On the Greatest Common Divisor of Shifted Sets

Randell Heyman and Igor E. Shparlinski
School of Mathematics and Statistics
University of New South Wales
Sydney, NSW 2052, Australia
{randell,igor.shparlinski}@unsw.edu.au

February 28, 2022

Abstract

Given a set of n positive integers $\{a_1, \ldots, a_n\}$ and an integer parameter H we study small additive shift of its elements by integers h_i with $|h_i| \le H$, $i = 1, \ldots, n$, such that the greatest common divisor of $a_1 + h_1, \ldots, a_n + h_n$ is very different from that of $a_1, \ldots, a_n$. We also consider a similar problem for the least common multiple.

1 Introduction

Let $\mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{Z}^n$ be a nonzero vector. The *approximate common divisor problem*, introduced by Howgrave-Graham [13] for $n = 2$, can generally be described as follows. Suppose we are given two bounds $D > h \ge 1$. Assuming that for some h_i with $|h_i| \le H$, $i = 1, \ldots, n$, we have

$$\gcd(a_1 + h_1, \ldots, a_n + h_n) > D, \qquad (1)$$

the task is to determine the shifts $h_1, \ldots, h_n$. If it is also requested that $h_1 = 0$ then we refer to the problem as the *partial approximate common divisor problem* (certainly in this case the task is to find the shifts faster than via complete factorisation of $a_1 \ne 0$).

This problem has a strong cryptographic motivation as it is related to some attacks on the RSA and some other cryptosystems, see [4, 5, 13, 18] and references therein for various algorithms and applications. In particular, much of the current motivation for studying approximate common divisor problems stems from the search for efficient and reliable *fully homomorphic encryption*, that is, encryption that allows arithmetic operations on encrypted data, see [6, 11, 16].

Here we consider a dual question and show that for any $\mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{Z}^n$, there are shifts $|h_i| \le H$, $i = 1, \ldots, n$, for which (1) holds with a relatively large value of D. Throughout we use $\gcd(\mathbf{x})$ to mean $\gcd(x_1, \ldots, x_n)$ for any $\mathbf{x} \in \mathbb{Z}^n$.

We also denote the height of $\mathbf{x}$ with $\mathfrak{H}(\mathbf{x}) = \max\{|x_1|, \ldots, |x_n|\}$.

The implied constants in the symbols 'O', '$\ll$' and '$\gg$' may occasionally, where obvious, depend on the integer parameter n and the real positive parameter ε, and are absolute otherwise. We recall that the notations $U = O(V)$, $U \ll V$ and $V \gg U$ are all equivalent to the assertion that the inequality $|U| \le c|V|$ holds for some constant $c > 0$.

Our treatment of this question is based on some results of Baker and Harman [2] (see also [1]). For an integer $n \ge 1$ and real positive $\varepsilon < 1$, we define $\kappa(n, \varepsilon)$ as the solution $\kappa > 0$ to the equation

$$\frac{n(\varepsilon\kappa - 1)}{n - 1} = \frac{1}{2^{2+\max\{1,\kappa\}} - 4}. \tag{2}$$

The solution is unique as the left hand side of (2) is monotonically increasing (as a function of κ) from $-n/(n-1)$ to $+\infty$ on $[0, \infty)$ while the right hand side of (2) is positive and monotonically non-increasing.

We also set

$$\vartheta(n, \varepsilon) = \frac{1}{(n-1)} \left(1 - \frac{1}{\varepsilon\kappa(n, \varepsilon)}\right).$$

It easy to see from (2) that $\varepsilon\kappa(n, \varepsilon) < 1$, so $\vartheta(n, \varepsilon) > 0$.

Theorem 1. *For any vector $\mathbf{a} \in \mathbb{Z}^n$, any real positive $\varepsilon < 1$ and*

$$H \ge \mathfrak{H}(\mathbf{a})^{\varepsilon}$$

there exists a vector $\mathbf{h} = (h_1, \ldots, h_n) \in \mathbb{Z}^n$ of height

$$\mathfrak{H}(\mathbf{h}) \le H$$

such that

$$\gcd(\mathbf{a}+\mathbf{h}) \gg \mathfrak{H}(\mathbf{h}) H^{\vartheta(n,\varepsilon)}.$$

Next we are interested in asking for which $\mathbf{h}$ the shifted set is pairwise coprime.

