Bull. London Math. Soc. **46**
(2014), 1156-1166.

A refinement of the *abc* conjecture

by

O. Robert, C.L. Stewart* and G. Tenenbaum

Abstract. Based on recent work, by the first and third authors, on the distribution of the squarefree kernel of an integer, we present precise refinements of the famous *abc* conjecture. These rest on the sole heuristic assumption that, whenever a and b are coprime, then the kernels of a, b and $c = a + b$ are statistically independent.

Classification AMS: Primary: 11N25, 11D99, Secondary: 11N56.

1. Introduction

For any non-zero integer n let $k(n)$ denote the greatest squarefree factor of n, so that

$$k(n) = \prod_{p|n} p.$$

$k(n)$ is also called the core, the squarefree kernel and the radical of n. The *abc* conjecture, proposed by Oesterlé and Masser [9], is the conjecture that for each $\varepsilon > 0$ there exists a positive number $A_0(\varepsilon)$ such that for any pair (a, b) of distinct coprime positive integers

$$c < A_0(\varepsilon)k^{1+\varepsilon}, \tag{1·1}$$

where

$$c = a + b \quad \text{and} \quad k = k(abc). \tag{1·2}$$

The conjecture has a number of profound consequences [3], [8], [10], in particular in the study of Diophantine equations.

An explicit upper bound for c in terms of k was first established by Stewart and Tijdeman [16] in 1986. Subsequently Stewart and Yu [17] proved that there is an effectively computable positive number A_1 such that for all pairs (a, b) of coprime positive integers

$$c < \exp\left\{A_1 k^{1/3}(\log k)^3\right\}.$$

Several refinements or modifications to the *abc* conjecture have been put forward [1], [2], [11], [4], [5], [6]. For instance, van Frankenhuijsen, see (1.4) and (1.5) of [5], proposed that there exist positive numbers A_2 and A_3 so that (1·1) may be replaced by

$$c < k\exp\left(A_2\sqrt{\log k/\log_2 k}\right) \tag{1·3}$$

and that there exist infinitely many pairs (a, b) of distinct coprime positive integers for which

$$c > k\exp\left(A_3\sqrt{\log k/\log_2 k}\right). \tag{1·4}$$

Here and in the sequel, we let $\log_j$ denote for $j \geqslant 2$ the jth iterate of the function $x \mapsto \max(1, \log x)$ $(x > 0)$.

The purpose of this article is to provide a refinement which is more precise than those proposed previously. It is based on the recent work of Robert and Tenenbaum [13] on the function $N(x, y)$ which counts the number of positive integers n up to x whose greatest squarefree divisor is at most y. We shall base our conjecture on the heuristic assumption that whenever a and b are coprime positive integers $k(a + b)$ is statistically independent of $k(a)$ and $k(b)$. This is the only assumption that we require.

* The research of the second author is supported in part by the Canada Research Chairs Program and by Grant A3528 from the Natural Sciences and Engineering Research Council of Canada.

Conjecture *A*. *There exists a real number C_1 such that, if a and b are coprime positive integers, then, with c and k as in* (1·2),

$$c < k\exp\left(4\sqrt{\frac{3\log k}{\log_2 k}}\left(1+\frac{\log_3 k}{2\log_2 k}+\frac{C_1}{\log_2 k}\right)\right). \tag{1·5}$$

Furthermore, there exists a real number C_2 and infinitely many pairs of coprime positive integers a and b for which

$$c > k\exp\left(4\sqrt{\frac{3\log k}{\log_2 k}}\left(1+\frac{\log_3 k}{2\log_2 k}+\frac{C_2}{\log_2 k}\right)\right). \tag{1·6}$$

We remark that it follows from Conjecture A that for each $\varepsilon > 0$, we can select $A_2 = 4\sqrt{3}+\varepsilon$ in (1·3) for large k, and $A_3 = 4\sqrt{3}-\varepsilon$ in (1·4).

There have been several computational studies undertaken in order to test the plausibility of the *abc* conjecture. The most extensive is *Reken mee met ABC* [12],[7] based at the Universiteit Leiden. It is a distributed computing program involving many individuals. Associated with each triple (a, b, c) of coprime positive integers with $a + b = c$ are two quantities, the *quality* q defined by

$$q = (\log c)/\log k$$

and the *merit* m defined by

$$m = (q-1)^2(\log k)\log_2 k.$$

B. de Smit maintains a website [14] to keep track of exceptional triples, measured by the sizes of their quality and merit, which have been found by virtue of the above project.The largest known quality of a triple is ≈ 1.63 and the five triples known with quality larger than 1.55 have c at most 10^{16}. It follows from Conjecture A that the limit supremum of m as we range over all pairs (a, b) of distinct coprime positive integers is 48. To date nineteen triples have been found with merit larger than 30, each with c at least 10^{20}, and eighty-three with merit larger than 25. The triple with largest known merit was found by Ralf Bonse. It is

$$a = 2543^4\cdot 182587\cdot 2802983\cdot 85813163,\quad b = 2^{15}\cdot 3^{77}\cdot 11\cdot 173,\quad c = 5^{56}\cdot 245983,$$

and has merit ≈ 38.67.

