambda y - x > (\lambda - 1)y = \lambda y \cdot \frac{\lambda - 1}{\lambda} \geqslant \frac{3\lambda y}{4}.$$

Since $z|1-\lambda+\lambda^2$, we also get

$$\lambda^2 > 1 - \lambda + \lambda^2 \geqslant z > \frac{3\lambda y}{4},$$

and therefore that

$$\lambda > \frac{3y}{4}.$$

Thus,

$$z > \frac{3\lambda y}{4} > \frac{9y^2}{16}.$$

Since we also have $z|y^2 - yx + x^2$, we get that

$$\delta = \frac{y^2 - yx + x^2}{z}$$

is a positive integer. However,

$$\delta < \frac{y^2}{z} < \frac{16}{9} < 2,$$

therefore $\delta = 1$, and so

$$z = y^2 - yx + x^2.$$

Thus,

$$n = x^3 + y^3 + z^3 = (y+x)(y^2 - yx + x^2) + z^3 = z(y+x) + z^3,$$

therefore

$$\frac{n}{z} = y + x + z^2.$$

Looking at this last relation modulo y, we get $x+z^2 \equiv 0 \pmod{y}$. Since $y|x+z$, we also get $z \equiv -x \pmod{y}$ and therefore $z^2 \equiv x^2 \pmod{y}$. Thus, $x^2+x \equiv 0 \pmod{y}$; hence, $y|x(x+1)$. This is possible only when $y = x+1$ and $x = 2$. Thus, $x=2$, $y=3$, $z = 3^2 - 2\cdot 3 + 2^2 = 7$, so that $\lambda = 3$, contradicting the fact that $\lambda \geqslant 4$.

5. From now on, we may assume that $y \nmid z+x$ and therefore that $y|z^2-zx+x^2$. If $y=3$, then $x=2$, in which case $z|2^3+3^3=35$; hence, $z=7$ (because $z \equiv 1 \pmod{3}$), which is a case already treated. Thus, we may assume that $y>3$, and since $y|z^2-zx+x^2$, an argument similar to the one employed at step 2 tells us that $y \equiv 1 \pmod{3}$.

6. Here, we observe that $x \equiv 2 \pmod{3}$. Indeed, for if not, we must either have $x=3$, which is impossible because then $3|n$, but $x^3+y^3+z^3 \equiv 2 \pmod{3}$, or $x \equiv 1 \pmod{3}$, therefore $3 \nmid n$, while $x^3+y^3+z^3 \equiv 0 \pmod{3}$.

7. Write $n := x^\alpha y^\beta z$. Since we already know that $x \equiv 2 \pmod{3}$ and $y \equiv z \equiv 1 \pmod{3}$, we reduce the relation

$$x^3+y^3+z^3 = x^\alpha y^\beta z$$

modulo 3 to get $1 \equiv 2^\alpha \pmod{3}$. This shows that α is even.

8.1. Assume that $x=2$. We first show that $\alpha=2$. Indeed, for if not, we would first get $8 \mid y^3+z^3$ and hence that $8|(z+y)(z^2-zy+y^2)$. Since z^2-zy+y^2 is odd, we get $8|y+z$. Thus, $(y,z) \in \{(1,7),(7,1),(3,5),(5,3)\} \pmod{8}$.
We know that $z|y^3+2^3$, and $y|z^3+2^3$. In particular, $-2y \equiv (4/y)^2 \pmod{z}$, and so

$$\left(\frac{-2y}{z}\right) = 1,$$

and in a similar way one deduces that

$$\left(\frac{-2z}{y}\right) = 1.$$

Hence, we have

$$1 = \left(\frac{-1}{z}\right)\left(\frac{-1}{y}\right)\left(\frac{2}{y}\right)\left(\frac{2}{z}\right)\left(\frac{y}{z}\right)\left(\frac{z}{y}\right) = (-1)^{\left(\frac{z-1}{2}+\frac{y-1}{2}\right)+\left(\frac{z^2-1}{8}+\frac{y^2-1}{8}\right)+\left(\frac{(y-1)(z-1)}{4}\right)}$$
$$= (-1)^{1+0+0} = -1,$$

a contradiction. Therefore, $\alpha = 2$.