For $\mathbf{a} \in \mathbb{Z}^n$ we denote by $L(\mathbf{a})$ the smallest H such that there is a $\mathbf{h} \in \mathbb{Z}^n$ with $\mathfrak{H}(\mathbf{h}) = H$ such that

$$\gcd(a_i + h_i, a_j + h_j) = 1, \qquad 1 \le i < j \le n.$$

For $n = 2$, and thus $\mathbf{a} = (a_1, a_2) \in \mathbb{Z}^2$, Erdős [8, Equation (3)] has given the bound

$$L(\mathbf{a}) \ll \frac{\log \min\{|a_1|, |a_2|\}}{\log \log \min\{|a_1|, |a_2|\}}.$$

However the method of [8] does not seem to generalise to $n \ge 3$.

Theorem 2. *For an arbitrary* $\mathbf{a} \in \mathbb{Z}^n$ *we have*

$$L(\mathbf{a}) \ll \log^2 \mathfrak{H}(\mathbf{a}).$$

Note in fact our argument allows to replace $\mathfrak{H}(\mathbf{a})$ with a smaller qunatity

$$\mathfrak{H}^*(\mathbf{a}) = \min_{1 \le i \le n} \max_{\substack{1 \le j \le n \\ i \ne j}} |a_i|.$$

For $\mathbf{a} \in \mathbb{Z}^n$ we denote by $\ell(\mathbf{a})$ the smallest H such that there is a vector $\mathbf{h} \in \mathbb{Z}^n$ with $\mathfrak{H}(\mathbf{h}) = H$ and

$$\gcd(a_1 + h_1, \ldots, a_n + h_n) = 1.$$

A very simple argument, based on the Chinese Remainder Theorem, implies the following result, which generalises [8, Equation (2)].

Theorem 3. *For infinitely many* $\mathbf{a} \in \mathbb{Z}^n$ *we have*

$$\ell(\mathbf{a}) \gg \left(\frac{\log \mathfrak{H}(\mathbf{a})}{\log \log \mathfrak{H}(\mathbf{a})} \right)^{1/n}.$$

Note that Theorem 3 is essentially an explicit version of a result of Huck and Pleasants [14].

It is clear that for non-zero vector $\mathbf{a} \in \mathbb{Z}^n$ and arbitrary vectors $\mathbf{x}, \mathbf{y} \in \mathbb{Z}^n$ we have

$$\gcd(a_1, \ldots, a_m) \mid \gcd(\mathbf{a} \cdot \mathbf{x}, \mathbf{a} \cdot \mathbf{y}),$$

where

$$\mathbf{a} \cdot \mathbf{x} = \sum_{i=1}^{n} a_i x_i \qquad \text{and} \qquad \mathbf{a} \cdot \mathbf{y} = \sum_{i=1}^{n} a_i y_i.$$

Let $R(\mathbf{a}, h)$ be the number of vectors $\mathbf{x}, \mathbf{y} \in \mathbb{Z}^n$ with positive components and of height $\mathfrak{H}(\mathbf{x}), \mathfrak{H}(\mathbf{y}) \le h$ for which

$$\gcd(a_1, \ldots, a_m) = \gcd(\mathbf{a} \cdot \mathbf{x}, \mathbf{a} \cdot \mathbf{y}). \tag{3}$$

By [10, Theorem 3] we have

$$|R(\mathbf{a}, h) - \zeta(2)^{-1} h^{2n}| \le h^{2n-1/n} (h\mathfrak{H}(\mathbf{a}))^{o(1)},$$

where $\zeta(s)$ is the Riemann zeta function.

