GCD of Random Linear Forms

Joachim von zur Gathen[1] and Igor E. Shparlinski[2]

[1] Fakultät für Elektrotechnik, Informatik und Mathematik,
Universität Paderborn,
33095 Paderborn, Germany
gathen@upb.de
http://www-math.upb.de/~aggathen
[2] Department of Computing, Macquarie University,
NSW 2109, Australia
igor@comp.mq.edu.au
http://www.comp.mq.edu.au/~igor

Abstract. We show that for arbitrary positive integers $a_1, \ldots, a_m$, with probability at least $6/\pi^2 + o(1)$, the gcd of two linear combinations of these integers with rather small random integer coefficients coincides with $\gcd(a_1, \ldots, a_m)$. This naturally leads to a probabilistic algorithm for computing the gcd of several integers, with probability at least $6/\pi^2 + o(1)$, via just one gcd of two numbers with about the same size as the initial data (namely the above linear combinations). Naturally, this algorithm can be repeated to achieve any desired confidence level.

1 Introduction

For a vector $\mathbf{u} = (u_1, \ldots, u_m) \in \mathbb{R}^m$ we define its *height* as

$$h(\mathbf{u}) = \max_{i=1,\ldots,m} |u_i|.$$

We let $\mathbf{a} = (a_1, \ldots, a_m) \in \mathbb{N}^m$ be a vector of $m \geq 2$ positive integers, $\mathbf{x} = (x_1, \ldots, x_m), \mathbf{y} = (y_1, \ldots, y_m) \in \mathbb{N}^m$ be two integer vectors of the same length, where $\mathbb{N} = \{1, 2, \ldots\}$, and consider the linear combinations

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

Then clearly $\gcd(a_1, \ldots, a_m)$ divides $\gcd(\mathbf{a} \cdot \mathbf{x}, \mathbf{a} \cdot \mathbf{y})$, and we want to show that in fact, equality holds quite often.

For an integer M, we denote by $\rho_{\mathbf{a}}(M)$ the probability that, for $\mathbf{x}, \mathbf{y}$ chosen uniformly in $\mathbb{N}^m$ with height at most M,

$$\gcd(a_1, \ldots, a_m) = \gcd(\mathbf{a} \cdot \mathbf{x}, \mathbf{a} \cdot \mathbf{y}). \tag{1}$$

Assuming that $\mathbf{a} \cdot \mathbf{x}$ and $\mathbf{a} \cdot \mathbf{y}$ behave as independent random integer multiples of $\gcd(a_1, \ldots, a_m)$, it is reasonable to expect that (1) holds with probability

R. Fleischer and G. Trippen (Eds.): ISAAC 2004, LNCS 3341, pp. 464–469, 2004.

Joachim von zur Gathen & Igor E. Shparlinski (2004). GCD of Random Linear Forms. In *Proceedings of 15th International Symposium on Algorithms and Computation ISAAC '04*, Hong Kong, China, Rudolf Fleischer & Gerhard Trippen, editors, volume 3341 of *Lecture Notes in Computer Science*, 845–846. Springer-Verlag, Berlin, Heidelberg. ISBN 978-3-540-24131-7 (Print) 978-3-540-30551-4 (Online). ISSN 0302-9743. URL http://dx.doi.org/10.1007/978-3-540-30551-4_41.

$\zeta(2)^{-1} = 6/\pi^2$ where $\zeta(s)$ is the Riemann zeta function. Here we obtain a lower bound for $\rho_{\mathbf{a}}(M)$ which for a very wide range of m, M, and $N = h(\mathbf{a})$ shows that this quantity is asymptotically at least that big. The range in which this is established improves quite substantially the corresponding result of [2]. In particular, our result implies that one can choose M of order $\ln N$ in the algorithm of [2] rather than of order N as in Corollary 3 of [2], thus reducing quite dramatically the size of the operands which arise in the algorithm of [2].

The lower bound on $\rho_{\mathbf{a}}(M)$ plays a crucial role in the analysis of a fast probabilistic algorithm for computing the gcd of several integers which has been studied in [2]. This algorithm, for any $\delta > 0$, requires only about

$$\frac{1}{\ln(\pi^2/(\pi^2-6))} \ln \delta^{-1} = 1.06802\ldots \ln \delta^{-1} \tag{2}$$

pairwise gcd computations, to achieve success probability at least $1 - \delta$ (where $\ln z$ is the natural logarithm of $z > 0$). For comparison, it is noted that the naive deterministic approach may require up to $m - 1$ gcd computations. A drawback of the algorithm of [2] is that for its proof of correctness to work, the arguments given to the gcd computations have to be substantially larger than the original inputs. Our results now imply that one may choose the operands of that algorithm of approximately the same size as the inputs. An exact cost analysis depends on the cost of the particular gcd algorithm, a variety of which can be found in [3].