In [16] Stewart and Tijdeman proved that for each positive real number ε there exist infinitely many pairs (a, b) of coprime positive integers for which

$$c > k\exp\left\{(4-\varepsilon)\sqrt{\log k}/\log_2 k\right\}. \tag{1·7}$$

Subsequently, van Frankenhuijsen [5] improved $4-\varepsilon$ in (1·7) to 6.068.

2. Further refinements of Conjecture *A*

Conjecture A is based on our heuristic assumption, recall §1, and a careful analysis of the behaviour of the function $N(x,y)$ which counts the number of positive integers n up to x for which $k(n)$ is at most y. Thus

$$N(x,y) := \sum_{\substack{n\leqslant x\\ k(n)\leqslant y}} 1. \tag{2·1}$$

Set

$$\psi(m) := \prod_{p\mid m}(p+1) \quad (m\geqslant 1), \qquad F(t) := \frac{6}{\pi^2}\sum_{m\geqslant 1}\frac{\min(1,\mathrm{e}^t/m)}{\psi(m)} \qquad (t\geqslant 0). \tag{2·2}$$

As stated below (see Proposition 3.1), we have $N(x,y)\sim yF(v)$ with $v:=\log(x/y)$ in a wide range for the pair (x,y).

It was announced in Squalli's doctoral dissertation [15] and proved in [13] that there exists a sequence of polynomials $\{Q_j\}_{j=1}^{\infty}$ with $\deg Q_j\leqslant j$, such that, for any integer $N\geqslant 1$,

$$F(t) = \exp\left\{\sqrt{\frac{8t}{\log t}}\left(1+\sum_{1\leqslant j\leqslant N}\frac{Q_j(\log_2 t)}{(\log t)^j}+O_N\left(\left(\frac{\log_2 t}{\log t}\right)^{N+1}\right)\right)\right\} \quad (t\geqslant 3). \tag{2·3}$$

In particular,

$$Q_1(X) := \tfrac12 X - \tfrac12\log 2 + 1, \quad Q_2(X) := \tfrac38 X^2 + (1-\tfrac34\log 2)X + 2 + \tfrac23\pi^2 + \tfrac38(\log 2)^2 - \log 2.$$

The following version of the conjecture, which is expressed in terms of the function F, is slightly more precise than Conjecture A. Indeed, it corresponds to the extra information that, for large k, we have

$$\max(C_1,C_2) < \lambda := 1 - \tfrac12\log(\tfrac43). \tag{2·4}$$

Conjecture *B*. *There exist positive numbers B_0 and B_1 such that if a and b are coprime positive integers, then, with c and k as in (1·2),*

$$c < B_0 k F\left(\tfrac23\log k\right)^{3-B_1/\log_2 k}. \tag{2·5}$$

Furthermore, there exists a positive number B_2 and infinitely many pairs (a,b) of distinct coprime positive integers with

$$c > kF\left(\tfrac23\log k\right)^{3-B_2/\log_2 k}. \tag{2·6}$$

To see that the two conjectures are equivalent provided one assumes (2·4), it suffices to appeal to (2·3) taking the form of Q_1 into account. Condition (2·4) corresponds to the condition that B_1 and B_2 are positive.

As will be seen in the final section, Conjecture B is itself a consequence of a further refined conjecture, involving the implicit function $\mathcal{H}(k)$ defined in (4·6) below in terms of solutions of certain transcendental equations. Using techniques developed in [13], it may be shown that, for any fixed integer J, we have

$$\log\mathcal{H}(k) = -\sqrt{\frac{\log k}{\log_2 k}}\left\{\sum_{1\leqslant j\leqslant J}\frac{R_j(\log_3 k)}{(\log_2 k)^j}+O\left(\left(\frac{\log_3 k}{\log_2 k}\right)^{J+1}\right)\right\} \qquad (k\to\infty) \tag{2·7}$$

where R_j is a polynomial of degree at most j. In particular, $R_1(X) = 8(\log 2)/\sqrt3$ is a positive constant.

Conjecture *C*. *Let $\varepsilon > 0$. There exists a positive number $B_3 = B_3(\varepsilon)$ such that, if a and b are coprime positive integers, then, with c and k as in* (1·2), *we have*

$$c \leqslant B_3 k F(\tfrac{2}{3}\log k)^3 \mathcal{H}(k)(\log k)^{11/2+\varepsilon}. \tag{2·8}$$

Furthermore, infinitely many such pairs (a, b) satisfy

$$c > kF(\tfrac{2}{3}\log k)^3 \mathcal{H}(k)/(\log k)^{3/2+\varepsilon} \tag{2·9}$$

Remarks. (i) We did not try to optimize the exponents of the log-factors in (2·8) and (2·9).

(ii) It follows from Conjecture C and the value of R_1 given above that, given any $\varepsilon > 0$, we may select $B_1 = \log 4 - \varepsilon$, $B_2 = \log 4 + \varepsilon$ in Conjecture B, and $C_1 = \beta + \varepsilon$, $C_2 = \beta - \varepsilon$, where $\beta := 1 + \log 3 - \frac{13}{6}\log 2$, in Conjecture A.