8.2. Here, we show that $\beta \in \{2,3\}$. If $\beta = 1$, we get

$$4yz = 2^3+y^3+z^3 > 3(2\cdot y\cdot z) = 6yz,$$

which is impossible, the above inequality following from the AGM-inequality. Using now the fact that $z|y^2-yx+x^2$ (see step 2), together with the fact that $y^2-xy+x^2 = y^2-x(y-x) < y^2$, we learn that $z<y^2$. Since

$$3z^3 > x^3+y^3+z^3 = 4y^\beta z,$$

we get

$$y^\beta < \frac{3z^2}{4} < \frac{3y^4}{4} < y^4,$$

and therefore that $\beta < 4$.

8.3. Assume that $\beta = 3$. Rewrite the equation

$$8 + y^3 + z^3 = 4y^3 z$$

as

$$y^3 = \frac{z^3+8}{4z-1}.$$

Let $D := 4z - 1$. Thus $z \equiv 4^{-1} \pmod D$. Since we also have $z^3 + 8 \equiv 0 \pmod D$, we get $4^{-3} + 8 \equiv 0 \pmod D$ and therefore that $D | 1 + 8 \cdot 4^3 = 513 = 3^3 \cdot 19$. Thus, $D \in \{1, 3, 3^2, 3^3, 19, 3 \cdot 19, 3^2 \cdot 19, 3^3 \cdot 19\}$. Since z must be at least the second prime number which is congruent to 1 modulo 3, we have that $D \geqslant 4 \cdot 13 - 1 = 51$, and since we also have that $D \equiv -1 \pmod 4$, it follows that in fact only the instance $D = 3^2 \cdot 19$ is possible. Therefore $z = \dfrac{D+1}{4} = \dfrac{3^2 \cdot 19 + 1}{4} = 43$. However, for this value of z, the number $\dfrac{z^3+8}{4z-1} = \dfrac{43^3+8}{4 \cdot 43 - 1} = 465$ is not the cube of a prime number.

8.4. Assume that $\beta = 2$. In this case,

$$z^3 < x^3 + y^3 + z^3 = 4y^2 z,$$

so that

$$z^2 < 4y^2,$$

which implies that $z < 2y$. But we also have that $y^2 | (x^3 + z^3)$, and since y does not divide $x + z$, it follows that $y^2 | z^2 - zx + x^2 = z^2 - 2z + 4$. Since $z \equiv 1 \pmod 3$, we also have that $3 | z^2 - 2z + 4$, and since $y > 3$, we have that $y^2 | (z^2 - 2z + 4)/3$. Now write

$$y^2 = \frac{z^2 - 2z + 4}{3\delta},$$

where δ is a positive integer. We then get

$$\delta = \frac{z^2 - 2z + 4}{3y^2} < \frac{z^2}{3y^2} < \frac{4y^2}{3y^2} = \frac{4}{3} < 2,$$

which means that $\delta = 1$. Thus, $3y^2 = z^2 - 2z + 4$. The original relation becomes

$$4y^2 z = 8 + y^3 + z^3 = y^3 + (z+2)(z^2 - 2z + 4) = y^3 + 3y^2(z+2),$$

so that

$$4z = y + 3(z+2) = 3z + y + 6,$$

which implies that $z = y+6$. Thus, $y \equiv -6 \pmod{z}$, and since $z \mid y^2 - yx + x^2 = y^2 - 2y + 4$, we get $z \mid (-6)^2 - 2(-6) + 4 = 52 = 4 \cdot 13$. Thus, $z = 13$, $y = z - 6 = 7$, and we have obtained the solution

$$2^3 + 7^3 + 13^3 = 2^2 \cdot 7^2 \cdot 13$$

mentioned in the statement of our theorem.

