Smooth solutions to the *abc* equation:
the *xyz* Conjecture

par JEFFREY C. LAGARIAS et KANNAN SOUNDARARAJAN

RÉSUMÉ. Cet article étudie les solutions entières de l'équation *abc* pour lesquelles ni A, ni B, ni C n'ont de grands facteurs premiers. On pose $H(A, B, C) = \max(|A|, |B|, |C|)$, et on considère les solutions primitives $(\gcd(A, B, C) = 1)$ n'ayant aucun facteur premier plus grand que $(\log H(A, B, C))^\kappa$, pour un κ fini donné. Nous montrons que la Conjecture *abc* entraine que pour tout $\kappa < 1$ l'équation n'a qu'un nombre fini de solutions primitives. Nous donnons aussi un résultat conditionnel, affirmant que l'hypothèse de Riemann généralisée (GRH) implique que pour tout $\kappa > 8$ l'équation *abc* a un nombre infini de solutions primitives. Nous esquissons la preuve de ce dernier résultat.

ABSTRACT. This paper studies integer solutions to the *abc* equation $A + B + C = 0$ in which none of A, B, C have a large prime factor. We set $H(A, B, C) = \max(|A|, |B|, |C|)$, and consider primitive solutions $(\gcd(A, B, C) = 1)$ having no prime factor larger than $(\log H(A, B, C))^\kappa$, for a given finite κ. We show that the *abc* Conjecture implies that for any fixed $\kappa < 1$ the equation has only finitely many primitive solutions. We also discuss a conditional result, showing that the Generalized Riemann hypothesis (GRH) implies that for any fixed $\kappa > 8$ the *abc* equation has infinitely many primitive solutions. We outline a proof of the latter result.

1. Introduction

The *abc* equation is the homogeneous linear ternary Diophantine equation $A + B + C = 0$, usually written $A + B = C$ (replacing C with $-C$). An integer solution (A, B, C) is *nondegenerate* if $ABC \neq 0$, and is *primitive* if $\gcd(A, B, C) = 1$. The Diophantine size of a solution can be measured by the *height* $H(A, B, C)$, given by

$$(1.1) \qquad H(A, B, C) := \max(|A|, |B|, |C|).$$

The well-known *abc* Conjecture (of Masser and Oesterlé [25]) relates the height (1.1) of solutions to the *radical* $R(A, B, C)$ of a solution, which is given by

$$(1.2) \qquad R(A, B, C) = \mathrm{rad}(ABC) := \prod_{p|ABC} p.$$

abc Conjecture (weak form). *There is a positive constant κ_1 such that for any $\epsilon > 0$ there are only finitely many primitive solutions (A, B, C) to the abc equation $A + B = C$ such that*

$$R(A, B, C) \leq H(A, B, C)^{\kappa_1 - \epsilon}.$$

The restriction to primitive solutions is needed to exclude infinite families of imprimitive solutions such as $2^n + 2^n = 2^{n+1}$, which have $R(A, B, C) = 2$ and $H(X, Y, Z) = 2^{n+1}$.

For any individual solution (A, B, C) we define its *abc-exponent* $\kappa_1(A, B, C)$ by

$$\kappa_1(A, B, C) := \frac{\log R(A, B, C)}{\log H(A, B, C)}.$$

Then the maximum allowable exponent in the *abc* Conjecture is given by

$$\kappa_1 := \liminf_{\substack{H(A,B,C) \to \infty \\ A+B=C, gcd(A,B,C)=1}} \kappa_1(A, B, C).$$

It is known that this exponent satisfies $\kappa_1 \leq 1$, and the *abc* Conjecture is often stated in the following strong form, e.g. in Bombieri and Gubler [3, Chap. 12].

abc Conjecture (strong form). *The abc Conjecture holds with $\kappa_1 = 1$, so that for any $\epsilon > 0$ there are only finitely many primitive solutions to $A + B = C$ satisfying*

$$R(A, B, C) \leq H(A, B, C)^{1-\epsilon}.$$

1.1. xyz Conjecture. In this paper we relate the height to another measure of solution size, the *smoothness* $S(A, B, C)$, given by

$$(1.3) \qquad S(A, B, C) := \max\{p : \ p|ABC\}.$$

We study the existence of primitive solutions having minimal smoothness as a function of the height $H(A, B, C)$. For convenience we switch variables from (A, B, C) to (X, Y, Z) and formulate the following conjecture.

xyz Conjecture (weak form-1). *There exists a positive constant κ_0 such that the following hold.*

(a) For each $\epsilon > 0$ there are only finitely many primitive solutions (X, Y, Z) to the equation $X + Y = Z$ with

$$S(X, Y, Z) < (\log H(X, Y, Z))^{\kappa_0 - \epsilon}.$$

(b) For each $\epsilon > 0$ there are infinitely many primitive solutions (X, Y, Z) to to the equation $X + Y = Z$ with

$$S(X, Y, Z) < (\log H(X, Y, Z))^{\kappa_0 + \epsilon}.$$

The restriction to primitive solutions is needed, to exclude the same infinite family $2^n + 2^n = 2^{n+1}$, which has $S(X, Y, Z) = 2$ and $H(X, Y, Z) = 2^{n+1}$.

For any solution (X, Y, Z) we define its *smoothness exponent $\kappa_0(X, Y, Z)$* to be

$$(1.4) \qquad \kappa_0(X, Y, Z) := \frac{\log S(X, Y, Z)}{\log \log H(X, Y, Z)}.$$

Define the *xyz-smoothness exponent*

$$(1.5) \qquad \kappa_0 := \liminf_{\substack{H(X,Y,Z) \to \infty \\ X+Y=Z, \gcd(X,Y,Z)=1}} \kappa_0(X, Y, Z).$$

A priori this exponent exists and satisfies $0 \le \kappa_0 \le +\infty$.

xyz Conjecture (weak form-2). *The xyz-smoothness exponent κ_0 is positive and finite.*

To establish the truth of the weak form of the *xyz* Conjecture, it suffices to show that property (a) holds for some $\kappa_0 > 0$ and that property (b) holds for some $\kappa_0 < \infty$. The fact that properties (a) and (b) are mutually exclusive, and are monotone in κ in appropriate direction, then shows a unique constant κ_0 exists satisfying both (a) and (b). This makes it possible in principle to prove that a nonzero constant κ_0 exists, without determining its exact value.

1.2. Strong form of *xyz* Conjecture. There is a simple heuristic argument which supports the *xyz*-Conjecture, and which suggests that the correct constant might be $\kappa_0 = \frac{3}{2}$. Consider triples (X, Y, Z) where X, Y and Z are pairwise relatively prime, all lie in the interval $[1, H]$ and are all composed of prime factors smaller than $(\log H)^\kappa$. For each such triple $X + Y - Z$ is an integer in the interval $[-H, 2H]$ and if $X + Y - Z$ is "randomly distributed" then we may expect that roughly $1/H$ of the triples (X, Y, Z) will satisfy $X + Y = Z$. We must now count how many triples (X, Y, Z) there are and expect to find a solution to $X + Y = Z$ when there are many more than H such triples, e.g. $H^{1+\epsilon}$ triples, We also expect to

find no solution when there are far fewer than H such triples, i.e. $H^{1-\epsilon}$ triples.

Let $\mathcal{S}(y)$ denote the set of all integers having only prime factors $p \leq y$, Recall that

(1.6)
$$\Psi(x, y) := \#\{n \leq x : x \in \mathcal{S}(y)\}$$

counts the number of such integers below x. It is known that for fixed $\kappa > 1$, one has

(1.7)
$$\Psi(x, (\log x)^\kappa) = x^{1-\frac{1}{\kappa}+o(1)},$$

as $x \to \infty$ (see Tenenbaum [29, Chap. III.5, Theorem 10], [23, Lemma 9.4]). Thus for $\kappa > 1$ the number of such triples (X, Y, Z) with $X, Y, Z \in [1, H]$ is at most $\Psi(H, (\log H)^\kappa)^3 = H^{3(1-\frac{1}{\kappa}+o(1))}$, and if $\kappa < \frac{3}{2}$ this is $< H^{1-\epsilon}$. Thus heuristically we expect few hits in the relatively prime case, which suggests that $\kappa_0 \geq \frac{3}{2}$.

We now give a lower bound for the number of triples (X, Y, Z). Take $X \in [1, H]$ to be a number composed of exactly $K := [\log H/(\kappa \log \log H)]$ distinct primes all below $(\log H)^\kappa$. Using Stirling's formula there are $\binom{\pi((\log H)^\kappa)}{K} = H^{1-1/\kappa+o(1)}$ such values of X all lying below H. Given X, choose $Y \in [1, H]$ to be any number composed of exactly K distinct primes below $(\log H)^\kappa$, but avoiding the primes dividing X. There are $\binom{\pi((\log H)^\kappa)-K}{K} = H^{1-1/\kappa+o(1)}$ such values of Y. Finally choose $Z \in [1, H]$ to be a number composed of exactly K distinct primes below $(\log H)^\kappa$ avoiding the primes dividing X and Y. There are $\binom{\pi((\log H)^\kappa)-2K}{K} = H^{1-1/\kappa+o(1)}$ such values of Z. We conclude therefore that there are at least $H^{3(1-1/\kappa+o(1))}$ desired triples, and hence we expect that $\kappa_0 \leq \frac{3}{2}$.

