CIRCULANT GRAPHS AND GCD AND LCM OF SUBSETS

JOACHIM VON ZUR GATHEN AND IGOR E. SHPARLINSKI

Abstract. Given two sets A and B of integers, we consider the problem of finding a set $S \subseteq A$ of the smallest possible cardinality such the greatest common divisor of the elements of $S \cup B$ equals that of those of $A \cup B$. The particular cases of $B = \emptyset$ and $\#B = 1$ are of special interest and have some links with graph theory. We also consider the corresponding question for the least common multiple of the elements. We establish NP-completeness and approximation results for these problems by relating them to the Minimum Cover Problem.

1. Introduction

1.1. **Description of the problem and motivation.** For a set A of integers, $\gcd(A)$ and $\mathrm{lcm}(A)$ denote the greatest common divisor (gcd) and the least common multiple (lcm) of the elements of A, respectively. We consider some questions of how gcd and lcm behave on various subsets S of the original set A.

We are interested in both designing algorithms to construct such sets S with prescribed properties of $\gcd(S)$ and $\mathrm{lcm}(S)$ and also in upper and lower bounds on what one can possibly achieve.

We consider the question of finding a subset $S \subseteq A$ of the smallest possible cardinality with minimal gcd, namely, $\gcd(S) = \gcd(A)$, or with maximal lcm, namely, $\mathrm{lcm}(S) = \mathrm{lcm}(A)$. We also consider a modification of this question where we impose that a specific set B of integers be contained in S. This B may contain elements of A. This question arises in the theory of *circulant graphs* and is a special case of graph *editing problems*, see [Damaschke & Molokov, 2012], [Golovachy, 2013], [Mathieson, 2010] and [Mathieson & Szeider, 2012] for the background and further references.

Date: February 25, 2014.

1991 *Mathematics Subject Classification.* 11Y16, 68R10.

Key words and phrases. gcd, lcm, minimum cover, circulant graphs, graph editing.

To explain this connection we recall that an (undirected) *circulant graph* $G(A, m)$ on m nodes, labelled $0, 1, \ldots, m-1$, is defined by a set A of integers called *links*, where the nodes i and j are connected if and only if $|i-j| \equiv a \bmod m$ for some $a \in A$. Clearly, $G(A, m)$ is connected if and only if $\gcd(A \cup \{m\}) = 1$. Thus it is natural to ask how many links can at most be removed from A so that the new circulant graph is still connected. This leads to the above question with $B = \{m\}$.

The above can be generalized as follows:

Question 1. Given two sets A and B of positive integers, find a subset $S \subseteq A$ of the smallest possible size with $\gcd(S \cup B) = \gcd(A \cup B)$.

Similarly, we also ask:

Question 2. Given two sets A and B of positive integers, find a subset $S \subseteq A$ of the smallest possible size with $\operatorname{lcm}(S \cup B) = \operatorname{lcm}(A \cup B)$.

We first formalize these questions as decision problems.

Problem 3. *Minimum subset with minimal* gcd, MinGcd

Input: Sets A and B of positive integers, positive integer k.
Question: Does A contain a subset S with $\#S \le k$ and $\gcd(S \cup B) = \gcd(A \cup B)$?

Problem 4. *Minimum subset with maximal* lcm, MaxLcm

Input: Sets A and B of positive integers, positive integer k.
Question: Does A contain a subset S with $\#S \le k$ and $\operatorname{lcm}(S \cup B) = \operatorname{lcm}(A \cup B)$?

The input size of an instance (A, B) for both MinGcd and MaxLcm is naturally defined as

$$I(A, B) = \sum_{a \in A \cup B} \lceil \log(a+1) \rceil$$

where $\log z$ denotes the binary logarithm of $z \ge 1$.

For each of these (and other similar) problems X, we denote as OPT-X the corresponding optimization problem, where one has to find subsets as described with minimal k.

1.2. **Main Results.** We can now formulate our main results.

Theorem 5. MinGcd *and* MaxLcm *are* NP-*complete.*

Furthermore, a combination of the classical greedy approximation algorithm of [Johnson, 1974, Theorem 4] and known inapproximability results, see, for example, [Alon & Moshkovitz & Safra, 2006, Theorem 7], yield the following.

Theorem 6. OPT-MinGcd *and* OPT-MaxLcm *can be approximated in polynomial time within a factor* $O(\log I(A,B))$*, but not within a factor* $o(\log I(A,B))$ *if* $\mathrm{P} \neq \mathrm{NP}$.