Furnishing an estimate for $c = a + b$ which is sharp up to a power of $\log k$, this last formulation has a nice probabilistic interpretation which brings some further insight into the problem: the F-factor takes care of the statistical distribution of the squarefree kernel, and the $\mathcal{H}$-factor corresponds to the condition that a and b should be coprime. Indeed, integers with a small core have a strong tendency to be divisible by many small primes; hence the probability that two such integers should be coprime is very small. Thus the factor $\mathcal{H}(k)$ above may be seen as playing the same rôle, for pairs (a, b) with maximal $k = k(abc)$, as the well-known probability $6/\pi^2$ for unconstrained random integers.

3. Estimates for $N(x, y)$

Let

$$f(\sigma) := \sum_{n\geqslant 1} \frac{1}{\psi(n)n^\sigma} = \prod_p \Big(1 + \frac{1}{(p+1)(p^\sigma - 1)}\Big) \qquad (\sigma > 0), \tag{3·1}$$

and put

$$g(\sigma) = \log f(\sigma).$$

For $v \geqslant 6$, we let σ_v denote the solution of the transcendental equation

$$-g'(\sigma) = \sum_p \frac{p^\sigma \log p}{(p^\sigma - 1)\{1 + (p+1)(p^\sigma - 1)\}} = v \tag{3·2}$$

and make the convention that $\sigma_v = \frac{1}{2}$ when $0 \leqslant v < 6$. Thus, for $v > 6$, $\sigma = \sigma_v$ renders the quantity $\mathrm{e}^{\sigma v} f(\sigma)$ minimal. The function σ_v has been extensively studied in [13]. For any given integer $K \geqslant 1$, we have

$$\sigma_v = \sqrt{\frac{2}{v\log v}}\Big\{1 + \sum_{1\leqslant k\leqslant K} \frac{P_k(\log_2 v)}{(\log v)^k} + O_K\Big(\frac{(\log_2 v)^{K+1}}{(\log v)^{K+1}}\Big)\Big\} \quad (v \geqslant 3), \tag{3·3}$$

where P_k is a suitable polynomial of degree at most k. In particular,

$$P_1(z) = \tfrac{1}{2}(z - \log 2), \quad P_2(z) = \tfrac{3}{8}z^2 - (\tfrac{3}{4}\log 2 + \tfrac{1}{2})z + \tfrac{1}{2}\log 2 + \tfrac{3}{8}(\log 2)^2 + \tfrac{2}{3}\pi^2. \tag{3·4}$$

Here and in the sequel, we put

$$v = \log(x/y), \quad \mathcal{Y}_x := \mathrm{e}^{\frac{1}{4}\sqrt{2\log x}(\log_2 x)^{3/2}}, \quad \mathfrak{E}_t(x, y) := \frac{\sqrt{v\sigma_v}\log y}{y^{\sigma_v/t}} + \frac{1}{x^{1/16}} \quad (t > 0).$$

We recall from [13] that $\mathcal{Y}_x$ is an approximation to the threshold of the phase transition of the asymptotic behaviour of $N(x,y)$: given any $\varepsilon > 0$, we have $N(x,y) \sim yF(v)$ for $y > \mathcal{Y}_x^{1+\varepsilon}$ and $N(x,y) = o\big(yF(v)\big)$ whenever $y \leqslant \mathcal{Y}_x^{1-\varepsilon}$. The following statement, which is a consequence of theorem 3.3 and proposition 10.1 of [13], provides the effective version we shall need.

We recall Vinogradov's notations $f \ll g$ and $f \gg g$, meaning, respectively, that $|f| \leqslant C|g|$ and $|f| \geqslant C'|g|$ for suitable positive constants C, C'. The symbol $f \asymp g$ then means that $f \ll g$ and $f \gg g$ hold simultaneously.

Proposition 3.1. *Let* $\varepsilon > 0$. *We have*

$$N(x,y) = yF(v)\{1 + O\big(\mathfrak{E}_1(x,y)\big)\} \qquad \big(x \to \infty,\ \mathcal{Y}_x^{1+\varepsilon} \leqslant y \leqslant x\big) \tag{3·5}$$

$$N(x,y) \ll yF(v) \qquad (x \geqslant y \geqslant 2). \tag{3·6}$$

We also make use of the following result concerning the size and variation of F. Here again, we state more than necessary for our present purpose, but less than proved in [13] (Theorem 8.6, Propositions 8.8 and 8.9).

Proposition 3.2. *We have*

$$F(v) \asymp \Big(\frac{\log v}{v}\Big)^{1/4} \mathrm{e}^{v\sigma_v} f(\sigma_v) = \mathrm{e}^{2v\sigma_v + O(v\sigma_v/\log v)} \qquad (v \geqslant 2), \tag{3·7}$$

$$F(v+h) \ll F(v)\mathrm{e}^{\sigma_v h} \quad (v \geqslant 0,\ v+h \geqslant 0), \tag{3·8}$$

$$F(v+h) - F(v) = \left\{1 + O\left(\frac{\log v + |h|}{\sqrt{v\log v}}\right)\right\} h\sigma_v F(v) \qquad \big(v \geqslant 2,\ h \ll \sqrt{v\log v}\big). \tag{3·9}$$

Finally, we state the following result, where, for $a \geqslant 1$, we employ the notation

$$N_a(x,y) := \sum_{\substack{n\leqslant x\\ (n,a)=1\\ k(n)\leqslant y}} 1, \quad F_a(v) := \frac{6}{\pi^2}\sum_{(m,a)=1} \frac{\min(1, \mathrm{e}^v/m)}{\psi(m)}, \quad r(a) := \prod_{p\mid a}\Big(1 + \frac{2}{\sqrt{p}}\Big),$$

and let φ denote Euler's totient.