We therefore formulate the following conjecture.

xyz Conjecture (strong form). *The xyz-smoothness exponent* $\kappa_0 = \frac{3}{2}$.

In analogy with the *abc* Conjecture, it is convenient to define the *xyz-quality* $Q^*(X, Y, Z)$ of an *xyz* triple (X, Y, Z) to be

(1.8)
$$Q^*(X, Y, Z) := \frac{3}{2}\left(\frac{\log \log H(X, Y, Z)}{\log S(X, Y, Z)}\right).$$

Since

$$Q^*(X, Y, Z) = \frac{3}{2\kappa_0(X, Y, Z)},$$

the strong form of the *xyz*-Conjecture asserts that

(1.9)
$$\limsup_{H(X,Y,Z)\to\infty} Q^*(X, Y, Z) = 1.$$

Thus triples (X, Y, Z) with *xyz*-quality exceeding 1 are exceptionally good. Unlike the *abc*-Conjecture case, we do not know if there exist infinitely many relatively prime triples (X, Y, Z) having quality exceeding 1.

S(X,Y,Z)	X	Y	Z	$\kappa_0(X,Y,Z)$	$Q^*(X,Y,Z)$
3	1	8	9	1.39560	1.07480
5	3	125	128	1.01902	1.47200
7	1	4374	4375	0.91517	1.63904
11	3584	14641	18225	1.05011	1.42841
13	91	1771470	1771561	0.96197	1.55930

TABLE 1.1. Extremal solutions having $S(X,Y,Z) \leq 13$

1.3. Examples.

Example 1. Take $(X,Y,Z) = (1, 2400, -2401)$. Then $Y = 2^5 \cdot 3 \cdot 5^2, |Z| = 7^4$. Here the height $H(X,Y,Z) = 2401$ while the smoothness $S(X,Y,Z) = 7$. We have $\log H(X,Y,Z) = 7.78364$ so this example has smoothness exponent

$$\kappa_0(X,Y,Z) := \frac{\log S(X,Y,Z)}{\log \log H(X,Y,Z)} \approx 0.94829$$

This triple is exceptionally good, having *xyz*-quality

$$Q^*(X,Y,Z) := \frac{3}{2}\left(\frac{\log \log H(X,Y,Z)}{\log S(X,Y,Z)}\right) \approx 1.58180.$$

Example 2. de Weger [31, Theorem 5.4] found the complete set of primitive solutions to the *xyz* equation having $S(X,Y,Z) \leq 13$; there are 545 such solutions. His table of large solutions ([31, Table IX]) yields the extremal values given above in Table 1.1. His values include the current record for smallest value of smoothness exponent, which is $\kappa_0 \approx 0.91517$ for $(1, 4374, -4375)$.

Example 3. Consider the elliptic modular function

$$j(\tau) = \frac{1}{q} = 744 + 196884q + 21493760q^2 + ...$$

where $q = e^{2\pi i \tau}$ with $Im(\tau) > 0$ so $|q| < 1$. A *singular modulus* is a value $j(\tau)$ where τ is an algebraic integer in an imaginary quadratic number field $K = \mathbb{Q}(\sqrt{-d})$; each such τ corresponds to an elliptic curve with complex multiplication by an order in the field K. The theory of complex multiplication then says that $j(\tau)$ is then an algebraic integer lying in an abelian extension of K. Gross and Zagier [13, Theorem 1.3] observe that the norms of differences of singular moduli factorize into products of small primes. Thus, we may consider equations

$$(1.10) \qquad (j(\tau_1) - j(\tau_2)) + (j(\tau_2) - j(\tau_3)) + (j(\tau_3) - j(\tau_1)) = 0$$

as giving candidate triples (X,Y,Z) of algebraic integers $X+Y+Z = 0$ that are smooth in their generating number field $K = \mathbb{Q}(X,Y,Z)$, in the sense that the prime ideals dividing XYZ all have small norms. (These triples

were noted in Granville and Stark [12, Sec. 4.2] in connection with the *abc* Conjecture.) In special cases where all the values $j(\tau_i)$ are rational integers, which occur for example when each of $\{\tau_k : 1 \leq k \leq 3\}$ correspond to elliptic curves with CM by an order of an imaginary quadratic field having class number one, then (1.10) gives interesting examples of integer triples with small smoothness exponents. Using ([13, p. 193]) we have $j(\frac{-1+i\sqrt{3}}{2}) = 0$, and

$$j\Big(\frac{1+i\sqrt{67}}{2}\Big) - j\Big(\frac{-1+i\sqrt{3}}{2}\Big) = -2^{15} \cdot 3^3 \cdot 5^3 \cdot 11^3$$

$$j\Big(\frac{-1+i\sqrt{3}}{2}\Big) - j\Big(\frac{1+i\sqrt{163}}{2}\Big) = 2^{18} \cdot 3^3 \cdot 5^3 \cdot 23^3 \cdot 29^3$$

$$j\Big(\frac{1+i\sqrt{163}}{2}\Big) - j\Big(\frac{1+i\sqrt{67}}{2}\Big) = -2^{15} \cdot 3^7 \cdot 5^3 \cdot 7^2 \cdot 13 \cdot 139 \cdot 331.$$

After removing common factors we obtain the relatively prime triple

$$(-11^3, \ 2^3 \cdot 23^3 \cdot 29^3, -3^4 \cdot 7^2 \cdot 13 \cdot 139 \cdot 331)$$
$$= (-1331, \ 2373927704, \ -2373926373).$$

This has height $H = 2373927704$ and smoothness $S = 331$, whence

$$\kappa_0(X, Y, Z) := \frac{\log S(X, Y, Z)}{\log \log H(X, Y, Z)} \approx 1.88863.$$

Its xyz-quality $Q^*(X, Y, Z) \approx 0.79422$. Here the radical $R(X, Y, Z) = 2 \cdot 3 \cdot 7 \cdot 11 \cdot 13 \cdot 23 \cdot 29 \cdot 139 \cdot 331 = 184312146018$, which is larger than $H(X, Y, Z)$.

Example 4. Consider the Diophantine equation $x^2 + y^3 = z^7$. Recently Poonen, Schaefer and Stoll [26] proved that, up to signs, it has exactly eight primitive integer solutions, and that the largest of these is $(x, y, z) = (15312283, 9262, 113)$. Here $15312283 = 7 \cdot 53 \cdot 149 \cdot 277$ and $9262 = 2 \cdot 11 \cdot 421$. Thus for $(X, Y, Z) = (x^2, y^3, z^7)$ we have $H = (113)^7$ and $S = 421$, whence $\kappa_0(X, Y, Z) \approx 1.72682$ and $Q^*(X, Y, Z) \approx 0.86864$.

Example 5. The Monster simple group has order $M = 2^{45} \cdot 3^{20} \cdot 5^9 \cdot 7^6 \cdot 11^2 \cdot 13^3 \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 47 \cdot 59 \cdot 71$. Its three smallest irreducible representations are the identity representation (of degree 1) plus representations of degrees $196883 = 47 \cdot 59 \cdot 71$ and $21296876 = 2^2 \cdot 31 \cdot 41 \cdot 59 \cdot 71$. Here we note the triple $(X, Y, Z) = (196882, 1, 196883)$ has $196882 = 2 \cdot 7^4 \cdot 41$, and $S(X, Y, Z) = 71$; additionally the triple $(21296875, 1, 21296876)$ has $21296875 = 5^6 \cdot 29 \cdot 47$, and $S(X, Y, Z) = 71$. Both these triples have all three terms dividing M. They have smoothness exponents $\kappa_0(196882, 1, 196883) \approx 1.70463$ and $\kappa_0(21296875, 1, 21296876) \approx 1.50849$, respectively.

1.4. Main results. The main results presented here are conditional results, concerning both parts of the xyz Conjecture.

In §2 we show that a nonzero lower bound on κ_0 follows from the *abc* Conjecture. Namely, we show the weak form of the *abc* Conjecture implies the lower bound $\kappa_0 \geq \kappa_1$.

Theorem 1.1. *If the weak form of the abc-Conjecture holds with exponent κ_1, then for each $\epsilon > 0$ there are only finitely many primitive solutions (X, Y, Z) to $X + Y = Z$ such that*

$$S(X, Y, Z) < (\log H(X, Y, Z))^{\kappa_1 - \epsilon}.$$

Thus the xyz-smoothness exponent κ_0 satisfies $\kappa_0 \geq \kappa_1$. In particular, the strong form of the abc Conjecture implies

$$\kappa_0 \geq 1.$$

The simple deduction of Theorem 1.1 is given in §2. We also give an unconditional result (Theorem 2.2), namely that there are only finitely many primitive solutions to

$$S(X, Y, Z) < (3 - \epsilon) \log \log H(X, Y, Z).$$

For the upper bound in the *xyz* Conjecture, we derive an upper bound in our paper [23], assuming appropriate Riemann hypotheses. We assume the Generalized Riemann hypothesis (GRH) in the following form:

Generalized Riemann Hypothesis (GRH**).** *The Riemann zeta function and all Dirichlet L-functions have all of their zeros in the critical strip $0 < Re(s) < 1$ lying on the critical line $Re(s) = \frac{1}{2}$.*

The main result of [23] is that the GRH implies the upper bound part of the *xyz*-Conjecture, establishing that $\kappa_0 \leq 8$.

