Lower Bounds for Numbers of ABC-Hits

Sander Dahmen

October 5, 2006

Abstract

By an ABC-hit, we mean a triple (a,b,c) of relatively prime positive integers such that a+b=c and $\operatorname{rad}(abc) < c$. Denote by N(X) the number of ABC-hits (a,b,c) with $c \leq X$. In this paper we discuss lower bounds for N(X). In particular we prove that for every $\epsilon > 0$ and X large enough $N(X) \geq \exp\left((\log X)^{1/2-\epsilon}\right)$.

1 Introduction

Definition 1. A triple $(a, b, c) \in \mathbb{Z}_{>0}^3$ is called an ABC-sum if a + b = c and gcd(a, b, c) = 1.

The ABC conjecture states that for every $\epsilon > 0$ there exist at most finitely many ABC-sums (a, b, c) such that $c > (\operatorname{rad}(abc))^{1+\epsilon}$.

Definition 2. An ABC-sum (a, b, c) is called an ABC-hit if c > rad(abc).

It is easy to construct infinitely many ABC-hits. Let for example $s \in \mathbb{Z}_{\geq 2}$ and p be a prime not dividing s. Now define for every $n \in \mathbb{Z}_{>0}$ an ABC-sum (a_n, b_n, c_n) by letting

$$a_n = s^{(p-1)p^n} - 1, \ b_n = 1, \ c_n = s^{(p-1)p^n}.$$

Since $s^{(p-1)p^n} = s^{\phi(p^{n+1})} \equiv 1 \pmod{p^{n+1}}$, we have $p^{n+1}|a_n$, hence

$$rad(a_n b_n c_n) \le \frac{a_n}{p^n} \cdot 1 \cdot s \le \frac{s}{p^n} c_n.$$

So for n large enough, the ABC-sums (a_n, b_n, c_n) are ABC-hits.

There are better so-called lower bounds in the ABC conjecture. C. L. Stewart and R. Tijdeman proved in [S-T] that if $C_0 < 4$, then there exist infinitely many ABC-sums (a, b, c) such that

$$c > R \exp\left(C_0 \frac{\sqrt{R}}{\log \log R}\right), \ R := \operatorname{rad}(abc).$$

Later, M. van Frankenhuysen improved this result by showing that it holds with $C_0 < 4$ replaced by $C_0 = 4\sqrt{2}$. Together with an idea of H. W. Lenstra, this $C_0 = 4\sqrt{2}$ can even be replaced by $C_0 < 4 \cdot 2^{0.2995} \sqrt[4]{2\pi/e}$; see [Fra].

Now define the counting function $N(x): \mathbb{R}_{>0} \to \mathbb{Z}_{>0}$ by

$$N(X) := |\{ABC - hits (a, b, c) | c \le X\}|.$$

In [S-T] and [Fra] methods from the geometry of numbers together with versions of the prime number theorem with error term are used to arrive at the above mentioned results. In this paper these methods are applied to prove the following theorem.

Theorem 3. For every $\epsilon > 0$ there exists an $X_0 > 0$ such that for all $X \geq X_0$

$$N(X) \ge \exp\left((\log X)^{\frac{1}{2} - \epsilon}\right).$$

Before we start with (the preliminaries of) the proof of this theorem, we would like to mention that according to [G-S] we have the following upper bound.

Theorem 4. For every $\epsilon > 0$ there exists an $X_0 > 0$ such that for all $X \geq X_0$

$$N(X) \le X^{\frac{2}{3} + \epsilon}$$
.

2 Preliminaries

In this section we discuss some isolated results used in the proof of Theorem 3.

2.1 Minkowski

We have the following generalized version of Minkowski's convex body theorem.

Theorem 5. Let $\Lambda \subset \mathbb{R}^n$ be a lattice of rank n and let V be a convex, centrally symmetric subset of \mathbb{R}^n . If

$$\operatorname{vol}_n(V) > m2^n \det \Lambda$$

for some $m \in \mathbb{Z}_{>0}$, then V contains at least m different pairs of nonzero lattice points $\pm v_i \in \Lambda$, i = 1, ..., m.

Proof. See [Cas, Ch. III, Theorem II].
$$\Box$$

The set V that will be used in the theorem above is (up to a scalar multiple) described in the following lemma.