9. From now on, we assume that $x > 2$. The relation $x|y^3 + z^3$ implies that $y^3 \equiv -z^3 \pmod{x}$ and therefore $-yz \equiv \left(z^2/y\right)^2 \pmod{x}$, and so

$$\left(\frac{-yz}{x}\right) = 1. \tag{3}$$

In a similar way, using the facts that $y|x^3 + z^3$ and $z|x^3 + y^3$, one gets

$$\left(\frac{-xz}{y}\right) = \left(\frac{-xy}{z}\right) = 1.$$

Thus,

$$1 = \left(\frac{-yz}{x}\right) = \left(\frac{-1}{x}\right) \cdot \left(\frac{y}{x}\right) \cdot \left(\frac{z}{x}\right) = (-1)^{\frac{x-1}{2}} \cdot \left(\frac{y}{x}\right) \cdot \left(\frac{z}{x}\right),$$

and similarly

$$1 = (-1)^{\frac{y-1}{2}} \cdot \left(\frac{x}{y}\right) \cdot \left(\frac{z}{y}\right),$$

and

$$1 = (-1)^{\frac{z-1}{2}} \left(\frac{x}{z}\right) \cdot \left(\frac{y}{z}\right).$$

Write $a := \dfrac{x-1}{2}$, $b := \dfrac{y-1}{2}$, $c := \dfrac{z-1}{2}$. Multiplying the three relations above side by side and using quadratic reciprocity we get

$$1 = (-1)^{a+b+c+ab+ac+bc},$$

which means that

$$S := a + b + c + ab + ac + bc$$

must be an even number. Let us notice that it is not possible that all three numbers a, b, c are even. Indeed, if this were so, then $x \equiv y \equiv z \equiv 1 \pmod{4}$, and reducing the equation

$$x^3 + y^3 + z^3 = n$$

modulo 4, we would get $3 \equiv 1 \pmod{4}$, which is impossible. Thus, at least one of the numbers a, b, c is odd. This, together with the fact that S is even implies that all three numbers a, b, c are odd, therefore $x \equiv y \equiv z \equiv 3 \pmod{4}$. We reduce now the relation

$$x^3 + y^3 + z^3 = x^\alpha y^\beta z$$

modulo 4, and since α is even (see step 7), we get $1 \equiv 3^{\beta+1} \pmod 4$ and therefore that β is odd. Thus, we may write our original equation as

$$x^3 + y^3 + z^3 = m^2yz, \tag{4}$$

where $m := x^{\alpha/2}y^{(\beta-1)/2}$ is an integer. Write $x + y = 2\ell$. Notice that since $x \equiv y \equiv 3 \pmod 4$, we have that ℓ is an odd number. Let p be an arbitrary prime divisor of ℓ. Reducing the above equation mod p, we get $z^3 \equiv m^2yz \pmod p$, therefore $y \equiv \left(z/m\right)^2 \pmod p$. Thus,

$$\left(\frac{y}{p}\right) = 1.$$

Since $y \equiv -x \pmod p$, we get that

$$1 = \left(\frac{y}{p}\right) = \left(\frac{-x}{p}\right) = \left(\frac{-1}{p}\right) \cdot \left(\frac{x}{p}\right) = (-1)^{\frac{p-1}{2}} \cdot (-1)^{\frac{x-1}{2}\cdot\frac{p-1}{2}} \cdot \left(\frac{p}{x}\right) = \left(\frac{p}{x}\right),$$

where in the above computation we used the quadratic reciprocity law together with the fact that $x \equiv 3 \pmod 4$. Since the above formula holds for all prime divisors p of ℓ, we get, by multiplying all these relations, that

$$1 = \left(\frac{\ell}{x}\right) = \left(\frac{(y+x)/2}{x}\right) = \left(\frac{4}{x}\right) \cdot \left(\frac{(y+x)/2}{x}\right) = \left(\frac{2y+2x}{x}\right) = \left(\frac{2y}{x}\right).$$

In the above argument, we used only equation (4) (which is symmetric in y and z), together with the fact that $x \equiv y \equiv z \equiv 3 \pmod 4$ (which is also symmetric in y and z), but we did not use size arguments (i.e. the fact that $y < z$). Thus, an identical argument can be carried through to show that

$$\left(\frac{2z}{x}\right) = 1.$$

Multiplying these last two relations we get

$$1 = \left(\frac{2y}{x}\right) \cdot \left(\frac{2z}{x}\right) = \left(\frac{4}{x}\right) \cdot \left(\frac{yz}{x}\right) = \left(\frac{yz}{x}\right),$$

which together with the fact that

$$\left(\frac{-yz}{x}\right) = 1$$

(see equation (3)), implies that

$$\left(\frac{-1}{x}\right) = 1,$$

contradicting the fact that $x \equiv 3 \pmod 4$.