Theorem 1.2. *If the Generalized Riemann Hypothesis (GRH) holds, then for each $\epsilon > 0$ there are infinitely many primitive solutions (X, Y, Z) to $X + Y + Z = 0$ such that*

$$S(X, Y, Z) < (\log H(X, Y, Z))^{8+\epsilon}.$$

In particular, the xyz-smoothness exponent satisfies $\kappa_0 \leq 8$.

We discuss this result and its proof in §3 and §4. Theorem 1.2 is a consequence of a stronger result, stated as Theorem 3.1 below, which gives a lower bound for the number of weighted primitive solutions to the equation, for fixed $\kappa > 8$, which is the correct order of magnitude according to the heuristic argument above. This result is proved by a variant of the Hardy-Littlewood method, combined with the Hildebrand-Tenenbaum saddle-point method for counting numbers all of whose prime factors are small. The generalized Riemann hypothesis is invoked mainly to control the minor arcs estimates. The Hardy-Littlewood method detects all integer solutions, and an inclusion-exclusion argument is needed at the end of the proof to count primitive integer solutions only.

Combining Theorems 1.1 and 1.2 above yields the following result conditionally proving the xyz Conjecture.

Theorem 1.3. (Alphabet Soup Theorem) *The weak form of the abc-Conjecture together with the GRH implies the weak form of the xyz Conjecture.*

This conditional result does not establish the strong form of the xyz Conjecture, since it only shows $\kappa_1 \le \kappa_0 \le 8$. Even if we assume the strong form of the abc Conjecture, we deduce only $\kappa_0 \ge 1$ which is weaker than the lower bound of the strong xyz Conjecture.

1.5. S-unit equations. The S-unit equation is $X + Y = Z$, subject to the constraint that all prime factors of XYZ belong to a fixed finite set of primes S. In 1988 Erdős, Stewart and Tijdeman [11] showed the existence of collections of primes S with $|S| = s$ such that the S-unit equation $X + Y = Z$ has "exponentially many" solutions, namely at least $\exp\left((4 - \epsilon)s^{\frac{1}{2}}(\log s)^{-\frac{1}{2}}\right)$ solutions, for $s \ge s_0(\epsilon)$ sufficiently large. Recently Konyagin and Soundararajan [22] improved this construction, to show that there exist S such that the S-unit equation has at least $\exp\left(s^{2-\sqrt{2}-\epsilon}\right)$ solutions.

In these constructions the sets of primes S were tailored to have large numbers of solutions, which entailed losing control over the relative sizes of primes in S. However it is natural to consider the special case where S is an inital segment of primes

$$(1.11) \qquad S_y = \mathcal{P}(y) := \{p : p \text{ prime}, \ p \le y\}.$$

Here Erdős, Stewart and Tijdeman conjectured ([11, p. 49, top]) a stronger exponential bound, which asserts that for $s = |S_y|$ and each $\epsilon > 0$ there should be at least $\exp(s^{\frac{2}{3}-\epsilon})$ S-unit solutions to $X + Y = Z$ and at most $\exp(s^{\frac{2}{3}+\epsilon})$ such solutions, for all $s > s_0(\epsilon)$.

The xyz Conjecture can be reformulated as an assertion about the maximal relatively prime solutions to an S-unit equation for $S = S_y$, as $y \to \infty$.

xyz Conjecture (weak form-3). *Take $S = S_y$ and let H_y denote the maximal height of any relatively prime solution to this S-unit equation. Then the constant*

$$(1.12) \qquad \kappa_0 := \liminf_{y \to \infty} \ \frac{\log y}{\log \log H_y}.$$

is positive and finite.

It is known that the complete set of solutions to an S-unit equation for fixed S can be determined effectively, and therefore H_y is effectively computable in principle (but only for small y in practice).

The conjecture of Erdős, Stewart and Tijdeman for the number of solutions to the S-unit equation for $S = S_y$ would imply that $\kappa_0 \geq \frac{3}{2}$. In fact the exponent $\frac{2}{3}$ conjectured by Erdős, Stewart and Tijdeman [11] is supported by a similar heuristic to that giving $\kappa_0 = \frac{3}{2}$ in the strong form of the *xyz* Conjecture.

We obtain the following result in the direction of the conjecture of Erdős, Stewart and Tijdeman, as an easy consequence of the quantitative bound given in Theorem 3.1 later in this paper.

Theorem 1.4. *Let $S = S_s^*$ denote the set of the first s primes, and let $N(S)$ count the number of primitive solutions (X, Y, Z) to the S-unit equation $X + Y = Z$. If the Generalized Riemann Hypothesis is valid, then for each $\epsilon > 0$,*

$$N(S_s^*) \gg_\epsilon \exp(s^{\frac{1}{8} - \epsilon}).$$

This is proved as Theorem 1.3 in [23].

1.6. Discussion.

We remark first on earlier work related to the approach used in Theorem 1.2. The Hardy-Littlewood method was used in 1984 by Balog and Sarkozy [1], [2] to show that the equation $X + Y + Z = N$ has solutions with $S(X, Y, Z) \leq \exp(3\sqrt{\log N \log \log N})$ for all sufficiently large N. Exponential sums taken over smooth numbers below x with all prime factors smaller than y were studied in detail by de la Bretèche ([4], [5]) and de la Bretèche and Tenenbaum ([7], [8], [9]), see also de la Bretèche and Granville [6]. Their various bounds are unconditional, and are valid for sufficiently large y, requiring at least $y \geq \exp(c(\log \log x)^2)$. Thus their results fail to apply in the range we need, which is $y = (\log x)^\kappa$ for fixed κ.

In the range of interest, it is a delicate problem even to count the number of y-smooth integers up to x. General results on $\Psi(x, y)$ are reviewed in Hildebrand and Tenenbaum [21]. In [23] we make use of the saddle-point method developed by Hildebrand and Tenenbaum [20] to estimate $\Psi(x, y)$. They showed that it provides an asymptotic formulas for $\Psi(x, y)$ when y is not too small, including a result of Hildebrand that for $y \geq \exp((\log \log x)^{\frac{5}{3} + \epsilon})$,

$$(1.13) \qquad \Psi(x, y) = x\rho(u)\Big(1 + O_\epsilon\Big(\frac{u \log(u + 1)}{\log x}\Big)\Big).$$

where u is defined by $y = x^{\frac{1}{u}}$ and $\rho(u)$ is Dickman's function. The range $y = (\log x)^\kappa$ for fixed κ lies outside the range covered by Hildebrand's (1.13), and in this range, the behavior of $\Psi(x, y)$ is known to be sensitive to the fine distribution of primes and location of the zeros of $\zeta(s)$. In 1984 Hildebrand [18] showed that the Riemann hypothesis is equivalent to the

assertion that for each $\epsilon > 0$ and $1 \le u \le y^{1/2-\epsilon}$ there is a uniform estimate

$$\Psi(x, y) = x\rho(u) \exp(O_\epsilon(y^\epsilon)).$$

Moreover, assuming the Riemann hypothesis, he showed that for each $\epsilon > 0$ and $1 \le u \le y^{1/2-\epsilon}$ the stronger uniform estimate

$$\Psi(x, y) = x\rho(u) \exp\left(O_\epsilon\left(\frac{\log(u+1)}{\log y}\right)\right)$$

holds. On choosing $y = (\log x)^\kappa$ for $\kappa > 2$, this latter estimate yields

$$\Psi(x, (\log x)^\kappa) \asymp x\rho(u),$$

which provides only an order of magnitude estimate for the size of $\Psi(x, y)$. Furthermore if the Riemann hypothesis is false then $\Psi(x, y)$ must sometimes exhibit large oscillations away from the value $x\rho(u)$ for some (x, y) in these ranges. In 1986 Hildebrand [19] obtained further results indicating that when $y < (\log x)^{2-\epsilon}$ one should not expect any smooth asymptotic formula for $\Psi(x, y)$ in terms of the y-variable to hold. We assume GRH, and because of these issues we retain $\Psi(x, y)$ in our final formulas, avoiding any asymptotic approximation to it (see Theorem 4.1 and 4.2). Note that the threshold value $\kappa = 2$ is relevant in the formula in Theorem 3.1, see (3.4) below.

Concerning the relative strength of the abc Conjecture versus the xyz Conjecture, the difficulty of the abc Conjecture lies entirely in the lower bound. Indeed it has the trivial upper bound $\kappa_1 \le 3$, and also has the unconditional upper bound $\kappa_1 \le 1$. The xyz Conjecture presents difficulties in both the upper bound and the lower bound, and we have no unconditional result for either bound. Concerning the lower bound, we do not know any implication from the xyz Conjecture in the direction of the abc Conjecture. Thus establishing a nontrivial lower bound (a) of the xyz Conjecture could be an easier problem than the abc-Conjecture. We note the coincidence that the abc implication that $\kappa_0 \ge 1$ lower bound occurs exactly at the threshold value of $\kappa = 1$ at which the behavior of the smooth number counting function $\Psi(x, x^\kappa)$ qualitatively changes. For $\kappa > 1$ one has $\Psi(x, x^\kappa) = x^{1-\frac{1}{\kappa}+o(1)}$, while for $0 < \kappa < 1$ one has $\Psi(x, x^\kappa) = x^{o(1)}$.

It remains an interesting question whether the GRH assumption can be removed to obtain an unconditional upper bound on κ_0, possibly worse than the conditional upper bound $\kappa_0 \le 8$ obtained in [23]. This problem does not seem entirely out of reach. We are not aware of any approach that might yield a positive unconditional lower bound for κ_0.