Lemma 6. Let $n \in \mathbb{Z}_{>0}$ and define the subset $V \subset \mathbb{R}^n$ as

$$V := \{ x \in \mathbb{R}^n \mid \sum_{\substack{i=1\\x_i > 0}}^n x_i \le 1 \text{ and } \sum_{\substack{i=1\\x_i < 0}}^n |x_i| \le 1 \}.$$

Then $\operatorname{vol}_n(V) = \frac{(2n)!}{n!^3}$.

Proof. Let p be an integer satisfying $0 \le p \le n$, define

$$K_p := \{x = (x_1, \dots, x_n) \in \mathbb{R}^n \mid x_i \ge 0 \text{ for } i \le p \text{ and } x_i \le 0 \text{ for } i > p\}$$

We will compute the volume of V contained in K_p . First define the 'm-dimensional hyperpyramid'

$$Y_m := \{x = (x_1, \dots, x_m) \in \mathbb{R}^m \mid x_1, \dots, x_m \ge 0 \text{ and } \sum_{i=1}^m x_i \le 1\},$$

which has volume 1/m!. Identify \mathbb{R}^n with $\mathbb{R}^p \times \mathbb{R}^{n-p}$, then it follows that $K_p \cap V_n = Y_p \times (-Y_{n-p})$. So

$$\operatorname{vol}_n(K_p \cap V_n) = \operatorname{vol}_p(Y_p).\operatorname{vol}_{n-p}(Y_{n-p}) = \frac{1}{p!(n-p)!}.$$

Now let $I \subset \{1, 2, \dots, n\}$ and define

$$K_I := \{ x = (x_1, \dots, x_n) \in \mathbb{R}^n \mid x_i \ge 0 \text{ for } i \in I \text{ and } x_i \le 0 \text{ for } i \notin I \}.$$

Adding up the volumes of V contained in K_I for all 2^n possible sets I gives the volume of V. Note that $K_p = K_{\{1,2,\ldots,p\}}$. If I contains p elements then by symmetry also $\operatorname{vol}_n(K_I \cap V) = 1/(p!(n-p)!)$. If I contains p elements, there are $\binom{n}{p}$ possibilities for the set I. Summing over all possible sets I we arrive at

$$\operatorname{vol}_n(V) = \sum_{p=0}^n \binom{n}{p} \frac{1}{p!(n-p)!} = \frac{1}{n!} \sum_{p=0}^n \binom{n}{p}^2.$$

From the identity $\sum_{k=0}^{n} {n \choose k}^2 = {2n \choose n}$ (which can be proven by considering the identity $(1+x)^n (1+x)^n = (1+x)^{2n}$ and comparing the coefficients of x^n on both sides) we conclude

$$\operatorname{vol}_n(V) = \frac{1}{n!} \binom{2n}{n} = \frac{(2n)!}{n!^3}.$$

2.2 The Prime Number Theorem

Let $\pi(x)$ denote the number of primes $\leq x$. Then for any $k \in \mathbb{Z}_{>0}$

$$\pi(x) = x \left(\frac{1}{\log x} + \frac{1!}{\log^2 x} + \ldots + \frac{(k-1)!}{\log^k x} + \mathcal{O}\left(\frac{1}{\log^{k+1} x}\right) \right).$$

For a proof see [Ing, p. 65]. Taking k=3 we obtain a formula with the precision we need:

$$\pi(x) = x \left(\frac{1}{\log x} + \frac{1}{\log^2 x} + \frac{2}{\log^3 x} + \mathcal{O}\left(\frac{1}{\log^4 x}\right) \right). \tag{1}$$

Using this version of the prime number theorem with error term, we can now derive two formulas that will be needed later.