This completes the proof of Theorem 1.

4. An upper bound for $T_0(x)$

Theorem 2. *As* $x \to \infty$, *we have*

$$T_0(x) \leqslant x \exp\left\{-(1+o(1))\sqrt{\frac{1}{6}\log x \log\log x}\right\}.$$

Proof. First recall the estimate

$$\Psi(x,y) := \#\{n \leqslant x : P(n) \leqslant y\} \ll x \exp\{-(1+o(1))u \log u\}, \qquad (5)$$

where $u = \log x / \log y$ (see for instance Tenenbaum [4]). Now let

$$y = \exp\left\{\sqrt{\frac{3}{2}\log x \log\log x}\right\} \qquad (6)$$

and set

$$u = \frac{\log x}{\log y} = \sqrt{\frac{2}{3}\frac{\log x}{\log\log x}} \quad \text{so that} \quad u \log u = (1+o(1))\sqrt{\frac{1}{6}\log x \log\log x}. \qquad (7)$$

It follows from (5), (6) and (7) that

$$\begin{aligned}\#\{n \leqslant x : n \in T_0,\ P(n) \leqslant y\} &\ll x \exp\{-(1+o(1))u \log u\} \qquad (8)\\ &\ll x \exp\left\{-(1+o(1))\sqrt{\frac{1}{6}\log x \log\log x}\right\}.\end{aligned}$$

We shall therefore assume from now on that $P(n) > y$.

Let x be a large number with the corresponding y and u defined by (6) and (7). Then, using Stirling's formula, as well as the fact that

$$\sum_{p \leqslant y} \frac{1}{p} = \log\log y + O(1)$$

holds as y tends to infinity, we get

$$\begin{aligned}\#\{n \leqslant x : \omega(n) \geqslant u\} &\leqslant \sum_{p_1 \ldots p_{\lfloor u \rfloor} \leqslant x} \frac{x}{p_1 \ldots p_{\lfloor u \rfloor}} \leqslant \frac{x}{\lfloor u \rfloor !}\left(\sum_{p \leqslant x} \frac{1}{p}\right)^{\lfloor u \rfloor} \qquad (9)\\ &\leqslant x \left(\frac{e \log\log x + O(1)}{\lfloor u \rfloor}\right)^{\lfloor u \rfloor}\\ &\leqslant x \exp\{-(1+o(1))u \log u\}\\ &\ll x \exp\left\{-(1+o(1))\sqrt{\frac{1}{6}\log x \log\log x}\right\}.\end{aligned}$$

Hence, from here on we may assume that $\omega(n) < u$.

We now neglect those integers $n \leqslant x$, $n \in T_0$ with $P(n) > y$ and such that $P(n)^2 | n$, since the number of such integers is

$$\ll \#\{n \leqslant x : P(n) > y,\ P(n)^2 | n\} \leqslant \sum_{p>y} \frac{x}{p^2} \tag{10}$$
$$\ll \frac{x}{y} = x \exp\left\{-\sqrt{\frac{3}{2}\log x \log\log x}\right\}.$$