The plan of the paper is as follows. In Section 2 we establish the lower bound for the xyz constant κ_0 assuming the abc Conjecture. In Section 3 we discuss the main upper bound obtained in [23] for the number of smooth

solutions when $\kappa_0 > 8$, given as Theorem 3.1, and discuss conjectures concerning asymptotics of the number of primitive solutions versus all solutions for $\kappa > 2$. Section 4 presents results from [23] giving asymptotic formulas counting weighted smooth solutions, from which Theorem 3.1 is derived, and in Section 5 we give some ideas of their proofs. In Section 6 we make concluding remarks, indicating directions for extending the results.

2. Lower bound assuming the *abc* Conjecture

The *abc* Conjecture was formulated by Masser and Oesterlé in 1985, cf. Stewart and Tijdeman [28], Oesterlé [25]. It has many equivalent formulations, see Bombieri and Gubler [3, Chap.XII].

Theorem 2.1. *Assume the weak form of the abc Conjecture holds. Then for any $\epsilon > 0$ there are only finitely many primitive solutions (X, Y, Z) to the equation $X + Y = Z$ such that*

$$S(X, Y, Z) \le (\kappa_1 - \epsilon) \log H(X, Y, Z).$$

In particular the xyz-smoothness exponent κ_0 satisfies $\kappa_0 \ge \kappa_1$.

Proof. Let $X + Y = Z$. The radical $R(X, Y, Z)$ satisfies

$$(2.1) \qquad R(X, Y, Z) = \mathrm{rad}(XYZ) = \prod_{p \mid XYZ} p \le \prod_{p \le S(X,Y,Z)} p$$

Now the prime number theorem implies that

$$\prod_{p \le y} p = e^{y + o(y)} \text{ as } y \to \infty,$$

so that

$$R(X, Y, Z) \le e^{S(X,Y,Z)(1 + o(1))}$$

The hypothesis that a solution has smoothness satisfying

$$S(X, Y, Z) \le (\kappa_1 - \epsilon) \log H(X, Y, Z)$$

now yields

$$(2.2) \qquad R(X, Y, Z) \le H(X, Y, Z)^{\kappa_1 - \epsilon + o(1)} \text{ as } H \to \infty.$$

Thus such a solution must satisfy

$$(2.3) \qquad R(X,Y,Z) \le H(X,Y,Z)^{\kappa_1 - \frac{1}{2}\epsilon},$$

if its height is sufficiently large, say $H \ge H_0$. The weak form of the abc Conjecture asserts that the inequality (2.3) has only finitely many primitive integer solutions (X,Y,Z), so the result follows. $\qquad\square$

Proof of Theorem 1.1. This follows immediately from Theorem 2.1. $\qquad\square$

Following the same proof as Theorem 2.1 above, but employing the current best bound on the abc equation, yields the following unconditional result.

Theorem 2.2. *For each $\epsilon > 0$ there are only finitely many primitive solutions (X,Y,Z) to $X + Y + Z = 0$ such that*

$$(2.4) \qquad S(X,Y,Z) \le (3-\epsilon)\log\log H(X,Y,Z).$$

Proof. The best unconditional bound in the direction of the abc Conjecture is currently that of Stewart and Yu [27]. They showed there is a constant c_1 such that all primitive solutions to the abc equation satisfy

$$H(X,Y,Z) \le \exp(c_1 R^{\frac{1}{3}}(\log R)^3),$$

where $R = R(X,Y,Z)$. Taking logarithms yields the constraint

$$(2.5) \qquad (\log H(X,Y,Z))^3 \le (c_1)^3 R(X,Y,Z)(\log R(X,Y,Z))^9.$$

Suppose that an abc solution (X,Y,Z) satisfies the bound (2.4). There are only finitely many abc solutions (X,Y,Z) having $S = S(X,Y,Z)$ taking a given value, since the abc-equation is then an S-unit equation, with S being all primes up to the given value. Thus by excluding a finite set of solutions we may assume $S \ge S_0$ for any fixed S_0. If $S = S(X,Y,Z)$ is sufficiently large (depending on ϵ) then the $o(1)$ term in (2.2) can be replaced by $\frac{1}{6}\epsilon$ and we obtain

$$R(X,Y,Z) \le e^{S(X,Y,Z)(1+\frac{1}{6}\epsilon)} \le e^{(3-\frac{1}{2}\epsilon)\log\log H} = (\log H)^{3-\frac{1}{2}\epsilon}.$$

Substituting this bound in (2.5) shows this solution must satisfy

$$(\log H)^3 \le (c_1)^3 (\log H)^{3-\frac{1}{2}\epsilon}(3\log\log H)^9.$$

However for all sufficiently large H, the right side of this inequality is smaller than $(\log H)^3$, giving a contradiction. We conclude that all such solutions have bounded H, and the finiteness result follows. $\square$

3. Upper bound assuming GRH

Let $\mathcal{S}(y)$ denote the set of all integers having all prime factors $\leq y$. We define the counting function $N(H, \kappa)$ of all smooth solutions $X + Y = Z$ to height H by

$$N(H, \kappa) := \#\{X + Y = Z : 0 \leq X, Y, Z \leq H, \max_{p|XYZ}\{p\} \leq (\log H)^\kappa\},$$

and the counting function $N^*(H, \kappa)$ of all primitive integer solutions to height H by

$$N^*(H, \kappa) := \# \{X + Y = Z \text{ primitive} : (X, Y, Z) \in N(H, \kappa)\}.$$

A main result of our paper [23] is the following lower bound for the number of primitive integer solutions for $\kappa > 8$ ([23, Theorem 1.3]).

Theorem 3.1. (Counting primitive smooth solutions) *Assume the truth of the Generalized Riemann Hypothesis (GRH). Then for each fixed $\kappa > 8$, the number of primitive integer solutions to $X + Y = Z$ below H satisfies*

(3.1)
$$N^*(H, \kappa) \geq \mathfrak{S}_\infty\left(1 - \frac{1}{\kappa}\right)\mathfrak{S}_f^*\left(1 - \frac{1}{\kappa}, (\log H)^\kappa\right)\frac{\Psi(H, (\log H)^\kappa)^3}{H}(1 + o(1)),$$

as $H \to \infty$. Here the "archimedean singular series" (more properly, "singular integral") is defined, for $c > \frac{1}{3}$, by

(3.2)
$$\mathfrak{S}_\infty(c) := c^3 \int_0^1 \int_0^{1-t_1} (t_1 t_2 (t_1 + t_2))^{c-1} dt_1 dt_2,$$

and the "primitive non-archimedean singular series" $\mathfrak{S}_f^*(c, y)$ is defined by

(3.3) $\mathfrak{S}_f^*(c, y) := \prod_{p \leq y} \left(1 + \frac{1}{p^{3c-1}}\left(\frac{p}{p-1}\left(\frac{p - p^c}{p-1}\right)^3 - 1\right)\right) \prod_{p > y} \left(1 - \frac{1}{(p-1)^2}\right).$

This result is derived using a variant of the Hardy-Littlewood method together with the Hildebrand-Tenenbaum saddle point method for counting smooth numbers. The method is first applied to integer solutions weighted using a smooth compactly-supported test function in the open interval $\mathbb{R}_{>0} = (0, \infty)$, where one can obtain asymptotic formulas, described in Section 4, for both integer solutions and for primitive integer solutions. We obtain only a one-sided inequality in (3.1) because the weight function corresponding to (3.1) is not compactly supported, but can be approximated from below by such functions to give a lower bound.

We discuss the terms appearing on the right side of this formula. Since for $\kappa > 1$, we have $\Psi(H, (\log H)^\kappa) = H^{1 - \frac{1}{\kappa} + o(1)}$ as $H \to \infty$, we see that the "main term" on the right side of the formula (3.1) is essentially of the form $H^{2 - \frac{3}{\kappa} + o(1)}$ given by the heuristic of section 1.2, up to the value of the constants in the "singular series". We believe that equality should actually

hold in (3.1) in the indicated range of H and κ, thus giving an asymptotic formula for $\kappa > 8$.