Lemma 7. Let $x \in \mathbb{R}_{>0}$ and denote by $n := \pi(x) - 1$ the number of odd primes $\leq x$ and by p_1, \ldots, p_n the first n odd primes. Then

$$\sum_{i=1}^{n} \log p_{i} = n \log \left(\frac{x}{e}\right) - \frac{x}{\log^{2} x} + \mathcal{O}\left(\frac{x}{\log^{3} x}\right),$$

$$\sum_{i=1}^{n} \log \log p_{i} = n \log \left(\frac{x}{n}\right) + \mathcal{O}\left(\frac{x}{\log^{3} x}\right).$$

Proof. For $f(y) = \log y$ or $f(y) = \log \log y$ we have

$$\sum_{i=1}^{n} f(p_i) = [f(y)\pi(y)]_2^x - \int_2^x f'(y)\pi(y)dy.$$
 (2)

Now let $f(y) = \log y$. Using (1) we have

$$[f(y)\pi(y)]_2^x = x\left(1 + \frac{1}{\log x} + \frac{2}{\log^2 x} + \mathcal{O}\left(\frac{1}{\log^3 x}\right)\right)$$
 (3)

and

$$\int_{2}^{x} f'(y)\pi(y)dy = \int_{2}^{x} \left(\frac{1}{\log y} + \frac{1}{\log^{2} y} + \mathcal{O}\left(\frac{1}{\log^{3} y}\right)\right)dy. \tag{4}$$

Partial integration shows that for $m \in \mathbb{N}$

$$\int_{a}^{b} \frac{1}{\log^{m} y} dy = \left[\frac{y}{\log^{m} y} \right]_{a}^{b} + m \int_{a}^{b} \frac{1}{\log^{m+1} y} dy.$$
 (5)

So (4) becomes

$$\int_{2}^{x} f'(y)\pi(y)dy = \frac{x}{\log x} + \frac{2x}{\log^{2} x} + \int_{2}^{x} \mathcal{O}\left(\frac{1}{\log^{3} y}\right)dy$$
$$= \frac{x}{\log x} + \frac{2x}{\log^{2} x} + \mathcal{O}\left(\frac{x}{\log^{3} x}\right).$$

Substituting this and (3) in equation (2) yields

$$\sum_{i=1}^{n} \log p_i = x + \mathcal{O}\left(\frac{x}{\log^3 x}\right). \tag{6}$$

By definition we have $n = \pi(x) - 1$. Rewriting equation (1) using the geometric series, we obtain

$$n = x \left(\frac{1}{\log x - 1} + \frac{1}{\log^3 x} + \mathcal{O}\left(\frac{1}{\log^4 x}\right) \right).$$

Multiplication with $\log x - 1$ now gives

$$n(\log x - 1) = x + \frac{x}{\log^2 x} + \mathcal{O}\left(\frac{x}{\log^3 x}\right). \tag{7}$$

From this and (6) we now obtain the first part of the lemma. Now let $f(y) = \log \log y$. Using equation (1) we have

$$[f(y)\pi(y)]_2^x = (\log\log x)(n+1) + \mathcal{O}(1) = n\log\log x + \mathcal{O}(\log\log x)$$
 (8)

and

$$\int_{2}^{x} f'(y)\pi(y)dy = \int_{2}^{x} \left(\frac{1}{\log^{2} y} + \mathcal{O}\left(\frac{1}{\log^{3} y}\right)\right)dy. \tag{9}$$

Partial integration (5) gives

$$\int_{2}^{x} f'(y)\pi(y)dy = \frac{x}{\log^{2} x} + \mathcal{O}\left(\frac{x}{\log^{3} x}\right). \tag{10}$$

Substituting this and (8) in equation (2) yields

$$\sum_{i=1}^{n} \log \log p_i = n \log \log x - \frac{x}{\log^2 x} + \mathcal{O}\left(\frac{x}{\log^3 x}\right). \tag{11}$$

On the other hand

$$n \log(\log x - 1) = n \left(\log \log x + \log \left(1 - \frac{1}{\log x} \right) \right)$$

$$= n \left(\log \log x - \left(\frac{1}{\log x} + \mathcal{O}\left(\frac{1}{\log^2 x} \right) \right) \right)$$

$$= n \log \log x - \frac{x}{\log^2 x} + \mathcal{O}\left(\frac{x}{\log^3 x} \right),$$

where the second equality follows from a first order Taylor expansion, and the third from (1). Together with (11) we obtain

$$\sum_{i=1}^{n} \log \log p_i = n \log(\log x - 1) + \mathcal{O}\left(\frac{x}{\log^3 x}\right). \tag{12}$$