Proposition 3.3. *We have*

$$F_a(v+h) - F_a(v) \gg \sum_{\substack{m\geqslant \mathrm{e}^{v+h}\\ (m,a)=1}} \frac{\mathrm{e}^v}{m\psi(m)} \qquad (a \geqslant 1,\ v \geqslant 2,\ h \asymp 1), \tag{3·10}$$

$$N_a(x,y) = \frac{yk(a)F_a(v)}{\psi(a)}\Big\{1 + O\Big(r(a)\mathfrak{E}_2(x,y)\Big)\Big\} \qquad (\mathcal{Y}_x^2 \leqslant y \leqslant x,\ a \leqslant x). \tag{3·11}$$

Proof. The bound (3·10) immediately follows from the definition of $F_a(v)$ by restricting the sum to $m > \mathrm{e}^{v+h}$.

Estimate (3·11) may be proved along the lines of proposition 10.1 in [13], which corresponds to $a = 1$. We avoid repeating the details here since they are identical to those of [13], simply carrying the condition $(m,a) = 1$ throughout the computations and appealing to the saddle-point estimate for $F_a(v)$. □

To state our next lemma, we introduce some further notation. Let us define

$$H(s,z) := \prod_p \Big(1 + \frac{1}{(p+1)(p^s-1)} + \frac{1}{(p+1)(p^z-1)}\Big) \qquad (\Re e\, s > 0,\ \Re e\, z > 0). \tag{3·12}$$

For $v>0$, we denote by $\vartheta_v>0$ the unique solution to the equation

$$\sum_p \frac{p^\sigma \log p}{(p^\sigma-1)\{2+(p+1)(p^\sigma-1)\}}=v, \tag{3·13}$$

so that $(s,z)=(\vartheta_v,\vartheta_v)$ is a real saddle-point for $(s,z)\mapsto \mathrm{e}^{(s+z)v}H(s,z)$. Moreover, it can be checked that

$$\vartheta_v=\sigma_v\{1+O(1/\log v)\}\qquad (v\geqslant 2). \tag{3·14}$$

Finally, we set

$$h(\sigma):=\log H(\sigma,\sigma)\qquad (\sigma>0) \tag{3·15}$$

and note that

$$H(\sigma,\sigma)=\mathrm{e}^{h(\sigma)}=f(\sigma)^2\prod_p\Big(1-\frac{1}{\{1+(p^\sigma-1)(p+1)\}^2}\Big)\qquad (\sigma>0). \tag{3·16}$$

Proposition 3.4. *Let $\kappa\in(0,\frac12)$, $\mu>0$. For $x^\kappa\leqslant y\leqslant x^{1-\kappa}$, and suitable $B=B(\kappa)$, we have*

$$\sum_{\substack{x<a\leqslant \mathrm{e}^\mu x\\ a/\mathrm{e}^\mu<b<a,\,(a,b)=1\\ k(a)\leqslant y,\,k(b)\leqslant y}} 1 \gg \frac{y^2\mathrm{e}^{2v\vartheta_v+h(\vartheta_v)}}{v^{3/2}(\log v)^{5/2}}\gg y^2F(v)^{2-B/\log v}. \tag{3·17}$$

Proof. Let $D(x,y)$ denote the double sum to be estimated. By (3·11) and (3·10), we have

$$D(x,y)\geqslant D_1-R_1$$

with

$$D_1\gg \mathrm{e}^v y\sum_{\substack{x<a\leqslant \mathrm{e}^\mu x\\ k(a)\leqslant y}}\frac{k(a)}{\psi(a)}\sum_{\substack{m>\mathrm{e}^{v+\mu}\\ (m,a)=1}}\frac{1}{m\psi(m)}\gg \frac{y\mathrm{e}^v}{\log v}\sum_{\substack{x<a\leqslant \mathrm{e}^\mu x\\ k(a)\leqslant y}}\sum_{\substack{m>\mathrm{e}^{v+\mu}\\ (m,a)=1}}\frac{1}{m\psi(m)},$$
$$R_1\ll y^2F(v)^{2-\kappa_1},$$

for some positive constant κ_1 depending only on κ. Next, we invert summations in our lower bound for D_1 and appeal to (3·11) and (3·10) again. We get $D_1\geqslant D_2-R_2$ with

$$D_2\gg\frac{y^2\mathrm{e}^{2v}}{\log v}S,\quad S:=\sum_{\substack{m,n>\mathrm{e}^{v+\mu}\\ (m,n)=1}}\frac{k(m)}{mn\psi(m)^2\psi(n)},\quad R_2\ll y^2F(v)^{2-\kappa_1}.$$

It remains to bound S from below. To this end, we restrict the sum to pairs (m,n) in $(\mathrm{e}^{v+\mu},\mathrm{e}^{v+2\mu}]^2$ to get $\mathrm{e}^{2v}S\gg T/\log v$ with

$$T:=\sum_{\substack{\mathrm{e}^{v+\mu}<m,n\leqslant \mathrm{e}^{v+2\mu}\\ (m,n)=1}}\frac{1}{\psi(m)\psi(n)}$$
$$=\frac{1}{(2\pi i)^2}\int_{(\sigma_v+i\mathbb{R})^2}\frac{H(s,z)\mathrm{e}^{(v+\mu)(s+z)}(\mathrm{e}^{\mu s}-1)(\mathrm{e}^{\mu z}-1)}{sz}\,\mathrm{d}s\,\mathrm{d}z$$

where $H(s,z)$ is defined by (3·12).