A well-known fact says that $\gcd(a_1, \ldots, a_m)$ equals 1 with probability $\zeta^{-1}(m)$ for random integers $a_1, \ldots, a_m$; see [4], Theorem 332, for a precise formulation in the case $m = 2$. It is important to not confuse our result which holds for arbitrary ("worst-case") inputs with the "average-case" result which follows from this fact.

2 Main Result

We show that for a wide choice of parameters $\rho_{\mathbf{a}}(M) \geq 0.607$. More precisely, we have the following.

Theorem 1. *Let $\mathbf{a} \in \mathbb{Z}^m$ be of height at most N. Then for any $M > m$, we have $\rho_{\mathbf{a}}(M) \geq \zeta(2)^{-1} - \Delta$, where $\Delta = O\left(\ln^{-1}(M/m) + M^{-1}\ln(MN)\right)$.*

Proof. Without loss of generality we can assume that M/m is large enough because otherwise the result is trivial. As in [2], we remark that it is enough to consider only the case $\gcd(a_1, \ldots, a_m) = 1$.

We define Q as the largest integer with the condition

$$\prod_{p \leq Q} p \leq (M/m)^{1/2},$$

where the product is taken over all primes $p \leq Q$. By the Prime Number Theorem, see Theorem 4.4 of [1], we have $Q = (1/2 + o(1))\ln(M/m)$.

Let $\mathcal{L}$ be the set of all pairs of integer vectors $\mathbf{x}, \mathbf{y} \in \mathbb{N}^m$ with $h(\mathbf{x}), h(\mathbf{y}) \le M$. For an integer $k \ge 2$, we denote by $P(k)$ the largest prime divisor of k, and set $P(1) = 1$. We define the following subsets:

- $\mathcal{Q} = \{(\mathbf{x}, \mathbf{y}) \in \mathcal{L} \mid Q \ge P(\gcd(\mathbf{a} \cdot \mathbf{x}, \mathbf{a} \cdot \mathbf{y})) > 1\}$,
- $\mathcal{R} = \{(\mathbf{x}, \mathbf{y}) \in \mathcal{L} \mid M > P(\gcd(\mathbf{a} \cdot \mathbf{x}, \mathbf{a} \cdot \mathbf{y})) > Q\}$,
- $\mathcal{S} = \{(\mathbf{x}, \mathbf{y}) \in \mathcal{L} \mid P(\gcd(\mathbf{a} \cdot \mathbf{x}, \mathbf{a} \cdot \mathbf{y})) \ge M\}$,
- $\mathcal{T} = \{(\mathbf{x}, \mathbf{y}) \in \mathcal{L} \mid p | \gcd(\mathbf{a} \cdot \mathbf{x}, \mathbf{a} \cdot \mathbf{y})$ for some $p \le Q\}$.

Obviously $\mathcal{Q} \subseteq \mathcal{T}$, and

$$1 - \rho_{\mathbf{a}}(M) = M^{-2m} \left(\#\mathcal{Q} + \#\mathcal{R} + \#\mathcal{S}\right) \le M^{-2m} \left(\#\mathcal{T} + \#\mathcal{R} + \#\mathcal{S}\right).$$

For an integer $d \ge 1$, let us denote by $\mathcal{U}_d(M)$ the set of all integer vectors $\mathbf{x} \in \mathbb{N}^m$ with $h(\mathbf{x}) \le M$ and $d | \mathbf{a} \cdot \mathbf{x}$, and put $U_d(M) = \#\mathcal{U}_d(M)$. Because $\gcd(a_1, \ldots, a_m) = 1$, we obviously have $U_p(p) = p^{m-1}$ for any prime p. Then, for any squarefree d, by the Chinese Remainder Theorem, we conclude that $U_d(d) = d^{m-1}$, and $U_d(dK) = K^m d^{m-1}$ for any integer K. Finally, using $U_d(d \lfloor M/d \rfloor) \le U_d(M) \le U_d(d \lceil M/d \rceil)$, we obtain that for $d = o(M/m)$,

$$\begin{aligned} U_d(M) &= (M/d + O(1))^m d^{m-1} = \frac{M^m}{d} \left(1 + O(d/M)\right)^m \\ &= \frac{M^m}{d} \exp\left(O(dm/M)\right) = \frac{M^m}{d} \left(1 + O(md/M)\right). \end{aligned} \tag{3}$$