We conclude:

$$\sum_{i=1}^{n} \log \log p_{i} = n \log \left(\frac{n(\log x - 1)}{n} \right) + \mathcal{O}\left(\frac{x}{\log^{3} x} \right)$$

$$= n \log \left(\frac{x \left(1 + \mathcal{O}\left(\frac{1}{\log^{2} x} \right) \right)}{n} \right) + \mathcal{O}\left(\frac{x}{\log^{3} x} \right)$$

$$= n \left(\log \left(\frac{x}{n} \right) + \log \left(1 + \mathcal{O}\left(\frac{1}{\log^{2} x} \right) \right) \right) + \mathcal{O}\left(\frac{x}{\log^{3} x} \right)$$

$$= n \log \left(\frac{x}{n} \right) + \mathcal{O}\left(\frac{x}{\log x} \right) \cdot \mathcal{O}\left(\frac{1}{\log^{2} x} \right) + \mathcal{O}\left(\frac{x}{\log^{3} x} \right)$$

$$= n \log \left(\frac{x}{n} \right) + \mathcal{O}\left(\frac{x}{\log^{3} x} \right),$$

where the first equality of course follows from equation (12), the second from (7) and the fourth from (1) and a first order Taylor expansion. This completes the proof of (the second part of) the lemma.

3 The Proof

We are now ready to give the proof of Theorem 3.

Proof. For $q = b/c \in \mathbb{Q}^*$, with $b, c \in \mathbb{Z}$ and $\gcd(b, c) = 1$, define the height $h(q) := \log(\max(|b|, |c|))$, where |.| denotes the standard Archimedean valuation. Let $x \geq 5$ and define $n := \pi(x) - 1$ the number of odd primes $\leq x$. Denote by p_1, \ldots, p_n the first n odd primes. Consider the subgroup of $\mathbb{Q}_{>0}^*$ generated by the first n odd primes

$$Q_n := \{p_1^{a_1} \dots p_n^{a_n} \mid a_i \in \mathbb{Z}\}$$

and the subset of elements of bounded height

$$Q_x := \{ q \in \mathcal{Q}_n \mid h(q) \le B(x) \},$$

where $B(x): \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ is some function to be specified later. Define the injective group homomorphism

$$\varphi_n: \mathcal{Q}_n \to \mathbb{R}^n: p_1^{a_1} \dots p_n^{a_n} \mapsto (a_1 \log p_1, \dots, a_n \log p_n).$$

Then

$$\Lambda_n := \varphi_n(\mathcal{Q}_n) = \{ (a_1 \log p_1, \dots a_n \log p_n) \in \mathbb{R}^n \mid a_i \in \mathbb{Z} \}$$

is a lattice of rank n. Define two more sets

$$L_x := \varphi_n(Q_x) = \{ y \in \Lambda_n \mid \sum_{\substack{i=1\\ y_i > 0}}^n y_i \le B(x) \text{ and } \sum_{\substack{i=1\\ y_i < 0}}^n |y_i| \le B(x) \},$$

$$L_x \subset V_x := \{ y \in \mathbb{R}^n \mid \sum_{\substack{i=1 \ y_i > 0}}^n y_i \le B(x) \text{ and } \sum_{\substack{i=1 \ y_i < 0}}^n |y_i| \le B(x) \}.$$

An important fact that we will use is that there is a 1-1 relation between (unordered) pairs $\pm y \in \Lambda_n - \{0\}$ and ABC-sums (a,b,c) with $\operatorname{rad}(bc) | \prod_{i=1}^n p_i$, given by $(a,b,c) \mapsto \{\varphi_n(b/c), -\varphi_n(b/c)\}$. Pairs $\pm y \in L_x - \{0\}$ correspond under this 1-1 relation to ABC-sums (a,b,c) with $\operatorname{rad}(bc) | \prod_{i=1}^n p_i$ and $\log c \leq B(x)$. Define $\mathcal{Q}_{n,m} := \{b/c \in \mathcal{Q}_n \mid b \equiv c \pmod{2^m}; b,c \in \mathbb{Z}_{>0} \text{ and } \gcd(b,c) = 1\}$ and $\Lambda_{n,m} := \varphi_n(\mathcal{Q}_{n,m})$. Since 3 and 5 $\pmod{2^m}$ generate $(\mathbb{Z}/2^m\mathbb{Z})^*$, we have a surjective homomorphism $\mathcal{Q}_n \to (\mathbb{Z}/2^m\mathbb{Z})^*$ with kernel equal to $\mathcal{Q}_{n,m}$. So $\mathcal{Q}_n/\mathcal{Q}_{n,m} \simeq (\mathbb{Z}/2^m\mathbb{Z})^*$. Which gives us $|\mathcal{Q}_n/\mathcal{Q}_{n,m}| = 2^{m-1}$, hence