Theorem 4. *Let $n \ge 2$ and let $\mathbf{a} \in \mathbb{Z}^n$. Then, for $\max\{h, \mathfrak{H}(\mathbf{a})\} \to \infty$,*

$$|R(\mathbf{a}, h) - \zeta(2)^{-1} h^{2n}| \le h^{2n-n/(n^2-n+1)} (h\mathfrak{H}(\mathbf{a}))^{o(1)}.$$

2 Proof of Theorem 1

We use the following [2, Theorem 1], see also [2, Equation (2.1)] that gives an explicit formula for constant $\gamma(K)$ below.

Lemma 5. *Suppose that for some fixed $K > 0$ and some sufficiently large real positive Q and R we have*

$$\left(\sum_{i=1}^{n} a_i^2 \right)^{1/2} \le R^K$$

and

$$C_1(K, n) \le Q \le R^{\gamma(K)},$$

where

$$\gamma(K) = \frac{1}{2^{2+\max\{1,K\}} - 4}.$$

Let $\psi_1, \ldots, \psi_n$ be positive integers with

$$\psi_i \le c_2(K,n)(\log Q)^{-n}, \quad i = 1, \ldots, n,$$

and

$$\psi_1 \cdots \psi_n = Q^{-1}.$$

Then

$$\left\| \frac{a_i}{r} \right\| \le \psi_i, \quad i = 1, \ldots, n,$$

$$R \le r \le 2QR.$$

where $C_1(K,n)$ and $c_2(K,n)$ depend at most on K and n.

To prove Theorem 1, we choose some parameters Q and R that satisfy Lemma 5 with $K = \kappa(n,\varepsilon)$, where $\kappa(n,\varepsilon)$ is given by (2), and then we set $\psi_i = Q^{-1/n}$, $i = 1, \ldots, n$. Then by Lemma 5, there exist an integer r with $R \le r \le 2QR$ such that

$$\left\| \frac{a_i}{r} \right\| \le Q^{-1/n}, \qquad i = 1, \ldots, n,$$

where $\|\xi\|$ is distance between a real ξ and the closest integer. So for some integers h_i with $|h_i| \le rQ^{-1/n}$ we have

$$a_i + h_i \equiv 0 \pmod r, \qquad i = 1, \ldots, n.$$

Suppose that for some constant $A > 0$ we choose R such that for $Q = (0.5)^{n/(n-1)} A^{-1} R^{\gamma(K)}$, we have

$$2Q^{1-1/n} R = H. \tag{4}$$

Then

$$R = A^{(n-1)/(n\gamma(K)+n)} H^{n/(n\gamma(K)-\gamma(K)+n)}.$$

Then, taking A to satisfy

$$A^{(n-1)/(n\gamma(K)+n)} = n^{1/2K}$$

due to our choice of $K = \kappa(n,\varepsilon)$, we have

$$R = n^{1/2K} H^{n/(n\gamma(K)-\gamma(K)+n)} = n^{1/2K} H^{1/\varepsilon K}. \tag{5}$$

Hence for $\mathbf{h} = (h_1, \ldots, h_n)$ we have

$$\mathfrak{H}(\mathbf{h}) \le rQ^{-1/n} \le 2Q^{1-1/n}R = H$$

and

$$\gcd(\mathbf{a}+\mathbf{h}) \ge r \ge \mathfrak{H}(\mathbf{h})Q^{1/n}. \tag{6}$$

Using (5), we derive

$$\left(\sum_{i=1}^{n} a_i^2\right)^{1/2} \le n^{1/2}H^{1/\varepsilon} = R^K.$$

Thus Lemma 5 indeed applies. We also have

$$Q^{1/n} \gg R^{\gamma(K)/n} \gg H^{\gamma(K)/\varepsilon nK}. \tag{7}$$

We now see from (2) that

$$\frac{\gamma(K)}{\varepsilon nK} = \frac{\varepsilon K - 1}{\varepsilon(n-1)K},$$

which together with (6) and (7) completes the proof.