We estimate the last integral by two-dimensional saddle-point method. Since similar calculations have been extensively described in [13], we only sketch the proof.

Writing $s = \vartheta_v + i\tau$, $z = \vartheta_v + it$, we deduce from lemma 5.13 and formula (7.7) of [13] that, for a suitable absolute constant η, we have

$$|H(s,z)| \leqslant \mathrm{e}^{-\eta(\log v)^2} H(\vartheta_v, \vartheta_v)$$

provided $(\log v)^{5/4}/v^{3/4} \ll \max(|\tau|, |t|) \leqslant \exp\{(\log v)^{38/37}\}$. Truncating the larger values by standard effective Perron formula (see, for instance, [18], theorem II.2.3), we may evaluate the double integral on the remaining small domain by saddle-point analysis, taking advantage of the fact that

$$\mathfrak{h}(s,z) := \sum_p \log\Big(1 + \frac{1}{(p+1)(p^s-1)} + \frac{1}{(p+1)(p^z-1)}\Big), \tag{3·18}$$

where the complex logarithms are understood in principal branch, defines a holomorphic continuation of $\mathfrak{h}(s,z)$ in a poly-disc of centre $(\vartheta_v, \vartheta_v)$ and radii $\frac{1}{2}\vartheta_v$.(1)

We thus arrive at

$$T \sim \frac{\mu^2 \mathrm{e}^{2v\vartheta_v} H(\vartheta_v, \vartheta_v)}{2\pi j(\vartheta_v)} \qquad (v \to \infty),$$

with

$$j(\sigma) := \sum_p \frac{p^\sigma (\log p)^2 \{(p+1)(p^{2\sigma}-1) + p^\sigma + 2\}}{(p^\sigma - 1)^2\{2 + (p^\sigma - 1)(p+1)\}^2} \asymp \frac{1}{\sigma^3 \log(1/\sigma)} \qquad (\sigma \to 0+).$$

This plainly yields the first lower bound in (3·17).

To prove the second lower bound, we appeal to (3·16), note that the estimate (3·14) implies $2v\vartheta_v + h(\vartheta_v) = 2v\sigma_v + h(\sigma_v) + O\big(v\sigma_v/\log v\big)$, and insert the lower bound

$$\prod_p \bigg(1 - \frac{1}{\{1 + (p^{\sigma_v} - 1)(p+1)\}^2}\bigg) \gg F(v)^{-c_0/\log v},$$

for a suitable absolute constant $c_0 > 0$. □

4. Justification for Conjectures *B* and *C*

We shall establish Conjectures B and C under the heuristic assumption that, whenever a and b are coprime integers, the kernel $k(a+b)$ is distributed as if $a+b$ was a typical integer of the same size. Albeit conjecture B formally follows from Conjecture C and (2·7), we shall provide a direct, simple proof. Notice that if $(a,b) = 1$ and $a + b = c$, then $k(abc) = k(a)k(b)k(c)$.

We start with the upper bounds. Under the above assumption, we may write

$$\mathcal{P}(x,z) := \sum_{\substack{x<a\leqslant 2x \\ b<a,\, (a,b)=1 \\ k(abc)\leqslant z}} 1 \leqslant \sum_{\substack{x<a\leqslant 2x \\ b<a,\, (a,b)=1}} \frac{1}{x}\Big\{N\Big(4x, \frac{z}{k(a)k(b)}\Big) - N\Big(x, \frac{z}{k(a)k(b)}\Big)\Big\}.$$

To prove (2·5), it suffices to show that, for $z = Z_x := x/F(\frac{2}{3}\log x)^{3-B_4/\log_2 x}$ and suitable $B_4 > 0$,we have

$$\sum_{r\geqslant 1} \mathcal{P}(2^r, Z_{2^r}) < \infty. \tag{4·1}$$

Indeed, this plainly implies that the conditions $k(abc) \leqslant z$ for some pair (a,b) with $x < a \leqslant 2x$, $b < a$, are realized only for a bounded number of integers x. This argument is similar to that of the Borel-Cantelli lemma.

1. See [13], lemma 8.4 for the details, in a similar situation, of the continuation, and theorem 8.6, for those of the saddle-point analysis.