2. Reductions Between Various Problems

2.1. Reduction to $B = \varnothing$.

We start by constructing from $A, B \subseteq \mathbb{Z}$ a set $A_B \subseteq \mathbb{Z}$ so that

$$\text{OPT-MinGcd}(A,B) = \text{OPT-MinGcd}(A_B, \varnothing).$$

This reduces the general case to the special situation where $B = \varnothing$. Moreover, given a minimum solution set S for one of the two problems, one can easily find a solution for the other one.

For any integer a, we define the nonnegative integer

$$a_B = \gcd(\{a\} \cup B),$$

and apply this element-wise to any $S \subseteq \mathbb{Z}$:

$$S_B = \{a_B \colon a \in S\}.$$

We claim that for any $S \subseteq A$ we have

$$\gcd(S \cup B) = \gcd(S_B). \tag{1}$$

For any $c \in \mathbb{Z}$, we have

$$\begin{aligned} c \mid \gcd(S \cup B) &\iff \forall a \in S\ \forall b\ \in B \quad c \mid a \text{ and } c \mid b \\ &\iff (\forall a \in S \quad c \mid a) \text{ and } c \mid \gcd(B) \\ &\iff \forall a \in S \quad c \mid a_B \iff c \mid \gcd(S_B). \end{aligned}$$

In particular, we have $\gcd(A \cup B) = \gcd(A_B)$.

Distinct $a \in A$ may yield the same a_B. However, if $S \subseteq A$ has minimal size with $\gcd(S \cup B) = \gcd(A \cup B)$, then $a \mapsto a_B$ is injective on S, and $\#S = \#S_B$. Thus

$$\text{OPT-MinGcd}(A,B) \geq \text{OPT-MinGcd}(A_B, \varnothing).$$

For the reverse direction, we take a section σ of $a \mapsto a_B$ on A, so that $\sigma(b) \in A$ and $(\sigma(b))_B = b$ for $b \in A_B$. For any $T \subseteq A_B$ of minimal size with $\gcd(T) = \gcd(A_B)$, we claim that

$$\gcd(\sigma(T) \cup B) = \gcd(A \cup B).$$

This follows from (1), since $(\sigma(T))_B = T$ and

$$\gcd(A \cup B) = \gcd(A_B) = \gcd(T) = \gcd(\sigma(T) \cup B).$$

We have $\#\sigma(T) \leq \#T$ and $\gcd((\sigma(T))_B) = \gcd(A_B)$. The minimality of $\#T$ implies that $\#\sigma(T) = \#T$ and thus

$$\text{OPT-MinGcd}(A,B) \leq \text{OPT-MinGcd}(A_B, \varnothing).$$

Overall, it follows that the minimal solution sizes for (A, B) and A_B are equal, and that the solution sets are related by the above correspondence. Clearly the set A_B can be constructed in time polynomial in $I(A, B)$. Thus both the decision and the optimization versions of the general and the special cases are polynomial-time equivalent.

So from now on we assume that the input consists of one set A and denote by $I(A) = I(A, \emptyset)$ the input size.

2.2. **Minimum Cover Problem.** We present polynomial time reductions between MinGcd, MaxLcm and the following problem, which is well studied in complexity theory.

Problem 7. *Minimum cover*, MinCover

> **Input:** List $\mathcal{C}$ of subsets of a finite set X, positive integer k.
> **Question:** Does $\mathcal{C}$ contain a cover for X of size k or less, that is, a subset $\mathcal{D} \subseteq \mathcal{C}$ with $\#\mathcal{D} \leq k$ such that every element of X belongs to at least one member of $\mathcal{D}$?

Furthermore, let n be the input size, usually about $\#\mathcal{C} \cdot \log m$ if $X = \{1, \ldots, m\}$. Then OPT-MinCover can be approximated in polynomial time within a factor of $O(\log n)$, but no smaller factor (unless $P = NP$), see [Alon & Moshkovitz & Safra, 2006].

It is well known that MinCover is NP-complete, see, for example, [Garey & Johnson, 1979, Problem SP5, Section A.3.1]. In the next subsections, we present various reductions between MinCover and our problems. The latter are trivially in NP, and their reduction to MinCover transfers approximation algorithms for the latter to approximation algorithms for our problems. On the other hand, the reductions from MinCover to our problems show that the latter cannot be approximated too well.