3 Proof of Theorem 2

We recall the following well-known result of Iwaniec [15] on the *Jacobsthal problem*. For a given r, let $C(r)$ be the maximal length of a sequence of consecutive integers, each divisible by one of r arbitrarily chosen primes. Then Iwaniec [15] gives the following bound on $C(r)$:

Lemma 6. *For a given $r > 1$ we have,*

$$C(r) \ll (r \log r)^2.$$

We are now ready to prove Theorem 2.

We now set $h_1 = 0$ and chose h_i, $i = 2, \ldots, n$ as the smallest non-negative integer with

$$\gcd\left(\prod_{j=1}^{i-1}(a_j + h_j), a_i + h_i\right) = 1.$$

We show that if n is a positive integer and $a = \mathfrak{H}(\mathbf{a})$ then

$$\mathfrak{H}(\mathbf{h}) \ll \log^2 a. \tag{8}$$

For $n = 2$ we note that a_1 has $\omega(a_1)$ distinct prime factors, where $\omega(a)$ is the number of distinct prime divisors of an integer $a \geq 1$.

So, by Lemma 6,

$$C(\omega(a_1)) \ll (\omega(a_1)\log(\omega(a_1))^2 \ll \log^2 a_1 = \log^2 a$$

for all a_1, and from the trivial bound $\omega(k)! \leq k$ and the Stirling formula we have

$$\omega(k) \ll \frac{\log k}{\log(2 + \log k)}$$

for any integer $k \geq 1$. Now a straight forward inductive argument, after simple calculations, implies (8) and concludes the proof.

4 Proof of Theorem 3

Let us choose a sufficiently large parameter H and the first $(2H+1)^n$ primes $p_{i_1,\ldots,i_n} > H$ for $-H \leq i_1, \ldots, i_n \leq H$.

For each $k = 1, \ldots, n$ we define a_k as the smallest positive integer with

$$a_k \equiv i_k \pmod{p_{i_1,\ldots,i_n}}, \qquad -H \leq i_1, \ldots, i_n \leq H.$$

Set $\mathbf{a} = (a_1, \ldots, a_n)$. Clearly, for any $\mathbf{h} \in \mathbb{Z}^n$ with $\mathfrak{H}(\mathbf{h}) \leq H$, we have

$$p_{h_1,\ldots,h_n} \mid \gcd(a_1 + h_1, \ldots, a_n + h_n).$$

This implies that $\ell(\mathbf{a}) \geq H$.

It remains to estimate $\mathfrak{H}(\mathbf{a})$. Clearly, we have $p_{i_1,\ldots,i_n} \ll H^n \log H$ for $-H \leq i_1, \ldots, i_n \leq H$. Therefore,

$$\mathfrak{H}(\mathbf{a}) \leq \prod_{-H \leq i_1,\ldots,i_n \leq H} p_{i_1,\ldots,i_n} = \exp(O(H^n \log H)) = \exp(O(\ell(\mathbf{a})^n \log \ell(\mathbf{a}))),$$

which completes the proof.

5 Proof of Theorem 4

Clearly, it is enough to consider the case where $\gcd(a_1, \ldots, a_n) = 1$.

We can certainly assume that $n \le \log h$ for otherwise the bound is trivial.

Let μ denote the Möbius function, that is $\mu(1) = 1$, $\mu(d) = 0$ if $d \ge 2$ is not squarefree, and $\mu(d) = (-1)^{\omega(d)}$ otherwise, where $\omega(d)$, as before, is the number of prime divisors of an integer $d \ge 1$.

As in the proof of [10, Theorem 3], by the inclusion exclusion principle we have

$$R(\mathbf{a}, h) = \sum_{d \ge 1} \mu(d) U_d(\mathbf{a}, h)^2,$$

where for an integer $d \ge 1$, we denote by $U_d(\mathbf{a}, h)$ the number of vectors $\mathbf{x} \in \mathbb{Z}^n$ with positive components and of height $\mathfrak{H}(\mathbf{x}) \le h$ for which $d \mid \mathbf{a} \cdot \mathbf{x}$.

We now recall from [10] some estimates on $U_d(\mathbf{a}, h)$.