$$|\Lambda_n/\Lambda_{n,m}| = 2^{m-1}. (13)$$

Let $\alpha \in \mathbb{Q}$ with $0 < \alpha < 1$, denote by d the denominator of α and define $\beta := 1 - \alpha$. Let $m \in \mathbb{Z}$ be such that d|m and

$$2^{m-d} < \frac{\operatorname{vol}_n(V_x)}{2^n \det \Lambda_n} \le 2^m. \tag{14}$$

Together with Lemma 6 and det $\Lambda_n = \prod_{i=1}^n \log p_i$ we get

$$2^{m} \ge \frac{(2n)!B(x)^{n}}{2^{n}(n!)^{3} \prod_{i=1}^{n} \log p_{i}} = \exp\left(\log\left(\frac{(2n)!B(x)^{n}}{2^{n}(n!)^{3}}\right) - \sum_{i=1}^{n} \log\log p_{i}\right). \quad (15)$$

Stirling's formula, $\log n! = n \log n - n + \mathcal{O}(\log n)$, gives

$$\log\left(\frac{(2n)!}{2^n(n!)^3}\right) = n\log\left(\frac{2e}{n}\right) + \mathcal{O}\left(\log n\right).$$

Using this identity and Lemma 7 we obtain from (15)

$$2^{m} \ge \exp\left(n\log\left(\frac{2eB(x)}{x}\right) + \mathcal{O}\left(\frac{x}{\log^{3} x}\right)\right). \tag{16}$$

Now note that $\alpha m, \beta m \in \mathbb{Z}$ and that from (13) and (14) we get

$$\operatorname{vol}_n(V_x) > 2^{m+n-d} \det \Lambda_n = 2^{\beta m+1-d} 2^n 2^{\alpha m-1} \det \Lambda_n = 2^{\beta m+1-d} 2^n \det \Lambda_{n,\alpha m}.$$

We will see later that $2^m \to \infty$ when $x \to \infty$, hence for x large enough $2^{\beta m+1-d} \in \mathbb{Z}_{>0}$. So by Theorem 5 we have that for x large enough at least $2^{\beta m+1-d}$ different pairs of nonzero lattice points $\pm y$ of $\Lambda_{n,\alpha m}$ are contained in V_x and hence in L_x . Under the 1-1 relation mentioned earlier, these pairs of points correspond to $2^{\beta m+1-d}$ different ABC-sums (a,b,c) with $\log c \leq B(x)$ and

$$\operatorname{rad}(bc) | \prod_{i=1}^{n} p_i, \ 2^{\alpha m} | c - b = a.$$
 (17)

We claim that for x large enough, these ABC-sums are in fact ABC-hits. From (17), (16) and Lemma 7 we obtain

$$\operatorname{rad}(abc) \leq \frac{2a}{2^{\alpha m}} \prod_{i=1}^{n} p_{i} \leq 2c \left(\frac{1}{2^{m}}\right)^{\alpha} \prod_{i=1}^{n} p_{i}$$

$$\leq c \exp\left(-\alpha n \log\left(\frac{2eB(x)}{x}\right) + \sum_{i=1}^{n} \log p_{i} + \mathcal{O}\left(\frac{x}{\log^{3} x}\right)\right)$$

$$= c \exp\left(n \log\left(\frac{x}{e}\left(\frac{x}{2eB(x)}\right)^{\alpha}\right) - \frac{x}{\log^{2} x} + \mathcal{O}\left(\frac{x}{\log^{3} x}\right)\right).$$

Now define B(x) such that $x/e \cdot (x/(2eB(x)))^{\alpha} = 1$, i.e.

$$B(x) := \frac{1}{2} \left(\frac{x}{e} \right)^{1 + \frac{1}{\alpha}}.$$

Then for x large enough

$$\operatorname{rad}(abc) \le c \exp\left(-\frac{x}{\log^2 x} + \mathcal{O}\left(\frac{x}{\log^3 x}\right)\right) < c.$$