From here on, we shall therefore assume that $Q := P(n) \| n$ and write $n = mQ$. Now, writing (2) as

$$n = mQ = p_1^{b_1} + \ldots + p_k^{b_k}, \tag{11}$$

where $p_1 < \ldots < p_k = Q$ are the prime factors of n and each b_i is non negative, we get from (11) that

$$p_1^{b_1} + \ldots + p_{k-1}^{b_{k-1}} + \delta \equiv 0 \pmod{Q}, \tag{12}$$

where δ is 0 or 1, depending if $b_k > 0$ or $b_k = 0$. The number appearing on the left hand side of (12) depends only on the prime factors of m and does not depend on Q, and moreover, each one of these numbers has at most $\log x$ factors. Thus, we may fix $m \leqslant x/y$ and count how many candidates there may be for a given prime number Q. Since n is not a prime power, we have $k \geqslant 2$, and therefore the left hand side of congruence (12) is a positive integer. Since $p_i^{b_i} < n \leqslant x$, it follows that $b_i \ll \log x + 1$. In fact, $b_i < \log x + 1$ always holds except when $i = 1$ and $p_1 = 2$, in which case $b_1 \leqslant \frac{\log x + 1}{\log 2}$. Thus, the total number of integers which can appear on the left hand side of (12) is $\ll (\log x + 1)^{\omega(n)} \ll (\log x + O(1))^u \ll \exp\{(1 + o(1))u \log\log x\}$, which means that

$$\#\{n \leqslant x : n \in T_0, P(n) > y, P(n) \| n, \omega(n) < u\} \tag{13}$$
$$\ll \frac{x \log x}{y} \exp\{(1 + o(1))u \log\log x\}$$
$$\ll x \exp\left\{-(1 + o(1))\sqrt{\frac{1}{6}\log x \log\log x}\right\}.$$

Theorem 2 then follows from (8), (9), (10) and (13).

5. Empirical lower and upper bounds for $T(x)$

Although we cannot prove that T is an infinite set, a heuristic argument shows that

$$\exp\left(\frac{2}{e}(1 + o(1))\frac{\log x}{(\log\log x)^2}\right) \leqslant T(x) \leqslant x^{1/2+o(1)}. \tag{14}$$

Our argument goes as follows. First, we will show that, heuristically,

$$T(x)=\frac{1}{2}\sum_{n\leqslant x} f(n), \qquad \text{where} \quad f(n):=\frac{1}{n}\prod_{p|n}\left\lfloor\frac{\log n}{\log p}\right\rfloor, \tag{15}$$

from which we will show that (14) follows.

Indeed, given a positive integer n such that $\omega(n)$ is odd and writing $n=q_1^{a_1}\dots q_r^{a_r}$, then in order to have $n\in T$, we must find a representation of the form

$$n=q_1^{\alpha_1}+\dots+q_r^{\alpha_r}. \tag{16}$$

Now, for each exponent α_i, there are $\lfloor\log n/\log q_i\rfloor$ possible choices. Hence, if a representation of the form (16) is possible, then the exponents α_i have been chosen in the interval $[1,\lfloor\log n/\log q_i\rfloor]$. Therefore, since there are $\prod_{i=1}^{r}\lfloor\log n/\log q_i\rfloor$ possible choices for the right hand side of (16), we should 'expect' that a representation of the form (16) will be possible with a 'probability' equal to $\frac{1}{n}\prod_{p|n}\left\lfloor\frac{\log n}{\log p}\right\rfloor$, thus establishing (15); note that the factor $\frac{1}{2}$ comes from the fact that a randomly chosen number has an odd "$\omega(n)$" with a probability $\frac{1}{2}$.

It remains to prove that (14) follows from (15).

First we prove the lower bound. Let x be a large positive real number and let $k\geqslant 1$ be an integer.

Let $p_1<\dots<p_k$ be the first k primes. We shall consider only the contribution to $T(x)$ of those positive integers $n=p_1\dots p_kp\leqslant x$, where $p>p_k$ is a prime number. We first get rid of the integer parts. Clearly, if $i\in\{1,\dots,k\}$, then

$$\left\lfloor\frac{\log n}{\log p_i}\right\rfloor=\frac{\log n}{\log p_i}-\left\{\frac{\log n}{\log p_i}\right\}>\frac{\log n}{\log p_i}\left(1-\frac{\log p_i}{\log n}\right)>\frac{\log n}{\log p_i}\exp\left(-2\frac{\log p_i}{\log n}\right),$$