Now we consider the "main term" on the right side of (3.1) for smaller values of κ. It is well-defined in the range $\kappa > \frac{3}{2}$ where the heuristic in section 1.2 is expected to apply. Here $\kappa > \frac{3}{2}$ corresponds to $c > \frac{1}{3}$, and the "archimedean singular integral" (3.2) defines an analytic function on the half-plane $Re(c) > \frac{1}{3}$ which diverges at $c = \frac{1}{3}$, while the "non-archimedean singular series" $\mathfrak{S}_f(c, y)$ is well-defined for all $c > 0$. The archimedean singular series is uniformly bounded on any half-plane $Re(c) > \frac{1}{3} + \epsilon$. For the non-archimedean singular series, we find that its limiting behavior as $y = (\log H)^\kappa \to \infty$ changes at the threshold value $\kappa = 2$, corresponding to $c = \frac{1}{2}$. Namely, one has

$$(3.4) \qquad \lim_{H \to \infty} \mathfrak{S}_f^*\Big(1 - \frac{1}{\kappa}, (\log H)^\kappa\Big) = \begin{cases} \mathfrak{S}_f^*(1 - \frac{1}{\kappa}) & \text{for } \kappa > 2, \\ \\ 0 & \text{for } 0 < \kappa \le 2, \end{cases}$$

where for $c > \frac{1}{2}$ we set

$$(3.5) \qquad \mathfrak{S}_f^*(c) := \prod_p \Big(1 + \frac{1}{p^{3c-1}} \Big(\frac{p}{p-1}\Big(\frac{p - p^c}{p-1}\Big)^3 - 1\Big)\Big).$$

The Euler product (3.5) converges absolutely and defines an analytic function $\mathfrak{S}_f^*(c)$ on the half-plane $Re(c) > \frac{1}{2}$; this function is uniformly bounded on any half-plane $Re(c) > \frac{1}{2} + \epsilon$. Furthermore for values corresponding the $2 < \kappa < \infty$ (i.e. $\frac{1}{2} < c < 1$) the "non-archimedean singular series" $\mathfrak{S}_f^*(c, y)$ remains bounded away from 0. We conclude that for $2 < \kappa < \infty$ the "main term" estimate for $N^*(H, \kappa)$ agrees with the prediction of the heuristic argument given earlier. In the remaining region $1 < \kappa \le 2$ one can show that although $\mathfrak{S}_f^*(1 - \frac{1}{\kappa}, (\log H)^\kappa) \to 0$ as $H \to \infty$, one has

$$\mathfrak{S}_f^*\Big(1 - \frac{1}{\kappa}, (\log H)^\kappa\Big) \gg \exp\Big(-(\log H)^{2-\kappa}\Big).$$

This bound implies that $\mathfrak{S}_f^*(1 - \frac{1}{\kappa}, (\log H)^\kappa) \gg H^{-\epsilon}$ for any $\epsilon > 0$, which for $\frac{3}{2} < \kappa \le 2$ shows the "main term" is still of the same order $H^{2 - \frac{3}{\kappa} + o(1)}$ as the heuristic predicts. Thus it could still be the case that the "main term" on the right side of (3.2) gives a correct order of magnitude estimate for $N^*(H, \kappa)$ in this range.

In terms of limiting behavior as $H \to \infty$ there appears to be a second threshold value at $\kappa = 3$, which concerns the relative density of primitive smooth solutions in the set of all smooth solutions. We formulate this as the following conjecture.

Conjecture 3.1. (Relative density Conjecture) *The relative density of primitive smooth solutions satisfies*

$$(3.6) \qquad \lim_{H \to \infty} \frac{N^*(H, \kappa)}{N(H, \kappa)} = \begin{cases} \zeta(2 - \frac{3}{\kappa}), & \text{for } 3 < \kappa < \infty, \\ \\ 0 & \text{for } 1 < \kappa \leq 3. \end{cases}$$

Assuming GRH, our paper [23] shows that a weighted version of this conjecture holds for $\kappa > 8$, given here as Theorem 4.3. Furthermore for each $\kappa > 3$ the ratios of the conjectured "main terms" in the asymptotic formulas for these quantities have this limiting value, a result implied by (4.8) below.

We remark that the number $N(H, \kappa)$ of all smooth integer solutions already has a contribution from smooth multiples of the solution $(X, Y, Z) = (1, 1, 2)$ that gives

$$(3.7) \qquad N(H, \kappa) \geq \Psi(\tfrac{1}{2}H, (\log H)^{\kappa}) \geq H^{1 - \frac{1}{\kappa} + o(1)}, \text{ as } H \to \infty.$$

For $1 \leq \kappa < 2$ this lower bound exceeds the heuristic argument estimate $H^{2 - \frac{3}{\alpha} + o(1)}$ for $N^*(H, \kappa)$ by a positive power of H. This fact indicates that the heuristic in Section 1.2 fails for $N(H, \kappa)$ for $1 < \kappa < 2$, and supports the truth of Conjecture 3.1 in this range.

4. Asymptotic formulas for weighted smooth solutions

We now describe results giving asymptotic formulas that count weighted smooth solutions, when $\kappa > 8$. Let $\Phi(x) \in C_c^{\infty}(\mathbb{R}_{>0})$ be a smooth compactly supported (real-valued) function on the positive real axis. We define the weighted sum

$$(4.1) \qquad N(x, y; \Phi) := \sum_{\substack{X, Y, Z \in \mathcal{S}(y) \\ X + Y = Z}} \Phi\left(\frac{X}{x}\right) \Phi\left(\frac{Y}{x}\right) \Phi\left(\frac{Z}{x}\right),$$

which counts both primitive and imprimitive solutions. We also set

$$(4.2) \qquad N^*(x, y; \Phi) := \sum_{\substack{X, Y, Z \in \mathcal{S}(y) \\ X + Y = Z, \gcd(X, Y, Z) = 1}} \Phi\left(\frac{X}{x}\right) \Phi\left(\frac{Y}{x}\right) \Phi\left(\frac{Z}{x}\right),$$

which counts only primitive solutions. The asymptotic estimates below are obtained for a general Φ. For applications one may take $\Phi(x)$ to be a non-negative real-valued "bump function" supported on $[0, 1]$, which is equal to 1 on $[\epsilon, 1 - \epsilon]$ and is equal to 0 outside $[\frac{\epsilon}{2}, 1 - \frac{\epsilon}{2}]$. This choice of weight function will ensure that we essentially count solutions $X + Y = Z$ in y-smooth numbers restricted to the range $[1, x]$.

In [23, Theorem 2.1] we obtain the following asymptotic formula, valid for $\kappa > 8$, which counts both primitive and imprimitive weighted integer solutions.

Theorem 4.1. (Counting weighted smooth integer solutions) *Assume the truth of the GRH. Let $\Phi(x)$ be a fixed smooth compactly supported, real-valued weight function in $C_c^\infty(\mathbb{R}^+)$. Let x and y be large, with $(\log x)^{8+\delta} \leq y \leq \exp((\log x)^{1/2-\delta})$ for some $\delta > 0$. Define κ by the relation $y = (\log x)^\kappa$. Then, we have*

$$(4.3)\ \ N(x, y; \Phi) = \mathfrak{S}_\infty\Big(1 - \frac{1}{\kappa}, \Phi\Big)\mathfrak{S}_f\Big(1 - \frac{1}{\kappa}, y\Big)\frac{\Psi(x,y)^3}{x} + O_\eta\left(\frac{\Psi(x,y)^3}{x \log y}\right),$$

Here the "weighted archimedean singular series" (more properly, "singular integral") is defined by

$$(4.4)\ \ \mathfrak{S}_\infty(c, \Phi) := c^3 \int_0^\infty \int_0^\infty \Phi(t_1)\Phi(t_2)\Phi(t_1 + t_2)t_1 t_2(t_1 + t_2)^{c-1} dt_1 dt_2,$$

and the "non-archimedean singular series" $\mathfrak{S}_f(c, y)$ is defined by

$$(4.5)\ \ \ \mathfrak{S}_f(c, y) := \prod_{p \leq y}\Big(1 + \frac{p - 1}{p(p^{3c-1} - 1)}\Big(\frac{p - p^c}{p - 1}\Big)^3\Big) \prod_{p > y}\Big(1 - \frac{1}{(p - 1)^2}\Big).$$

The compact support of $\Phi(x)$ away from 0 guarantees that the "weighted archimedean singular series" $\mathfrak{S}_\infty(c, \Phi)$ is defined for all real c. The "non-archimedean singular series" $\mathfrak{S}_f(c, y)$ is given by an Euler product which converges to an analytic function for $Re(c) > \frac{1}{3}$ and diverges at $c = \frac{1}{3}$. It has a phase change in its behavior as $y \to \infty$ at the threshold value $c = \frac{2}{3}$ corresponding to $\kappa = 3$. Namely, we have

$$(4.6)\ \ \ \ \ \ \ \ \lim_{y \to \infty} \mathfrak{S}_f\Big(1 - \frac{1}{\kappa}, y\Big) = \begin{cases} \mathfrak{S}_f\Big(1 - \frac{1}{\kappa}\Big) & \text{for } \kappa > 3, \\ \\ +\infty & \text{for } 0 < \kappa \leq 3, \end{cases}$$

where for $c > \frac{2}{3}$ we define

$$(4.7)\ \ \ \ \ \ \ \ \ \ \ \mathfrak{S}_f(c) := \prod_p \Big(1 + \frac{p - 1}{p(p^{3c-1} - 1)}\Big(\frac{p - p^c}{p - 1}\Big)^3\Big).$$

The Euler product (4.7) converges absolutely to an analytic function of c on the half-plane $Re(c) > \frac{2}{3}$, and diverges at $c = \frac{2}{3}$. Outside this half-plane, on the range $\frac{1}{2} < c \leq \frac{2}{3}$, although one has $\mathfrak{S}_f(1 - \frac{1}{\kappa}, y) \to \infty$ as $y \to \infty$, one can show that

$$\mathfrak{S}_f\Big(1 - \frac{1}{\kappa}, y\Big) \ll \exp\Big(y^{3/\kappa - 1}\Big).$$

This implies that for $2 < \kappa \leq 3$ one has $\mathfrak{S}_f(1 - \frac{1}{\kappa}, (\log H)^\kappa) \ll H^\epsilon$ for any positive ϵ, which suggests that the heuristic argument of section 1.2 may continue to apply to $N(H, \kappa)$ on this range.