This proves our claim and we conclude that for x large enough

$$N(\exp(B(x)) \ge 2^{\beta m + 1 - d}. (18)$$

Using (16), $2eB(x)/x = (x/e)^{1/\alpha}$ and (1) we obtain

$$2^{m} \geq \exp\left(\frac{n}{\alpha}\log\left(\frac{x}{e}\right) + \mathcal{O}\left(\frac{x}{\log^{3}x}\right)\right)$$

$$= \exp\left(\frac{x}{\alpha}\left(\frac{1}{\log x} + \frac{1}{\log^{2}x} + \frac{2}{\log^{3}x} + \mathcal{O}\left(\frac{1}{\log^{4}x}\right)\right)(\log x - 1)\right)$$

$$= \exp\left(\frac{x}{\alpha}\left(1 + \frac{1}{\log^{2}x} + \mathcal{O}\left(\frac{1}{\log^{3}x}\right)\right)\right).$$

Together with (18) we obtain that for x large enough

$$N(\exp(B(x)) \geq \exp\left(\frac{\beta}{\alpha}x\left(1 + \frac{1}{\log^2 x} + \mathcal{O}\left(\frac{1}{\log^3 x}\right)\right)\right) \geq \exp\left(\frac{\beta}{\alpha}x\right)$$
$$= \exp\left(C'_{\alpha}B(x)^{\frac{\alpha}{1+\alpha}}\right),$$

where $C'_{\alpha}:=e(1/\alpha-1)2^{\alpha/(1+\alpha)}>0$. Since $x\mapsto \exp(B(x)):]0,\infty[\to]1,\infty[$ is surjective and monotonously increasing, we have for X large enough

$$N(X) \ge \exp\left(C_{\alpha}'(\log X)^{\frac{\alpha}{1+\alpha}}\right). \tag{19}$$

Since $\alpha/(1+\alpha) \uparrow 1/2$ when $\alpha \uparrow 1$, we obtain that for every $\epsilon > 0$ there exists a $C_{\epsilon} > 0$ such that for X large enough

$$N(X) \ge \exp\left(C_{\epsilon}(\log X)^{\frac{1}{2}-\epsilon}\right).$$

For the final statement, note that for every $\epsilon > 0$ and X large enough

$$\log N(X) \geq C_{\frac{\epsilon}{2}}(\log X)^{\frac{1}{2}-\frac{\epsilon}{2}} = C_{\frac{\epsilon}{2}}(\log X)^{\frac{\epsilon}{2}}(\log X)^{\frac{1}{2}-\epsilon} \geq (\log X)^{\frac{1}{2}-\epsilon}.$$

We remark that instead of using lattices to find ABC-hits, we could have used the box principle like in [S-T]. Our method of proof then gives, that for X large enough (19) holds, but now with $C'_{\alpha} = e(1/\alpha - 1)$. With this smaller constant we of course also end up with Theorem 3. On the other hand, one can try to find more lattice points inside V_x . Heuristically, one could expect a factor of 2^n more than used in the proof. But this extra factor would only increase the constant C'_{α} in (19) and again would not change the final result. Using not only ABC-sums (a, b, c) with a divisible by a large power of 2 to construct ABC-hits,

but also ABC-sums with a divisible by a large power of other primes, would also not necessarily improve Theorem 3. We do however not expect that Theorem 3 is best possible (in some natural sense). It might be possible that with some extra effort the $(\log X)^{\epsilon}$ term could be replaced by some power of $\log \log X$. We leave this to the interested reader.

References

- [Cas] J. W. S. Cassels, An Introduction to the Geometry of Numbers, Reprint of the 1971 Edition, Springer-Verlag, Berlin Heidelberg, 1997.
- [Fra] M. van Frankenhuysen, A Lower Bound in the abc Conjecture, J. Number Theory 82 (2000), 91-95.
- [G-S] G. Geuze and B. de Smit, Reken Mee met ABC, On-line reference: http://www.math.leidenuniv.nl/~desmit/vakantiecursus2006.pdf.
- [Ing] A. E. Ingham, The Distribution of Prime Numbers, Cambridge University Press, 1964.
- [S-T] C. L. Stewart and R. Tijdeman, On the Oesterlé-Masser Conjecture, Mon. Math. 102 (1986), 251-257.