2.3. **Reduction from MaxLcm to MinCover.** Let us take an instance (A, k) of MaxLcm. We compute a *coprime basis* (B, e) of A, where B consists of pairwise coprime integers $b \geq 2$ and $e\colon A \times B \longrightarrow \mathbb{N}$ is such that $a = \prod_{b \in B} b^{e(a,b)}$ for all $a \in A$. By dropping the unneeded elements b where $e(a, b) = 0$ for all $a \in A$ from B, we may assume that

$$\forall b \in B \ \exists\, a \in A\colon e(a, b) \geq 1. \tag{2}$$

We recall that [Bach & Shallit, 1996, Section 4.8] discuss coprime bases (under the designation of *gcd-free basis*) and show that one can be computed with $O(I(A)^2)$ bit operations, where, as before, $I(A)$ is the input size. They use classical arithmetic. According to [Bernstein, 2005], fast arithmetic yields an algorithm using $I(A)(\log I(A))^{O(1)}$ operations.

By the above, the size of B is polynomial in that of A. We note that the size of B can actually be much smaller than that of A: Take the first m primes, all exponent vectors e in $\{1,2\}^m$, and then all 2^m values $a_e = \prod_{1\leq i\leq m} p_i^{e_i}$. Then the coprime basis B consists of just these m primes and size(A) is only logarithmic in size(B). That is no worry, since we only use this reduction to derive good approximations for MinCover (which do not exist by the hardness result mentioned above) from good approximations to our problems; hence the latter do not exist either.

For $b \in B$, we let $d(b) = \max\{e(a,b) \colon a \in A\}$. Thus $d(b) \geq 1$ by (2), and $\operatorname{lcm}(A) = \prod_{b\in B} b^{d(b)}$; see also [Bach & Shallit, 1996, Corollary 4.8.2]. For $a \in A$, we set

$$C_a = \{b \in B \colon e(a,b) = d(b)\}.$$

We now take a subset $E \subseteq A$ such that $\{C_a \colon a \in A\} = \{C_a \colon a \in E\}$ and the C_a in the latter set are pairwise distinct. Clearly this can be done in time polynomial in $I(A)$. It is also clear that $\operatorname{lcm}(E) = \operatorname{lcm}(A)$.

Now we consider the MinCover instance with $X = B$ and $\mathcal{C} = \{C_a \colon a \in E\}$. For $S \subseteq E$, we consider $\mathcal{D} = \{C_a \colon a \in S\}$. Then $\#\mathcal{D} = \#S$, and

$$\begin{aligned}
\operatorname{lcm}(S) = \operatorname{lcm}(E) = \operatorname{lcm}(A) &\iff \forall b \in B \quad b^{d(b)} \mid \operatorname{lcm}(S) \\
&\iff \forall b \in B\ \exists a \in S \quad b^{d(b)} \mid a \\
&\iff \forall b \in B\ \exists a \in S \quad e(a,b) = d(b) \\
&\iff \forall b \in B = X\ \exists a \in S \quad b \in C_a \\
&\iff \mathcal{D} \text{ covers } X.
\end{aligned}$$

Thus a solution S of MaxLcm with $\#S \leq k$ implies one of MinCover with $\#\mathcal{D} \leq k$.

Conversely, given a cover $\mathcal{D} = \{C_a \colon a \in S\}$ of X with $\#\mathcal{D} \leq k$ we conclude that $\#S \leq k$, since the sets C_a for $a \in E$ are pairwise distinct.

Thus the size of the smallest set $S \subseteq A$ with $\operatorname{lcm}(S) = \operatorname{lcm}(A)$ and the size of the smallest cover $\mathcal{D}$ of X coincide. This concludes the reduction.

2.4. Reduction from MinGcd to MinCover.

We replace $d(b)$ in the previous reduction by $g(b) = \min\{e(a,b) \colon a \in A\}$. Then

$$\gcd(A) = \prod_{b\in B} b^{g(b)}.$$

We see that $\gcd(E) = \gcd(A)$ divides $\gcd(S)$ and in the argument for $\gcd(E) = \gcd(S)$, a divisibility $b^{d(b)} \mid u$ has to be replaced by $b^{g(b)+1} \nmid u$. Otherwise the argument goes through unchanged.

2.5. **Reduction from MinCover to MaxLcm.** We are given a list $\mathcal{C}$ of sets $C_1, \ldots, C_l \subseteq X$, where $X = \{1, \ldots, m\}$, and $k \geq 1$. We may assume that $X = \bigcup_{i \leq l} C_i$ and $k \leq l$, otherwise the MinCover problem is trivial. We let $p_1 < p_2 < \cdots < p_m$ be the first m prime numbers, $a = \prod_{j \in X} p_j$,

$$a_i = \prod_{j \in C_i} p_j,$$

for $i \leq l$ and $A = \{a_1, \ldots, a_l\}$. Thus $a = \operatorname{lcm}(A)$. We use the same value of k for both problems. Since $p_m \leq (1 + o(1))\, m \ln m$ as $m \to \infty$, the bit size of (A, k) is in $O(lm \log m)$. The set A can be computed in time polynomial in lm, using the sieve of Eratosthenes for generating the primes.