Applying (3·6) and (3·8) taking (2·3) and (3·3) into account, we obtain

$$\mathcal{P}(x,z) \ll \frac{z}{x} \sum_{\substack{x<a\leqslant 2x\\ b<a,\,(a,b)=1}} \frac{F\big(\log(xk(a)k(b)/z)\big)}{k(a)k(b)} \ll \frac{zF(v)}{x} \sum_{\substack{x<a\leqslant 2x\\ b<a,\,(a,b)=1}} \frac{x^{-2\sigma_v/3}}{k(a)^{1-\sigma_v}k(b)^{1-\sigma_v}}$$

with $v := \frac{2}{3}\log x$. By Rankin's method, we thus infer, writing $P(n)$ for the largest prime factor of an integer n with the convention that $P(1) = 1$,

$$\begin{aligned}\mathcal{P}(x,z) &\ll \frac{zF(v)}{x} \sum_{P(a)\leqslant x} \frac{x^{2\sigma_v/3}}{a^{\sigma_v}k(a)^{1-\sigma_v}} \sum_{\substack{P(b)\leqslant x\\ (b,a)=1}} \frac{x^{2\sigma_v/3}}{b^{\sigma_v}k(b)^{1-\sigma_v}}\\ &\ll \frac{zF(v)\mathrm{e}^{2v\sigma_v}}{x} \sum_{P(a)\leqslant x} \frac{1}{a^{\sigma_v}k(a)^{1-\sigma_v}} \prod_{\substack{p\leqslant x\\ p\nmid a}} \Big(1+\frac{1}{p(1-p^{-\sigma_v})}\Big).\end{aligned}$$

Since a standard computation yields, taking (3·7) into account,

$$\mathrm{e}^{v\sigma_v} \prod_{p\leqslant x} \Big(1+\frac{1}{p(1-p^{-\sigma_v})}\Big) \ll \frac{F(v)v^{5/4}}{(\log v)^{1/4}},$$

we get

$$\begin{aligned}\mathcal{P}(x,z) &\ll \frac{zF(v)^2\mathrm{e}^{v\sigma_v}v^{5/4}}{x(\log v)^{1/4}} \sum_{P(a)\leqslant x} \frac{1}{a^{\sigma_v}k(a)^{1-\sigma_v}} \prod_{p\mid a} \Big(1-\frac{1}{1+p(1-p^{-\sigma_v})}\Big)\\ &\ll \frac{zF(v)^2\mathrm{e}^{v\sigma_v}v^{5/4}}{x(\log v)^{1/4}} \prod_{p\leqslant x} \Big(1+\frac{1}{p(1-p^{-\sigma_v})}\Big)\Big(1-\frac{1}{\{1+p(1-p^{-\sigma_v})\}^2}\Big)\\ &\ll \frac{zF(v)^{3-K_0/\log v}}{x},\end{aligned}$$

where K_0 is a suitable positive constant.

This establishes the upper bound for c in Conjecture B.

We now embark on proving (2·8) and first define the quantity $\mathcal{H}(k)$, noticing that we shall now select in (4·1)

$$z = Z_x := \frac{x}{F(\frac{2}{3}\log x)^3\mathcal{H}(x)(\log x)^{11/2+\varepsilon}}.$$

Given $x \geqslant 2$, we let $u = u_x$ be the solution to the equation

$$\sigma_u = \vartheta_w \qquad (w := \log x - \tfrac{1}{2}u). \tag{4·2}$$

It is easy to see that

$$u = \tfrac{2}{3}\log x + O\Big(\frac{\log x}{\log_2 x}\Big), \quad w = \tfrac{2}{3}\log x + O\Big(\frac{\log x}{\log_2 x}\Big)$$

and a further computation actually yields $u - \frac{2}{3}\log x \sim 8(\log 2)(\log x)/9\log_2 x$. Recalling notation (3·15) and introducing $g(\sigma) := \log f(\sigma)$ $(\sigma > 0)$, we then put

$$\mathcal{H}_1(k) := \mathrm{e}^{2\sigma_u(w-u)} \prod_p \Big(1-\frac{1}{\{1+(p^{\vartheta_w}-1)(p+1)\}^2}\Big) = \mathrm{e}^{2\sigma_u(w-u)+h(\sigma_u)-2g(\sigma_u)}, \tag{4·3}$$

with $u := u_k$, $w := \log k - \frac{1}{2}u_k$.

We shall set out to prove

$$c \leqslant B_3 k F(u_k)^3 \mathcal{H}_1(k)(\log k)^{11/2+\varepsilon}, \tag{4·4}$$

and

$$c > kF(u_k)^3 \mathcal{H}_1(k)/(\log k)^{3/2+\varepsilon} \tag{4·5}$$

instead of (2·8) and (2·9) respectively. However, it can be shown that $F(u_k)/F(\frac{2}{3}\log k)$ satisfies a relation of type (2·7) with a different sequence of polynomials R_j. From this observation, the required result will follow with

$$\mathcal{H}(k) := F(u_k)^3 \mathcal{H}_1(k)/F(\tfrac{2}{3}\log k)^3. \tag{4·6}$$

Applying (2·3), (3·3), (3·6) and (3·8) again, we get

$$\begin{aligned} \mathcal{P}(x,z) &\ll \frac{z}{x} \sum_{\substack{x<a\leqslant 2x \\ b<a,\,(a,b)=1 \\ k(ab)\leqslant x}} \frac{F\big(\log\{xk(ab)/z\}\big)}{k(a)k(b)} \\ &\ll \frac{z}{x} \sum_{m+n\leqslant \log x} \frac{F(m+n)+F(\frac{1}{3}\log x)}{\mathrm{e}^{m+n}} S(m,n), \end{aligned}$$

with

$$S(m,n) := \sum_{\substack{a\leqslant 2x,\, b\leqslant 2x \\ (a,b)=1 \\ \mathrm{e}^{m-1}<k(a)\leqslant \mathrm{e}^{m+1},\, \mathrm{e}^{n-1}<k(b)\leqslant \mathrm{e}^{n+1}}} 1 \qquad (m\geqslant 1,\, n\geqslant 1).$$