More precisely, for $1 \le d \le 2h/3n$ we have

$$\left| U_d(\mathbf{a}, h)^2 - \frac{h^{2n}}{d^2} \right| \le 8nd^{-1}h^{2n-1}. \tag{9}$$

see [10, Equation (8)]. The proof of (9) also relies on the bound

$$U_d(\mathbf{a}, h) \le d^{n-1} \left(h/d + 1\right)^n. \tag{10}$$

that holds for any integer $d \ge 1$.

Furthermore, for any squarefree $d \ge 1$ we also have the bound

$$U_d(\mathbf{a}, h) \le h^{n-1} \left(hd^{-1/n} + 1\right). \tag{11}$$

see [10, Equation (10)].

Therefore, choosing some parameter D, we write

$$R(\mathbf{a}, h) = M + O(\Delta_1 + \Delta_2) \tag{12}$$

where

$$\begin{aligned}
M &= \sum_{d \le 2h/3n} \mu(d) U_d(\mathbf{a}, h)^2, \\
\Delta_1 &= \sum_{2h/3n < d \le D} \mu(d) U_d(\mathbf{a}, h)^2, \\
\Delta_2 &= \sum_{d > D} \mu(d) U_d(\mathbf{a}, h)^2.
\end{aligned}$$

Using (9), we derive

$$\begin{aligned} M &= \sum_{d\le 2h/3n} \mu(d) \left(\frac{h^{2n}}{d^2} + O\left(h^{2n-1}d^{-1}\right)\right) \\ &= h^{2n} \sum_{d\le 2h/3n} \frac{\mu(d)}{d^2} + O\left(h^{2n-1}\log h\right). \end{aligned}$$

Since

$$\sum_{d\le 2h/3n} \frac{\mu(d)}{d^2} = \sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} + O\left(D^{-1}\right) = \zeta(2)^{-1} + O\left(D^{-1}\right),$$

see [12, Theorem 287], we derive

$$M = h^{2n}\zeta(2)^{-1} + O\left(h^{2n-1}\log h\right). \tag{13}$$

To estimate Δ_1 we apply the bound (10), which for $d \ge 2h/3n$ can be simplified as $U_d(\mathbf{a}, h) = O(d^{n-1})$. Therefore,

$$\Delta_1 \ll \sum_{2h/3n<d\le D} d^{n-1} U_d(\mathbf{a}, h) \le D^{2n-1} \sum_{2h/3n<d\le D} U_d(\mathbf{a}, h). \tag{14}$$

Using the same argument as the proof of [10, Theorem 3], based on a bound of the divisor function $\tau(k)$, we obtain

$$\begin{aligned} \sum_{d>D} U_d(\mathbf{a}, h) &= \sum_{d>D} \sum_{\substack{\mathfrak{H}(\mathbf{x})\le h \\ d\mid \mathbf{a}\cdot\mathbf{x}}} 1 \\ &= \sum_{h(\mathbf{x})\le h} \sum_{\substack{d>D \\ d\mid \mathbf{a}\cdot\mathbf{x}}} 1 \le \sum_{h(\mathbf{x})\le h} \tau(\mathbf{a}\cdot\mathbf{x}) \le h^n (h\mathfrak{H}(\mathbf{a}))^{o(1)}, \end{aligned} \tag{15}$$

where $\mathbf{x}$ runs through integral vectors with positive components. Hence, we see that (14) yields the estimate

$$\Delta_1 \ll D^{n-1} h^n (h\mathfrak{H}(\mathbf{a}))^{o(1)}. \tag{16}$$

Finally, to estimate Δ_2 we apply the bound (11) and, as before derive

$$\Delta_2 \ll h^{n-1}\left(hD^{-1/n} + 1\right) \sum_{d>D} U_d(\mathbf{a}, h) \le h^{2n-1}\left(hD^{-1/n} + 1\right)(h\mathfrak{H}(\mathbf{a}))^{o(1)}. \tag{17}$$

Substituting the bounds (13), (16) and (17) into (12), we obtain

$$R(\mathbf{a}, h) = h^{2n}\zeta(2)^{-1} + O\left(\left(h^{2n-1} + D^{n-1}h^n + h^{2n}D^{-1/n}\right)(h\mathfrak{H}(\mathbf{a}))^{o(1)}\right).$$

Now, choosing

$$D = h^{n^2/(n^2-n+1)},$$

we conclude the proof.