It is also clear that for any prime p

$$U_p(M) \le (M/p + 1) M^{m-1} = M^m/p + M^{m-1}. \tag{4}$$

By the inclusion exclusion principle we have

$$M^{2m} - \#\mathcal{T} = \sum_{\substack{d \ge 1 \\ 1 \le P(d) \le Q}} \mu(d) U_d(M)^2$$

where μ is the Möbius function. We recall that $\mu(1) = 1$, $\mu(d) = 0$ if $d \ge 2$ is not squarefree, and $\mu(d) = (-1)^{\nu(d)}$ otherwise, where $\nu(d)$ is the number of prime divisors of d; see Section 2.1 of [1]. From the definition of Q we see that any squarefree d with $P(d) \le Q$ does not exceed $(M/m)^{1/2}$. Now from (3) we derive that for such d,

$$U_d(M)^2 = \frac{M^{2m}}{d^2} \left(1 + O(md/M)\right) = \frac{M^{2m}}{d^2} + O\left(mM^{2m-1}/d\right).$$

Therefore

$$\begin{aligned} M^{2m} - \#\mathcal{T}(M) &= \sum_{\substack{d\geq 1\\ 1\leq P(d)\leq Q}} \mu(d)\left(\frac{M^{2m}}{d^2} + O\left(mM^{2m-1}/d\right)\right) \\ &= M^{2m} \sum_{\substack{d\geq 1\\ 1\leq P(d)\leq Q}} \frac{\mu(d)}{d^2} + O\left(mM^{2m-1} \sum_{d\leq (M/m)^{1/2}} d^{-1}\right) \\ &= M^{2m} \prod_{p\leq Q}\left(1-\frac{1}{p^2}\right) + O\left(mM^{2m-1}\ln(M/m)\right). \end{aligned}$$

We now recall that

$$\prod_{p\leq Q}\left(1-\frac{1}{p^2}\right) = \prod_{p}\left(1-\frac{1}{p^2}\right) + O(Q^{-1}) = \zeta(2)^{-1} + O(Q^{-1})$$

see Section 11.4 of [1]. Thus

$$\#\mathcal{T} = (1-\zeta(2)^{-1})M^{2m} + O\left(M^{2m}Q^{-1} + mM^{2m-1}\ln(M/m)\right).$$

When $2M/\ln^2 M \geq m$, then the last term is smaller than the last but one term.

Thus

$$\#\mathcal{T} = (1-\zeta(2)^{-1})M^{2m} + O\left(M^{2m}Q^{-1}\right).$$

For $\#\mathcal{R}$, using (4), and the inequality $(a+b)^2 \leq 2(a^2+b^2)$ we get

$$\begin{aligned} \#\mathcal{R} &\leq \sum_{Q<p<M} U_p(M)^2 \leq 2\sum_{Q<p<M}\left(\frac{M^{2m}}{p^2} + M^{2m-2}\right) \\ &\leq 2M^{2m}\sum_{k>Q}\frac{1}{k^2} + 2M^{2m-2}\sum_{k<M} 1 \\ &= O\left(M^{2m}Q^{-1} + M^{2m-1}\right) = O\left(M^{2m}Q^{-1}\right). \end{aligned}$$

Finally, using (4) again, we derive

$$\begin{aligned} \#\mathcal{S} &\leq \sum_{p\geq M} U_p(M)^2 \leq M^{m-1}\sum_{p\geq M} U_p(M) \\ &= M^{m-1}\sum_{h(\mathbf{x})\leq M}\sum_{\substack{p\geq M\\ p\mid \mathbf{a}\cdot\mathbf{x}}} 1 = M^{m-1}\sum_{h(\mathbf{x})\leq M}\nu(\mathbf{a}\cdot\mathbf{x}) \\ &= O\left(M^{m-1}\sum_{h(\mathbf{x})\leq M}\ln \mathbf{a}\cdot\mathbf{x}\right), \end{aligned}$$

because for any integer $k \geq 2$ we have $\nu(k) = O(\ln k / \ln \ln k)$. Taking into account that $\mathbf{a} \cdot \mathbf{x} \leq mMN$ we finish the proof. □