For all $\Phi(x) \in C_c^\infty(\mathbb{R}_{>0})$ the "weighted archimedean singular series" (4.4) defines an entire function of c. For the special weight function $\Phi(x)$ being

the step function $\chi_{[0,1]}(x)$ which is 1 on $[0,1]$, and 0 elsewhere, the "weighted archimedean singular series" integral (4.4) becomes the "archimedean singular series" given in (3.2). This weight function is not compactly supported on $(0,\infty)$ so is not covered by Theorem 4.1. Indeed (3.2) defines an analytic function on the half-plane $Re(c) > \frac{1}{3}$, which diverges when approaching $c = \frac{1}{3}$.

To obtain a bound for primitive solutions, for a given range of y, we perform an inclusion-exclusion argument. The following result applies to arbitrary compactly supported smooth test functions ([23, Theorem 2.2]).

Theorem 4.2. (Counting weighted primitive integer solutions) *Assume the truth of the GRH. Let $\Phi(x)$ be a fixed, smooth, compactly supported, real valued function in $C_c^\infty(\mathbb{R}^+)$, Choose any $\delta > 0$ and let x and y be large with $(\log x)^{8+\delta} \le y \le \exp((\log x)^{1/2-\delta})$. Define κ by the relation $y = (\log x)^\kappa$. Then, we have*

$$N^*(x,y;\Phi) = \mathfrak{S}_\infty\left(1-\frac{1}{\kappa},\Phi\right)\mathfrak{S}_f^*\left(1-\frac{1}{\kappa},y\right)\frac{\Psi(x,(\log x)^\kappa)^3}{x} + O\left(\frac{\Psi(x,y)^3}{x(\log y)^{\frac{1}{4}}}\right),$$

where the primitive non-archimedean singular series $\mathfrak{S}_f^*(c,y)$ was defined in (3.3).

Theorem 3.1 is an immediate consequence of Theorem 4.2, taking an appropriate limit of nonnegative weight functions $\Phi(x)$ approaching the characteristic function $\chi_{[0,1]}(x)$.

The two theorems above imply the truth of a weighted analogue of Conjecture 3.1 for $\kappa > 8$, as follows.

Theorem 4.3. (Relative density of weighted smooth solutions) *Assume the truth of the GRH. For any nonnegative real-valued function $\Phi(x) \in C_c^\infty(\mathbb{R}_{>0})$ not identically zero, there holds*

$$\lim_{x\to\infty} \frac{N^*(x,(\log x)^\kappa;\Phi)}{N(x,(\log x)^\kappa;\Phi)} = \zeta\left(2-\frac{3}{\kappa}\right) \text{ for } \kappa > 8.$$

Proof. This result is based on the identity of Euler products
(4.8)
$$\mathfrak{S}_f^*(c) := \prod_p \left(\left(1+\frac{p-1}{p(p^{3c-1}-1)}\left(\frac{p-p^c}{p-1}\right)^3\right)\left(1-\frac{1}{p^{3c-1}}\right)\right) = \frac{1}{\zeta(3c-1)}\mathfrak{S}_f(c).$$

This identity shows that $\mathfrak{S}_f(c)$ has a meromorphic continuation to the half-plane $Re(c) > \frac{1}{2}$, with its only singularity on this region being a simple pole at $c = \frac{2}{3}$ having residue $\frac{1}{3}\mathfrak{S}_f^*(\frac{2}{3})$. In particular, for real $c = 1 - \frac{1}{\kappa} > \frac{2}{3} + \epsilon$ we have

$$\mathfrak{S}_f(c,y) = \mathfrak{S}_f(c)\left(1 + O_\epsilon\left(\frac{1}{y}\right)\right),$$

and for real $c > \frac{1}{2} + \epsilon$ we have

$$\mathfrak{S}_f^*(c, y) = \mathfrak{S}_f^*(c)\left(1 + O_\epsilon\left(\frac{1}{y}\right)\right).$$

Using these estimates in the main terms of Theorem 4.1 and Theorem 4.2, yields for $\kappa > 8 + \delta$, the estimate
(4.9)
$$N^*(x, (\log x)^\kappa; \Phi) = \frac{1}{\zeta(2 - \frac{3}{\kappa})} N(x, (\log x)^\kappa; \Phi))\left(1 + O_\delta\left(\frac{1}{(\log \log x)^{\frac{1}{4}}}\right)\right).$$

The positivity hypothesis on $\Phi(x)$ implies that $N(x, (\log x)^\kappa; \Phi) > 0$ so we may divide both sides of (4.9) by it to obtain the ratio estimate (3.7). $\square$

5.   Proof sketch for Theorem 4.1

Theorem 4.1 is established in [23] using the Hardy-Littlewood method ([16], [17]), in the modern form using finite exponential sums, see Vaughan [30]. We introduce the *weighted exponential sum*

$$E(x, y; \alpha) := \sum_{n \in \mathcal{S}(y)} e(n\alpha)\Phi\left(\frac{n}{x}\right),$$

where $e(x) := e^{2\pi i x}$. We have the identity

(5.1) $$N(x, y; \Phi) = \int_0^1 E(x, y; \alpha)^2 E(x, y; -\alpha)d\alpha,$$

because in multiplying out the exponential sums in the integral, only terms (n_1, n_2, n_3) with $n_1 + n_2 - n_3 = 0$ contribute. The Hardy-Littlewood method estimates the integral on the right side of (5.1) by splitting the integrand into small arcs centered around rational points with small denominators, and adding up the contributions of the arcs. The major contribution will come from those parts of the circle very near points $\frac{a}{q}$ with small denominator, the *major arcs*. The remainder of the circle comprises the *minor arcs*. Our choice of major arcs and minor arcs is given below.

5.1. General bound for $E(x, y, \alpha)$. To estimate the integral (5.1) we wish to determine how the function $E(x, y, \alpha)$ behaves for α near a rational number $\frac{a}{q}$ in lowest terms, and we write $\alpha = \frac{a}{q} + \gamma$. The main estimate for these is given by [23, Theorem 2.3].

Theorem 5.1. *Assume the truth of the GRH. Let $\delta > 0$ be any fixed real number. Let x and y be large with $(\log x)^{2+\delta} \leq y \leq \exp((\log x)^{1/2-\delta})$, and let κ be defined by $y = (\log x)^{\kappa}$. Let $\alpha \in [0,1]$ be a real number with $\alpha = a/q + \gamma$ where $q \leq \sqrt{x}$, $(a, q) = 1$, and $|\gamma| \leq 1/(q\sqrt{x})$.*

(1) If $|\gamma| \geq x^{\delta-1}$ then for any fixed $\epsilon > 0$, we have

$$E(x, y; \alpha) \ll x^{\frac{3}{4}+\epsilon}.$$

(2) If $|\gamma| \leq x^{\delta-1}$ then on writing $q = q_0 q_1$ with $q_0 \in \mathcal{S}(y)$ and all prime factors of q_1 being bigger than y, and writing $c_0 = 1 - 1/\kappa$, then for any fixed $\epsilon > 0$ we have

$$E(x, y; \alpha) = \frac{\mu(q_1)}{\phi(q_1)} \frac{c_0}{q_0^{c_0}} \prod_{p|q_0} \left(1 - \frac{p^{c_0} - 1}{p - 1}\right) \left(\int_0^\infty \Phi(w) e(\gamma x w) w^{c_0} \frac{dw}{w}\right) \Psi(x, y)$$

$$+ O\left(x^{\frac{3}{4}+\epsilon}\right) + O\left(\frac{\Psi(x, y) q_0^{-c_0+\epsilon} q_1^{-1+\epsilon}}{(1 + |\gamma|x)^2} \frac{(\log \log y)}{\log y}\right).$$

These bounds are obtained combining the Hildebrand-Tenenbaum saddle point method with bounds for partial Euler products, as we now explain. This estimate explicitly contains $\Psi(x, y)$ in main term, avoiding the problem that behavior of $\Psi(x, y)$ for very small values of y does not have a convenient simplifying formula.

5.2. Bounding partial Euler products. The Dirichlet series associated to the set $\mathcal{S}(y)$ of all integers having smoothness bound y is given by the partial Euler product

$$\zeta(s; y) = \prod_{p \leq y} (1 - p^{-s})^{-1} = \sum_{n \in \mathcal{S}(y)} n^{-s},$$

associated to the Riemann zeta function; this series converges absolutely on the half-plane $Re(s) > 0$. We invoke the GRH to control the size of $\zeta(s; y)$ and more generally for the partial Euler products associated to primitive Dirichlet L-functions,

$$L(s, \chi; y) := \prod_{p \leq y} \left(1 - \chi(p) p^{-s}\right)^{-1},$$

([23, Proposition 5.1]).

Proposition 5.1. *Assume the truth of the GRH. Let $\chi \bmod q$ be a primitive Dirichlet character. For any $\epsilon > 0$, and s a complex number with $Re(s) = \sigma \geq 1/2 + \epsilon$, we have*

$$|L(s, \chi; y)| \ll_\epsilon (q|s|)^\epsilon.$$

For the trivial character we have, with $\sigma = Re(s) \geq 1/2 + \epsilon$,

$$|\zeta(s; y)| \ll_\epsilon \exp\left(\frac{y^{1-\sigma}}{(1 + |t|) \log y}\right) |s|^\epsilon.$$

This result is analogous to a Lindelöf hypothesis bound. It is proved by using the GRH and the "explicit formula" techniques of prime number theory (see Davenport [10, Chaps. 17 and 19]) to estimate $\sum_{n \leq y} \Lambda(n)\chi(n)n^{-it}$, followed by partial summation to estimate $\log|L(\sigma + it, \chi; y)|$.