where in the above inequalities we used the fact that $\log p_i/\log n\leqslant 1/2$ and that the inequality $1-t>\exp(-2t)$ holds for $t\in(0,1/2)$. Together with the fact that $\lfloor\log n/\log p\rfloor\geqslant 1$, we get

$$f(n)\geqslant\left(\prod_{i=1}^{k}\frac{\log n}{\log p_i}\right)\exp\left(-2\sum_{i=1}^{k}\frac{\log p_i}{\log n}\right)>\exp(-2)\prod_{i=1}^{k}\frac{\log n}{\log p_i}\gg\frac{(\log p)^k}{\log p_1\dots\log p_k}.$$

This implies that

$$\begin{aligned} T(x)=\frac{1}{2}\sum_{n\leqslant x}f(n) &\gg \sum_{\substack{p_1\cdots p_kp\leqslant x\\ p>p_k}}\frac{1}{p_1\log p_1\dots p_k\log p_k}\frac{(\log p)^k}{p} \\ &=\frac{1}{p_1\log p_1\dots p_k\log p_k}\sum_{p_k<p\leqslant x/p_1\dots p_k}\frac{(\log p)^k}{p} \end{aligned}\tag{17}$$

$$\begin{aligned}
&\gg \frac{1}{p_1 \log p_1 \ldots p_k \log p_k} \int_{p_k}^{x/p_1 \ldots p_k} \frac{(\log t)^{k-1}}{t}\, dt \\
&\gg \frac{1}{p_1 \log p_1 \ldots p_k \log p_k} \frac{\left(\log(x/p_1 \ldots p_k)\right)^k}{k} \\
&= \exp\left(k \log\log x - \sum_{i=1}^{k} (\log p_i + \log\log p_i) + O\left(\frac{k(\log p_1 + \ldots + \log p_k)}{\log x}\right)\right).
\end{aligned}$$

The above chain of inequalities holds when k is such that

$$\log(x/p_1 \ldots p_k) - \log p_k \gg \log(x/p_1 \ldots p_k),$$

which in turn is true when

$$\log p_k + \frac{\log p_1 + \ldots + \log p_k}{\log x} = o(\log x),$$

which holds when

$$\log p_k + \frac{k \log p_k}{\log x} = o(\log x). \tag{18}$$

We now use the fact that, as k tends to infinity,

$$p_k \leqslant k \log k + k \log\log k - k + o(k)$$

(see Théorème A (v) in [1]), together with the well known estimate

$$\sum_{p \leqslant y} \log p = \sum_{n \leqslant y} \Lambda(n) + O(y^{1/2}) = y + O\left(\frac{y}{\exp(c\sqrt{\log y})}\right) = y + O\left(\frac{y}{(\log y)^2}\right),$$

where c is some positive constant and Λ denotes the von Mangoldt function, to conclude that

$$\sum_{i=1}^{k} \log p_i = p_k + O\left(\frac{p_k}{(\log k)^2}\right) \leqslant k \log k + k \log\log k - k + o(k). \tag{19}$$

Since $p_k < 2k \log k$ holds for all sufficiently large k, we also have that

$$\begin{aligned}
\sum_{i=1}^{k} \log\log p_i &\leqslant k \log\log p_k \leqslant k \log\left(\log k + \log(2 \log k)\right) \qquad (20) \\
&\leqslant k \log\log k + O\left(\frac{k \log\log k}{\log k}\right) \\
&= k \log\log k + o(k).
\end{aligned}$$

Introducing inequalities (19) and (20) into (17), we get

$$\begin{aligned} T(x) &\geqslant \exp\left(k\log\log x - k\log k - 2k\log\log k + k + o(k) + O\left(\frac{k^2\log k}{\log x}\right)\right) \quad (21)\\ &= \exp\left(k\log\left(\frac{\log x}{k(\log k)^2}\right) + k + o(k) + O\left(\frac{k^2\log k}{\log x}\right)\right). \end{aligned}$$