Suppose that $I \subseteq \{1, \ldots, m\}$ is such that $\#I \leq k$ and $\operatorname{lcm}(S) = \operatorname{lcm}(A)$, where $S = \{a_i \colon i \in I\}$. Let

$$\mathcal{D} = \{C_i \colon i \in I\}.$$

Then $\#\mathcal{D} \leq k$. Furthermore, for any $j \in X$, p_j divides $\operatorname{lcm}(A) = \operatorname{lcm}(S)$ and hence a_i for some $i \in I$. It follows that $j \in C_i \in \mathcal{D}$. Thus $\mathcal{D}$ covers X.

On the other hand, suppose that $I \subseteq \{1, \ldots, m\}$ is such that $\#I \leq k$ and $\mathcal{D} = \{C_i \colon i \in I\}$ covers X. Then $S = \{a_i \colon i \in I\}$ satisfies $\#S \leq k$ and $\operatorname{lcm}(S) = a = \operatorname{lcm}(A)$.

2.6. **Reduction from MinCover to MinGcd.** For an analogous reduction to MinGcd, we replace a_i by a/p_i in the above.

3. Proofs of Main Results

We start with the upper bounds claimed in Theorems 5 and 6. The fact that MaxLcm and MinGcd are NP-complete is trivial. Furthermore, the reductions of Sections 2.3 and 2.4 show that the know approximation algorithms for MinCover also yield ones for our problems.

Furthermore, our claimed lower bounds (NP-hardness and inapproximability) follow from the reductions in Sections 2.5 and 2.6 from our problems to MinCover, together with the NP-hardness and inapproximability of MinCover, as cited above.

Acknowledgments

The authors are grateful to Fedor Fomin for some references on the approximability and inapproximability of the set-cover problem.

The first author's work was supported by the B-IT Foundation and the Land Nordrhein-Westfalen. The second author's work was supported by the Australian Research Council grants DP110100628 and DP140100118.

References

[Alon & Moshkovitz & Safra, 2006] Noga Alon, Dana Moshkovitz and Shmuel Safra, 'Algorithmic construction of sets for k-restrictions', *ACM Trans. Algorithms*, **2** (2006), 153–177.

[Bach & Shallit, 1996] Eric Bach and Jeffrey Shallit, 'Algorithmic Number Theory, Vol.1: Efficient Algorithms', MIT, Cambridge-MA, 1996.

[Bernstein, 2005] Daniel J. Bernstein, 'Factoring into coprimes in essentially linear time', *J. of Algorithms*, **54** (2005), 1–30.

[Damaschke & Molokov, 2012] Peter Damaschke and Leonid Molokov, 'Parameterized reductions and algorithms for a graph editing problem that generalizes vertex cover', *Theor. Comp. Sci.*, **452** (2012), 39–46.

[Garey & Johnson, 1979] Michael R. Garey and David S. Johnson, 'Computers and intractability: A Guide to the Theory of NP-Completeness', W. H. Freeman and Co., 1979, San Francisco CA.

[Golovachy, 2013] Petr A. Golovachy, 'Editing to a connected graph of given degrees', *Preprint*, 2013, (available from `http://arxiv.org/abs/1308.180`).

[Johnson, 1974] David S. Johnson, 'Approximation algorithms for combinatorial problems', *J. Comput. Syst. Sci.*, **9** (1974), 256–278.

[Mathieson, 2010] Luke Mathieson, 'The parameterized complexity of editing graphs for bounded degeneracy', *Theor. Comp. Sci.*, **411** (2010), 34–36.

[Mathieson & Szeider, 2012] Luke Mathieson and Stefan Szeider, 'Editing graphs to satisfy degree constraints: A parameterized approach', *J. Comp. Syst. Sci.*, **78** (2012), 179–191.

B-IT, Universität Bonn, 53113 Bonn, Germany
E-mail address: `gathen@bit.uni-bonn.de`

Department of Pure Mathematics, University of New South Wales, Sydney, NSW 2052, Australia
E-mail address: `igor.shparlinski@unsw.edu.au`