5.3. Estimating the weighted exponential sum. We obtain the following estimates for the exponential sum at a value $\alpha = \frac{a}{q} + \gamma$ near a rational point ([23, Proposition 6.1]).

Proposition 5.2. *Assume the truth of the GRH. Let α be a real number in $[0,1]$ and write $\alpha = a/q + \gamma$ with $(a,q) = 1$, $q \leq \sqrt{x}$, and $|\gamma| \leq 1/(q\sqrt{x})$. Then*

$$E(x, y; \alpha) = M(x, y; q, \gamma) + O(x^{\frac{3}{4}+\epsilon}),$$

where the "local main term" $M(x, y; q, \gamma)$ is defined by

$$M(x, y; q, \gamma) := \sum_{n \in \mathcal{S}(y)} \frac{\mu(\frac{q}{(q,n)})}{\phi(\frac{q}{(q,n)})} e(n\gamma)\Phi\left(\frac{n}{x}\right).$$

This result is proved using an expansion in terms of Dirichlet characters

$$E(x, y; \alpha) = \sum_{\substack{d|q \\ d \in \mathcal{S}(y)}} \frac{1}{\phi(q/d)} \sum_{\chi \bmod q/d} \chi(a)\tau(\bar{\chi}) \sum_{m \in \mathcal{S}(y)} e(md\gamma)\chi(m)\Phi\left(\frac{md}{x}\right).$$

The contribution of the principal characters to this sum gives the "local main term" above. The contribution of each non-principal character is shown to be bounded by

$$\frac{\sqrt{q}}{\sqrt{d}} \sum_{m \in \mathcal{S}(y)} e(md\gamma)\chi(m)\Phi\left(\frac{md}{x}\right) \ll x^{\frac{3}{4}+\epsilon},$$

This is established using Proposition 5.1 for primitive characters, with an additional observation to handle imprimitive characters.

5.4. Estimating the "local main terms". To study the "local main term" we first observe that we may uniquely factor any integer $q = q_0 q_1$ where q_0 is divisible only by primes at most y, and q_1 is divisible only by primes larger than y. Then we have the identity

(5.2) $$M(x, y; q, \gamma) = \frac{\mu(q_1)}{\phi(q_1)} M(x, y; q_0, \gamma).$$

Thus we reduce to studying the local main term $M(x, y; q_0, \gamma)$ with $q_0 \in \mathcal{S}(y)$. Here use the Hildebrand-Tenebaum saddle-point method, expressing it as a contour integral. We have, for $\sigma = Re(s) > 1$,

(5.3) $$M(x, y; q_0, \gamma) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \zeta(s; y)H(s; q_0)x^s \check{\Phi}(s, \gamma x)ds.$$

in which

$$H(s; q_0) = q_0^{-s} \prod_{p \mid q_0} \left(1 - \frac{p^s - 1}{p - 1}\right),$$

and

$$\check{\Phi}(s, \lambda) := \int_0^\infty \Phi(w) e(\lambda w) w^{s-1} dw.$$

The Hilbert-Tenenbaum saddle point method applied to $\zeta(s; y)$ deforms the contour (5.3) to approximate near the real axis the vertical line $Re(s) = c$ where $c = c(x, y)$ is the *Hildebrand-Tenenbaum saddle point value*. This is defined to be the unique positive solution of the equation

$$\sum_{p \leq y} \frac{\log p}{p^c - 1} = \log x.$$

The root is unique because the function

$$g(c; y) := \sum_{p \leq y} \frac{\log p}{p^c - 1}$$

is strictly decreasing for $c > 0$, with limit $+\infty$ as $c \to 0^+$ and limit 0 as $c \to \infty$. We use this saddle-point value in the integral (5.3) and obtain [23, Proposition 6.2].

Proposition 5.3. *Let x and y be large, and assume that $(\log x)^{2+\delta} \leq y \leq \exp((\log x)^{1/2-\delta})$. Let $c = c(x, y)$ denote the Hildebrand-Tenenbaum saddle point value. Suppose $q_0 \in \mathcal{S}(y)$ and let γ be real with $|\gamma| \leq 1/(q_0\sqrt{x})$, and let $M(x, y; q_0, \gamma)$ be as in Proposition 5.2. Then we have:*

(1) If $|\gamma| \geq x^{\delta-1}$ then for any fixed $\epsilon > 0$, we have

$$|M(x, y; q_0, \gamma)| \ll x^{\frac{3}{4}+\epsilon} q_0^{-\frac{3}{4}+\epsilon}.$$

(2) If $|\gamma| \leq x^{\delta-1}$ then for any fixed $\epsilon > 0$, we have

$$M(x, y; q_0, \gamma) = \frac{1}{q_0^c} \prod_{p \mid q_0} \left(1 - \frac{p^c - 1}{p - 1}\right)(c\check{\Phi}(c, \gamma x))\Psi(x, y) + O_\epsilon(x^{\frac{3}{4}+\epsilon} q_0^{-\frac{3}{4}+\epsilon})$$

$$+ O_\epsilon\left(\frac{\Psi(x, y) q_0^{-c+\epsilon}}{(\log y)(1 + |\gamma|x)^2}\right).$$

We have the formula, valid for $\kappa \geq 1 + \delta$,

(5.4) $$c(x, y) = 1 - \frac{1}{\kappa} + O_\delta\left(\frac{\log \log y}{\log y}\right),$$

deducible from [20, Theorem 2]. This estimate is used in replacing c by c_0 in Theorem 5.1 above.

5.5. Hardy-Littlewood method estimates. For parameter values (x, y) we dissect the unit interval in the integral (5.1) into Farey arcs depending only on the parameter x. These are enumerated by rational numbers in the Farey sequence $\mathcal{F}(Q)$ of order $Q = x^{\frac{1}{2}}$. Here $\mathcal{F}(x^{\frac{1}{2}}) := \{\frac{a}{q} :$ $(a, q) = 1$ and $1 \leq q \leq x^{\frac{1}{2}}\}$. The Farey interval assigned to $\frac{a}{q}$ is the arc between it and the mediant $\frac{a+a'}{q+q'}$ of its left neighbor $\frac{a'}{q'} < \frac{a}{q}$, and the same for its right neighbor; these intervals partition the unit interval. We extract from some of these Farey intervals the *major arcs*. The definition of the major arcs depends on an initially specified *cutoff parameter δ* satisfying $0 < \delta \leq \frac{1}{4}$. Decreasing this parameter makes the major arcs smaller. We will eventually let δ become arbitrarily small. The set of *major arcs $\mathfrak{M}$* consists of an interval associated with each $\frac{a}{q}$ with $1 \leq q \leq x^{\frac{1}{4}}$, which consists of that part of the Farey interval of $\frac{a}{q}$ satisfying

$$\mathfrak{M}(\frac{a}{q}) = \Big\{\alpha : \ |\alpha - \frac{a}{q}| \leq x^{\delta - 1}\Big\}.$$

The set of *minor arcs $\mathfrak{m}$* consist of the rest of the interval $[0, 1]$ not covered by the major arcs.

We do not use all of the Farey intervals in $\mathcal{F}(x^{\frac{1}{2}})$ as major arcs, but we carry out some intermediate estimates for all such intervals, for possible later uses.

The estimates of Proposition 5.2(1) and Proposition 5.3(1) with (5.2) lead to a bound $O_\epsilon\Big(x^{\frac{3}{4} + \epsilon} \Psi(x, y)\Big)$ for the minor arcs contribution. Next Proposition 5.3(2) is used to simplify the integral over the major arcs to

$$N(x, y; \Phi) \approx \sum_{\substack{1 \leq q_0 \leq x^{1/4} \\ q_0 \in \mathcal{S}(y)}} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \int_{-x^{\delta-1}}^{x^{\delta-1}} E(x, y; q_0, \frac{a}{q} + \gamma)^2 M(x, y; q_0; -\frac{a}{q} - \gamma) d\gamma,$$

$$\approx \sum_{\substack{1 \leq q_0 \leq x^{1/4} \\ q_0 \in \mathcal{S}(y)}} \phi(q) \int_{-x^{\delta-1}}^{x^{\delta-1}} M(x, y; q_0, \gamma)^2 M(x, y; q_0; -\gamma) d\gamma,$$

on noting that all fractions $\frac{a}{q}$ contribute identical main terms to the latter integral, with the minor arcs estimates used to control the remainder terms.

We then estimate $M(x, y; q, \gamma)$ in the formula above using Proposition 5.3(2). It remains to integrate out the γ-variable in this formula over the major arcs, which produces the archimedean singular series contribution $\mathfrak{S}_\infty(c, \Phi)$ appearing in the final answer in Theorem 4.1, and the sum over $q = q_0 q_1$ in the major arcs terms in (5.2) , which is multiplicative, produces the non-archimedean singular series contribution $\mathfrak{S}_f(c, y)$ appearing in Theorem 4.1. The difference between the contribution of $q_0 \in \mathcal{S}(y)$ and

that of q_1 divisible only by primes greater than y in (5.2) accounts for the difference in Euler product factors for $p \leq y$ and $p > y$ in the formula (3.3) for $\mathfrak{S}_f^*(y, c)$. Finally the estimate (5.4) allows the replacement of c by $1 - \frac{1}{\kappa}$ in the two singular series, to prove Theorem 4.1.