In order to maximize the main term of the above inequality, we should choose k versus x in such a way that the expression $k\log\left(\frac{\log x}{k(\log k)^2}\right)$ should be as large as possible. Thus, we choose $k := \left\lfloor \frac{1}{e}\frac{\log x}{(\log\log x)^2}\right\rfloor$. We note that k is in the acceptable range; i.e., $p_1 \dots p_k < x$, that condition (18) is fulfilled, that with this choice of k we have

$$k\log\left(\frac{\log x}{k(\log k)^2}\right) = (1+o(1))k,$$

and that the error term is

$$\frac{k^2\log k}{\log x} = \frac{k}{\log k}\frac{k(\log k)^2}{\log x} = O\left(\frac{k}{\log k}\right) = o(k).$$

Hence, we may replace (21) by

$$T(x) \geqslant \exp\left(2(1+o(1))k\right) = \exp\left(\frac{2}{e}(1+o(1))\frac{\log x}{(\log\log x)^2}\right),$$

which proves the left hand side of inequality (14).

We now prove the upper bound.

Fix a large number k and write

$$T(x) < \sum_{\substack{n\leqslant x\\ \omega(n)<k}} \frac{1}{n}\prod_{p\mid n}\frac{\log n}{\log p} + \sum_{\substack{n\leqslant x\\ \omega(n)\geqslant k}} \frac{1}{n}\prod_{p\mid n}\frac{\log n}{\log p} = T_1(x) + T_2(x), \qquad (22)$$

say. We have

$$T_1(x) \leqslant \sum_{\substack{n\leqslant x\\ \omega(n)<k}} \frac{1}{n}(\log n)^{\omega(n)} \leqslant \sum_{n\leqslant x}\frac{(\log n)^k}{n} \ll \frac{(\log x)^{k+1}}{k+1}. \qquad (23)$$

In particular,

$$T_1(x) < (\log x)^{k+1} \qquad (24)$$

holds if k is sufficiently large.

In the sequel, we shall be using the fact that, if k is sufficiently large, then

$$\prod_{i=1}^{k} \log p_i > (\log k)^k. \tag{25}$$

Indeed, since $p_i \geqslant i \log i$ holds for all $i \geqslant 2$ (see [3]), one gets

$$\log p_i \geqslant \log i + \log\log i \qquad (i \geqslant 3). \tag{26}$$

The inequality $\log(1+t) > t/2$ holds for all $t \in (0, 1/2)$. The function $t \longmapsto \log\log t/\log t$ is decreasing for $t > e^e$ and its value at e^e is $1/e < 1/2$. Hence,

$$\log(\log i + \log\log i) = \log\log i + \log\left(1 + \frac{\log\log i}{\log i}\right) > \log\log i + \frac{1}{2}\frac{\log\log i}{\log i}$$

$(i > e^e \approx 15.2)$.

We thus get

$$\begin{aligned}
\sum_{i=1}^{k} \log\log p_i &\geqslant \sum_{i=16}^{k} \log\left(\log i + \log\log i\right) + O(1)\\
&= \sum_{i=16}^{k} \log\log i + \sum_{i=16}^{k} \log\left(1 + \frac{\log\log i}{\log i}\right) + O(1)\\
&\geqslant \sum_{i=16}^{k} \log\log i + \frac{1}{2}\sum_{i=16}^{k} \frac{\log\log i}{\log i} + O(1)\\
&\geqslant \int_{16}^{k} \log\log t dt + \frac{1}{2}\int_{16}^{t} \frac{\log\log t}{\log t} dt + O(1)\\
&= t\log\log t\Big|_{t=16}^{t=k} - \int_{16}^{t} \frac{1}{\log t} dt + \frac{1}{2}\int_{16}^{t} \frac{\log\log t}{\log t} dt + O(1)\\
&> k\log\log k,
\end{aligned}$$

where the last inequality follows for large enough k due to the fact that the function $\int_{16}^{k}\left(\frac{1}{2}\frac{\log\log t}{\log t} - \frac{1}{\log t}\right) dt$ tends to infinity with k, thus establishing (25).