6 Comments

We remark that it is also interesting to study analogous questions for polynomials with integer coefficients or over finite fields, see [7, 17, 9] for some polynomial versions of the approximate common divisor problem. Some of out techniques can be extended to this case, however some important ingredients, such as the results of Baker and Harman [1, 2] are missing.

Acknowledgment

The authors are grateful to Etienne Fouvry and Michel Laurent for very useful discussions.

This work was supported in part by the ARC Grant DP130100237.

References

[1] R. C. Baker and G. Harman, 'The sequence x/n and its subsequences', *Rocky Mount. J. Math.*, **26** (1996), 795–814.

[2] R. C. Baker and G. Harman, 'Small remainder of a vector to suitable modulus', *Math. Zeit.*, **221** (1996), 59–71.

[3] J. Bourgain and M. Z. Garaev, 'Sumsets of reciprocals in prime fields and multilinear Kloosterman sums', *Preprint*, 2012, (available from `http://arxiv.org/abs/1211.4184`).

[4] Y. Chen and P. Q. Nguyen, 'Faster algorithms for approximate common divisors: Breaking fully-homomorphic-encryption challenges over the integers', *Lecture notes in Computer Science*, **7237**, Springer-Verlag, Berlin, 2012, 502–519.

[5] H. Cohn and N. Heninger, ‘Approximate common divisors via lattices’, *Preprint*, 2011, (available from `http://arxiv.org/abs/1108.2714`).

[6] M. van Dijk, C. Gentry, S. Halevi and V. Vaikuntanathan, ‘Fully homomorphic encryption over the integers’, *Lecture notes in Computer Science*, **6110**, Springer-Verlag, Berlin, 1996, 24–43.

[7] M. Elkadi, A. Galligo and T. L. Ba, ‘Approximate GCD of several univariate polynomials with small degree perturbations’, *J. Symbolic Comput.*, **47** (2012), 410–421.

[8] P. Erdős, ‘On an elementary problem in number theory’, *Canadian Math. Bull.*, **1** (1958), 5–8.

[9] J. von zur Gathen, M. Mignotte and I. E. Shparlinski, ‘Approximate polynomial GCD: Small degree and small height perturbations’, *J. Symbolic Comput.*, **45** (2010), 879–886.

[10] J. von zur Gathen and I. E. Shparlinski, ‘GCD of random linear combinations’, *Algorithmica*, **46** (2006), 137–148.

[11] C. Gentry, ‘Fully homomorphic encryption using ideal lattices’, *Proc. of the 2009 ACM Intern. Symp. on Theory of Comp.*, ACM, New York, 2009, 169–178.

[12] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, Oxford Univ. Press, Oxford, 1979.

[13] N. Howgrave-Graham, ‘Approximate integer common divisors’, *Lecture notes in Comput. Sci.*, **2146**, Springer-Verlag, Berlin, 2001, 51–66.

[14] C. Huck and P. A. B. Pleasants, ‘Entropy and diffraction of the k-free points in n-dimensional lattices’, *Discrete Comput. Geom.*, **50** (2013), 39–68.

[15] H. Iwaniec, ‘On the problem of Jacobsthal’, *Demonstratio Math.*, **11** (1978), 225–231.

[16] D. Micciancio, ‘A first glimpse of cryptography’s holy grail’, *Comm. of the Assoc. of Comp. Machinery*, **53** (2010), 96–96.

[17] K. Nagasaka, 'Approximate polynomial GCD over integers', *J. Symbolic Comput.*, **46** (2011), 1306–1317.

[18] S. Sarkar and S. Maitra, 'Approximate integer common divisor problem relates to implicit factorization', *IEEE Trans. Inform. Theory*, **57** (2011), 4002–4013.