Now, for all m, n and any $\vartheta \in]0,1[$, we may write

$$\begin{aligned} S(m,n) &\leqslant \sum_{\substack{a\leqslant 2x,\, b\leqslant 2x \\ (a,b)=1}} \Big(\frac{2x}{a}\Big)^{\vartheta} \Big(\frac{2x}{b}\Big)^{\vartheta} \Big(\frac{\mathrm{e}^{m+1}}{k(a)}\Big)^{1-\vartheta} \Big(\frac{\mathrm{e}^{n+1}}{k(b)}\Big)^{1-\vartheta} \\ &\ll x^{2\vartheta} \mathrm{e}^{(1-\vartheta)(m+n)} \prod_{p\leqslant 2x} \Big(1+\frac{2}{p^{1-\vartheta}(p^{\vartheta}-1)}\Big) \\ &\ll x^{2\vartheta} \mathrm{e}^{(1-\vartheta)(m+n)} H(\vartheta,\vartheta)(\log x)^2. \end{aligned}$$

Writing $s := m+n$, $t := \log x - \frac{1}{2}s$, we infer that

$$\frac{F(m+n)S(m,n)}{\mathrm{e}^{m+n}} \ll \Big(\frac{\log s}{s}\Big)^{1/4} \mathrm{e}^{s\sigma_s + g(\sigma_s) + 2t\vartheta_t + h(\vartheta_t)} (\log x)^2.$$

By (4·2) and the definition of ϑ_v, the argument of the exponential is maximal when $s = u := u_x$, $t = w := \log x - \frac{1}{2}u_x$. For this choice, the last upper bound is equally valid when $F(m+n)$ is replaced by $F(\frac{1}{3}\log x) \ll F(u)x^{-\sigma_u/4}$.

Selecting the above values for s, t and carrying back our estimates in the upper bound for $\mathcal{P}(x,z)$, we thus obtain that

$$\mathcal{P}(x,z) \ll \frac{zF(u)\mathrm{e}^{2w\vartheta_w + h(\vartheta_w)} u^4}{x} \asymp \frac{zF(u)^3 \mathcal{H}_1(x) u^{9/2}}{x\sqrt{\log u}}. \tag{4·7}$$

The bound (4·7) is sufficient to ensure the convergence of the series (4·1) provided $\varepsilon > 0$. This completes our argument in favour of the upper bound in conjecture C.

In order to justify the lower bounds, we show that, still under the assumption that $k(c)$ behaves independently of $k(a)$ and $k(b)$, we have $\mathcal{P}(x,z)\to\infty$ for an appropriate value $z=z_x$.

Let us start with Conjecture B. According to the above hypothesis, we may write, for $x^{2/3+\varepsilon}<z\leqslant x$

$$\begin{aligned}\mathcal{P}(x,z)\geqslant&\sum_{\substack{x<a\leqslant 2x\\ a/2<b<a,\,(a,b)=1\\ k(a)\leqslant x^{1/3},\,k(b)\leqslant x^{1/3}}}\frac{2}{3x}\Big\{N\Big(3x,\frac{z}{k(a)k(b)}\Big)-N\Big(\frac{3x}{2},\frac{z}{k(a)k(b)}\Big)\Big\}\\ &\gg\frac{z}{x}F\big(\tfrac23\log x\big)^{2-(B+1)/\log_2 x}F\big(\tfrac53\log x-\log z\big)\gg\frac{z}{x}F\big(\tfrac23\log x\big)^{3-(B+1)/\log_2 x},\end{aligned}$$

where we successively appealed to (3·5), (3·9) and (3·17). Selecting

$$z=x/F(\tfrac23\log x)^{3-(B+2)/\log_2 x},$$

we obtain the required estimate.

Finally, we establish the lower bound in Conjecture C. For $x^{2/3+\varepsilon}<z\leqslant x$, $u:=u_x$, $y:=\mathrm{e}^{u/2}$, $w:=\log x-u/2$, we have

$$\begin{aligned}\mathcal{P}(x,z)\geqslant&\sum_{\substack{x<a\leqslant 2x\\ a/2<b<a,\,(a,b)=1\\ k(a)\leqslant y,\,k(b)\leqslant y}}\frac{2}{3x}\Big\{N\Big(3x,\frac{z}{k(a)k(b)}\Big)-N\Big(\frac{3x}{2},\frac{z}{k(a)k(b)}\Big)\Big\}\\ &\gg\frac{z\sigma_u}{x}\sum_{\substack{x<a\leqslant 2x\\ a/2<b<a,\,(a,b)=1\\ k(a)\leqslant y,\,k(b)\leqslant y}}\frac{F(\log\{xk(a)k(b)/z\})}{k(a)k(b)}.\end{aligned}$$

At this stage, we observe that, for sufficiently large x, we have

$$F(u)\leqslant F\big(\log(x\mathrm{e}^u/z)\big)\ll F\big(\log\{xk(a)k(b)/z\}\big)\frac{\mathrm{e}^{u/2}}{\sqrt{k(a)k(b)}}\tag{4·8}$$

uniformly for all a,b in the last range of summation. Indeed, the first inequality readily follows from the fact that $z\leqslant x$, and the second bound is obtained by applying (3·8) with $v=v(a,b,x,z):=\log(xk(a)k(b)/z)$ and $h=h(a,b,x,z):=\log\big(\mathrm{e}^u/k(a)k(b)\big)$: since $h\geqslant 0$ and $v\to\infty$ uniformly in a,b as $x\to\infty$, we plainly have $\sigma_v\leqslant 1/2$ for large x, which implies (4·8).