Using (25), we have

$$T_2(x) \leqslant \sum_{\substack{n \leqslant x\\ \omega(n) \geqslant k}} \frac{1}{n}\frac{(\log n)^k}{\prod_{i=1}^{\omega(n)} \log p_i} \leqslant \sum_{\substack{n \leqslant x\\ \omega(n) \geqslant k}} \frac{1}{n}\left(\frac{\log n}{\log \omega(n)}\right)^{\omega(n)}. \tag{27}$$

Using the fact that

$$\omega(n) \leqslant \frac{\log n}{\log\log n} + (1 + o(1))\frac{\log n}{(\log\log n)^2}$$

(see Pomerance [2]), together with the fact that the function $t \longmapsto \left(\frac{\log n}{\log t}\right)^t$ is increasing for $t \leqslant \log n$, it follows, from (27), that

$$\begin{aligned} T_2(x) &\leqslant \sum_{\substack{n \leqslant x \\ \omega(n) \geqslant k}} \frac{1}{n} \cdot n \cdot e^{O\left(\frac{\log n \cdot \log\log\log n}{\log\log n}\right)} \\ &\ll \mathcal{N}_k(x) \exp\left\{O\left(\frac{\log x \log\log\log x}{\log\log x}\right)\right\}, \end{aligned} \tag{28}$$

where

$$\mathcal{N}_k(x) = \#\{n \leqslant x \mid \omega(n) \geqslant k\}.$$

It is easy to see, using Stirling's formula, that

$$\mathcal{N}_k(x) \leqslant x \sum_{\substack{q_1 < \dots < q_k \\ q_1 \cdots q_k \leqslant x}} \frac{1}{q_1 \dots q_k} \leqslant \frac{x}{k!}\left(\sum_{q \leqslant x} \frac{1}{q}\right)^k \ll \frac{x}{\sqrt{k}}\left(\frac{e \log\log x + O(1)}{k}\right)^k. \tag{29}$$

In particular, combining (28) and (29), for large x and k, we have that

$$T_2(x) < x \cdot \left(\frac{(\log\log x)^{3/2}}{k}\right)^k \exp\left\{O\left(\frac{\log x \log\log\log x}{\log\log x}\right)\right\}. \tag{30}$$

We now choose k such that $k := \left\lfloor \frac{1}{2} \frac{\log x}{\log\log x} \right\rfloor$. It is clear that k is in the acceptable range; i.e., $k = \omega(n)$ for some $n \leqslant x$. Furthermore, inequality (24) shows that

$$T_1(x) < x^{1/2+o(1)}, \tag{31}$$

while inequality (30) shows that

$$\begin{aligned} T_2(x) &< x \exp\left(\frac{3}{2} k \log\log\log x - k \log k - O\left(\frac{\log x \log\log\log x}{\log\log x}\right)\right) \\ &= x \exp\left(-\frac{\log x}{2} + O\left(\frac{\log x \log\log\log x}{\log\log x}\right)\right) = x^{1/2+o(1)}. \end{aligned} \tag{32}$$

Using (31) and (32) in (22), we obtain the upper bound in (14).

Acknowledgement. The authors would like to thank the referee for some very helpful suggestions.

References

[1] J. Massias and G. Robin, 'Bornes effectives pour certains fonctions concernant les nombres premiers', *J. Théorie Nombres Bordeaux* **8** (1996), 215–242.

[2] C. Pomerance, 'On the distribution of round numbers', Ootacamund, India 1984, K. Alladi, ed., Lecture Notes in Math. **1122** (1985), 173–200.

[3] J.B. Rosser and L. Schoenfeld, 'Approximate formulas for some functions of prime numbers', *Illinois J. Math.* **6** (1962), 64–94.

[4] G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory*, Cambridge University Press, 1995.

Addresses: Jean-Marie De Koninck, Départment de mathématiques, Université Laval, Québec G1K 7P4, Canada

Florian Luca, Mathematical Institute, UNAM, Ap. Postal 61-3 (Xangari), CP 58 089, Morelia, Michoacán, Mexico

E-mail: `jmdk@mat.ulaval.ca`; `fluca@matmor.unam.mx`

Received: 30 July 2004; **revised:** 20 December 2004