Corollary 2 *Let $\mathbf{a} \in \mathbb{Z}^m$ be of height at most N. Then for any M such that $M/\max\{m, \ln N\} \to \infty$, we have*

$$\rho_{\mathbf{a}}(M) \geq \zeta(2)^{-1} + o(1).$$

3 Algorithmic Implications

It is easy to see that Corollary 2 implies that for any $a_1, \ldots, a_m$ one can compute $\gcd(a_1, \ldots, a_m)$ probabilistically as the gcd of two integers of asymptotically the same bit lengths as the original data, while the result of [2] only guarantees the same for two integers of bit lengths twice more. The probability of success in both cases is, asymptotically, at least $\zeta(2)^{-1} = 6/\pi^2 = 0.6079\ldots$. Repeating this several times and choosing the smallest result one gets an efficient and reliable algorithm to compute the above gcd which is an attractive alternative to the m-step (deterministic) chain of computation

$$\begin{aligned} \gcd(a_1, \ldots, a_m) &= \gcd(\gcd(a_1, a_2), a_3, \ldots, a_m) \\ &= \gcd(\ldots(\gcd(\gcd(a_1, a_2), a_3), \ldots, a_m). \end{aligned}$$

For illustration, we take l-bit primes $p_1, \ldots, p_m$, $a = p_1 \cdots p_m$, and $a_i = a/p_i$ for $i \leq m$. Then indeed $m-1$ steps are necessary until the gcd, which equals 1, is found.

After $i-1$ steps, the current value of the gcd has about $(m-i)l$ bits, and the reduction of the $(m-1)l$-bit a_{i+1} modulo this gcd takes about $2l^2(m-i)(i-1)$ operations in naive arithmetic; see [3], Section 2.4. This comes to a total of about $l^2m^3/3$ operations. If one gcd of n-bit integers costs about cn^2 operations, for a constant c, then all the gcds required amount to $cm^3/3$, for a grand total of $l^2m^3(1+c)/3$ operations.

In our algorithm, we can choose x_i and y_i of $\ln(ml)$ bits. The inner products together cost just over $2lm^2 \ln(ml)$ operations, and the single gcd about cl^2m^2. The latter is the dominant cost, and thus our algorithm is faster by a factor of about $m/3$ than the standard one.

In other words, if $k < m/3$, maybe $k \approx \sqrt{m}$, and confidence at least $1-\zeta(2)^{-k}$ is sufficient, then the k-fold repetition of our algorithm is faster. (In practice, one would not just repeat, but reduce the inputs modulo the gcd candidate obtained so far, and either find that it divides all of them and thus is the true gcd, or continue with the smaller values.)

The advantage of our method evaporates when one uses fast arithmetic.

The worst-case example is not quite as esoteric as it may look. In resultant and subresultant computations with several integer polynomials in several variables, nontrivial gcds occur with definite patterns.

4 Conclusion and Open Questions

It would be interesting to evaluate the constant implicit in the bound of Theorem 1. This should be possible, but may involve some nontrivial amount of technical details.

We believe that in fact $\rho_{\mathbf{a}}(M) \sim \zeta(2)^{-1}$ under the condition of Corollary 2 (or some similar conditions maybe marginally more restrictive). We believe that better sieving technique should produce such a result. Although it may have no algorithmic application it is a natural question which would be interesting to resolve.

Finally, we remark that the approach of [2] leads to an algorithm for an the extended gcd problem; see [3] for the background on this problem. Namely, solving the the extended gcd problem for $\mathbf{a} \cdot \mathbf{x}$ and $\mathbf{a} \cdot \mathbf{x}$ we obtain a relation

$$c_1 a_1 + \ldots + c_n a_n = d$$

for some integers $c_1, \ldots, c_n, d$ with $d > 0$. Repeating this the appropriate number of times, given by (2), and choosing the relation with the smallest value of d, we solve he extended gcd problem with probability at least $1 - \delta$.

References

1. T. M. Apostol, *Introduction to analytic number theory*, Springer-Verlag, NY, 1976.
2. G. Cooperman, S. Feisel, J. von zur Gathen and G. Havas, 'GCD of many integers', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **1627** (1999), 310–317.
3. J. von zur Gathen and J. Gerhard, *Modern computer algebra*, Cambridge University Press, Cambridge, 2003.
4. G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, Oxford Univ. Press, Oxford, 1979.