Inserting (4·8) in our previous lower bound for $\mathcal{P}(x,z)$ yields

$$\begin{aligned}\mathcal{P}(x,z)&\gg\frac{z\sigma_uF(u)}{x}\sum_{\substack{x<a\leqslant 2x\\ a/2<b<a,\,(a,b)=1\\ k(a)\leqslant y,\,k(b)\leqslant y}}\frac{1}{\sqrt{k(a)k(b)}\mathrm{e}^{u/2}}\gg\frac{z\mathrm{e}^{2w\vartheta_w+h(\vartheta_w)}F(u)}{xu^2(\log u)^3}\\ &\gg\frac{z\mathrm{e}^{2w\vartheta_w+h(\vartheta_w)+u\sigma_u+g(\sigma_u)}}{xu^{9/4}(\log u)^{11/4}}\asymp\frac{z\mathrm{e}^{3u\sigma_u+3g(\sigma_u)+2(w-u)\sigma_u+h(\sigma_u)-2g(\sigma_u)}}{xu^{9/4}(\log u)^{11/4}}\\ &\asymp\frac{zF(u)^3\mathcal{H}_1(x)}{xu^{3/2}(\log u)^{7/2}},\end{aligned}$$

where we successively appealed to (3·5), (3·9), (3·8), (3·17) and (3·7). Selecting

$$z=x(\log x)^{3/2+\varepsilon}/F(u)^3\mathcal{H}_1(x),$$

completes the proof.

References

[1] A. Baker, Logarithmic forms and the *abc*-conjecture, in: *Number Theory* (*Diophantine, computational and algebraic aspects*), Proceedings of the international conference, Eger, Hungary (K. Győry et al., ed.), de Gruyter, Berlin, 1998, 37–44.
[2] A. Baker, Experiments on the *abc*-conjecture, *Publ. Math. Debrecen* **65** (2004), 253–260.
[3] E. Bombieri and W. Gubler, Heights in Diophantine geometry, Cambridge University Press, Cambridge, 2007.
[4] M. van Frankenhuijsen, Hyperbolic spaces and the *abc* conjecture, thesis, Katholieke Universiteit Nijmegen, 1995.
[5] M. van Frankenhuijsen, A lower bound in the *abc* conjecture, *J. of Number Theory* **82** (2000), 91–95.
[6] M. van Frankenhuijsen, About the ABC conjecture and an alternative, *Number theory, analysis and geometry,* Springer, New York, 2012, 169-180.
[7] G. Geuze and B. de Smit, Reken mee met ABC, *Nieuw Arch. voor Wiskunde* **8** (2007), 26-30.
[8] A. Granville and T.J. Tucker, It's as easy as *abc*, *Notices of the A.M.S.* **49** (2002), 1224–1231.
[9] D.W. Masser, Open problems, *Proc. Symp. Analytic Number Theory*, (W.W.L. Chen, ed.), Imperial Coll. London, 1985.
[10] A. Nitaj, http://www.math.unicaen.fr/nitaj/abc.html, 2012.
[11] C. Pomerance, Computational number theory, Princeton Companion to Mathematics, W.T. Gowers, ed., Princeton U. Press, Princeton, New Jersey, 2008, 348–362.
[12] Reken mee met ABC, http://www.rekenmeemetabc.nl/, 2012.
[13] O. Robert and G. Tenenbaum, Sur la répartition du noyau d'un entier, *Indag. Math.* **24** (2013), 802–914.
[14] B. de Smit, http://www.math.leidenuniv.nl/∼desmit/abc/, 2012.
[15] H. Squalli, *Sur la répartition du noyau d'un entier*, Thèse de doctorat de troisième cycle, Université Nancy I, November 18, 1985.
[16] C.L. Stewart and R. Tijdeman, On the Oesterlé-Masser conjecture, *Monatshefte Math.* **102** (1986), 251–257.
[17] C.L. Stewart and K. Yu, On the *abc* conjecture, II, *Duke Math. Journal* **108** (2001), 169–181.
[18] G. Tenenbaum, *Introduction à la théorie analytique et probabiliste des nombres*, third ed., coll. Échelles, Belin, 2008.

Olivier Robert
Universités de Lyon & Saint-Étienne
Institut Camille Jordan (UMR 5208)
23, rue du Dr P. Michelon
F-42000 Saint-Étienne
France
olivier.robert@univ-st-etienne.fr

Cameron L. Stewart
Department of Pure Mathematics,
University of Waterloo,
Waterloo, Ontario,
Canada N2L 3G1
cstewart@uwaterloo.ca

Gérald Tenenbaum
Institut Élie Cartan
Université de Lorraine
BP 70239
54506 Vandœuvre-lès-Nancy Cedex
France
gerald.tenenbaum@univ-lorraine.fr