Finally we note that the inclusion-exclusion argument that produces Theorem 4.2 leads to the "primitive non-archimedean singular series" $\mathfrak{S}_f^*(1 - \frac{1}{\kappa}, y)$.

6. Concluding Remarks

We discuss various complementary questions and related problems.

(1) *Coefficients and side congruence constraints.* The Hardy-Littlewood method approach for the upper bound should apply to other linear additive problems involving smooth numbers. One can certainly treat smooth solutions of homogeneous linear ternary Diophantine equations having arbitrary coefficients (a, b, c), i.e.

$$aX + bY + cZ = 0.$$

The singular series will need to be modified appropriately. One could also impose congruence side conditions, on the allowable prime factors. For example, one could consider smooth solutions with all prime factors $p \equiv 1 \,(\mathrm{mod}\ 4)$. In this situation there may occur local congruence obstructions to existence of solutions.

(2) *Additional variables.* One can also treat by this method smooth solutions to linear homogeneous Diophantine equations in more variables, i.e. $X_1 + X_2 + \cdots + X_n = 0$, for $n \geq 4$. Here it is natural to restrict to primitive solutions which also have the non-degeneracy property that no sum of a proper subset of the variables vanishes.

The heuristic given in Section 1.2 is easily modified to show that the cutoff smoothness bound for infinitely many non-degenerate primitive solutions should be

$$S(X_1, X_2, ..., X_n) \ll (\log H(X_1, ..., X_n))^{1 + \frac{1}{n-1} + \epsilon}.$$

Here we can again add coefficients $\sum a_j X_j$ and put congruence restrictions on the allowed prime divisors of the X_i.

(3) *Binary additive variant.* Consider the (inhomogeneous) binary equation

(6.1) $$X + Y = 1.$$

Here the relative primality of solutions is built in to the equation. A similar heuristic to that in §1.2 applies to indicate there should be a threshold value κ_0^* such that there are infinitely many solutions $(X, 1, X + 1)$ having smoothness $S(X, 1, X + 1) \ll (\log X)^{\kappa_0^* + \epsilon}$, but only finitely many solutions

having $S(X, 1, X + 1) \ll (\log X)^{\kappa_0^* - \epsilon}$. This heuristic predicts that $\kappa_0^* = 2$. The examples in §1.3 include several "unusually good" triples with $\kappa_0(X, 1, X + 1) < 2$. However obtaining lower bounds for the number of solutions to (6.1), even conditionally under GRH, seems out of reach. The circle method is unable to effectively control minor arc estimates for binary problems.

(4) *Extension to algebraic number fields.* The *abc*-Conjecture has been generalized to number fields. For the latest results in this direction see Masser [24], Győry and Yu [15] and Győry [14]. One might formulate an extension of the xyz-Conjecture to algebraic number fields, considering triples (X, Y, Z) of algebraic integers with $X + Y + Z = 0$. In this case the smoothness bound will involve the (absolute) norms of prime ideals dividing the algebraic integers in the equation, and should be scaled to be independent of the field of definition of the equation.

(5) *Reverse implications for zeros of L-functions.* It is known that a uniform generalization of the *abc*-Conjecture for number fields has implications concerning location of zeros of L-functions, at least concerning nonexistence of "Siegel zeros" cf. Granville and Stark [12]. Their results are based on analyzing algebraic integer solutions to modular equations, cf. (1.10). Since the lower bound for the xyz-Conjecture is a weaker implication than the *abc*-Conjecture, one may ask whether a suitable uniform xyz-Conjecture for number fields is sufficient to imply nonexistence of Siegel zeros, by a similar approach.

(6) *Function field case.* Let K be a function field of one variable over a suitable non-algebraically closeld k, e.g $k = \mathbb{Q}$ or $k = \mathbb{F}_q$. In this context the *abc*-Conjecture is an unconditional theorem, and the GRH is unconditional over finite fields. It seems plausible that some analogues of the xyz Conjecture might be true and provable unconditionally in this context.

References

[1] A. Balog and A. Sárközy, *On sums of integers having small prime factors. I.* Stud. Sci. Math. Hungar. **19** (1984), 35–47.

[2] A. Balog and A. Sárközy , *On sums of integers having small prime factors. II.* Stud. Sci. Math. Hungar. **19** (1984), 81–88.

[3] E. Bombieri and W. Gubler, *Heights in Diophantine Geometry.* Cambridge University Press, Cambridge, 2006.

[4] R. de la Bretèche, *Sommes d'exponentielles et entiers sans grand facteur premier.* Proc. London Math. Soc. **77** (1998), 39–78.

[5] R. de la Bretèche, *Sommes sans grand facteur premier.* Acta Arith. **88** (1999), 1–14.

[6] R. de la Bretèche and A. Granville, *Densité des friables.* Preprint (2009).

[7] R. de la Bretèche and G. Tenenbaum, *Séries trigonométriques à coefficients arithmétiques.* J. Anal. Math. **92** (2004), 1–79.

[8] R. de la Bretèche and G. Tenenbaum, *Propriétés statistiques des entiers friables,* Ramanujan J. **9** (2005), 139–202.

[9] R. DE LA BRETÈCHE AND G. TENENBAUM, *Sommes d'exponentielles friables d'arguments rationnels*, Funct. Approx. Comment. Math. **37** (2007), 31–38.

[10] H. DAVENPORT, *Multiplicative Number Theory. Second Edition* (Revised by H. L. Montgomery). Springer-Verlag, New York, 1980.

[11] P. ERDŐS, C. STEWART AND R. TIJDEMAN, *Some diophantine equations with many solutions*. Compositio Math. **66** (1988), 37–56.

[12] A. GRANVILLE AND H. M. STARK, *abc implies no "Siegel zeros" for L-functions of characters with negative discriminant*. Invent. Math. **139** (2000), 509–523.

[13] B. GROSS AND D. B. ZAGIER, *On singular moduli*. J. reine Angew. Math. **355** (1985), 191–220.

[14] K. GYŐRY, *On the abc Conjecture in algebraic number fields*. Acta Arith. **133**, (2008), no. 3, 283–295.

[15] K. GYŐRY AND K. YU, *Bounds for the solutions of S-unit equations and decomposable form equations*. Acta Arith. **123** (2006), no. 1, 9–41.

[16] G. H. HARDY AND J. E. LITTLEWOOD, *Some problems in partitio numerorum III. On the expression of a number as a sum of primes*. Acta Math. **44** (1923), 1–70.

[17] G.H. HARDY AND J. E. LITTLEWOOD, *Some problems in partitio numerorum V. A further contribution to the study of Goldbach's problem*. Proc. London Math. Soc., Ser. 2, **22** (1924), 46–56.

[18] A. HILDEBRAND, *Integers free of large prime factors and the Riemann hypothesis*. Mathematika **31** (1984), 258–271.

[19] A. HILDEBRAND, *On the local behavior of $\Psi(x, y)$*. Trans. Amer. Math. Soc. **297** (1986), 729–751.

[20] A. HILDEBRAND AND G. TENENBAUM, *On integers free of large prime factors*. Trans. Amer. Math. Soc. **296** (1986), 265–290.

[21] A. HILDEBRAND AND G. TENEBAUM, *Integers without large prime factors*. J. Theor. Nombres Bordeaux **5** (1993), 411–484.

[22] S. KONYAGIN AND K. SOUNDARARAJAN, *Two S-unit equations with many solutions*. J. Number Theory **124** (2007), 193–199.

[23] J. C. LAGARIAS AND K. SOUNDARARAJAN, *Counting smooth solutions to the equation $A + B = C$*. Proc. London Math. Soc., to appear.

[24] D. W. MASSER, *On abc and discriminants*. Proc. Amer. Math. Soc. **130** (2002), 3141–3150.

[25] J. OESTERLÉ, *Nouvelles approches du "théorème" de Fermat*. Sém. Bourbaki, Exp. No. 694, Astérisque No. **161-162** (1988), 165–186 (1989).

[26] B. POONEN, E. F. SCHAEFER AND M. STOLL, *Twists of $X(7)$ and primitive solutions to $x^2 + y^3 = z^7$*. Duke Math. J. **137** (2007), 103–158.

[27] C. L. STEWART AND KUNRUI YU, *On the abc Conjecture II*. Duke Math. J. **108** (2001), 169–181.

[28] C. L. STEWART AND R. TIJDEMAN, *On the Oesterlé-Masser Conjecture*. Monatshefte Math. **102** (1986), 251–257.

[29] G. TENENBAUM, *Introduction to analytic and probabilistic number theory*. Cambridge Univ. Press, Cambridge, 1995.

[30] R. C. VAUGHAN, *The Hardy-Littlewood Method, Second Edition*. Cambridge Tracts in Mathematics **125**, Cambridge Univ. Press, 1997.

[31] B. M. M. DE WEGER, *Solving exponential Diophantine equations using lattice basis reduction algorithms*. J. Number Theory **26** (1987), no. 3, 325–367.

Jeffrey C. Lagarias
University of Michigan
Department of Mathematics
530 Church Street
Ann Arbor, MI 48109-1043, USA
E-mail: `lagarias@umich.edu`

Kannan Soundararajan
Department of Mathematics
Stanford University
Department of Mathematics
Stanford, CA 94305-2025,USA
E-mail: `ksound@stanford.edu`