ON THE KERNEL OF THE COPRIME GRAPH OF INTEGERS

J.W. Sander
Institut für Mathematik und Angewandte Informatik, Universität Hildesheim,
D-31141 Hildesheim, Germany
sander@imai.uni-hildesheim.de

T. Sander
Institut für Mathematik, Technische Universität Clausthal,
Clausthal-Zellerfeld, Germany
torsten.sander@math.tu-clausthal.de

Received: 4/27/09, Accepted: 5/25/09, Published: 11/19/09

Abstract
Let \((V, E)\) be the coprime graph with vertex set \(V = \{1, 2, \ldots, n\}\) and edges \((i, j) \in E\) if \(\gcd(i, j) = 1\). We determine the kernels of the coprime graph and its loopless counterpart as well as so-called simple bases for them (in case such bases exist), which means that basis vectors have entries only from \([-1, 0, 1]\). For the loopless version knowledge about the value distribution of Mertens’ function is required.

1. Introduction
For each integer \(n > 1\) the “traditional” coprime graph \(TCG_n = (V, E)\) has the vertex set \(V = \{1, 2, \ldots, n\}\) and edges \((i, j) \in E\) if and only if \(\gcd(i, j) = 1\). Obviously, \(TCG_n\) has a loop at 1. Since one usually prefers loopless graphs, we also consider the slightly modified loopless coprime graph \(LCG_n = (V', E')\) with \(V' = V\) and \(E' = E \setminus \{(1, 1)\}\). With regard to what we intend to prove \(LCG_n\) requires more involved techniques than \(TCG_n\). For that reason we shall mainly deal with \(LCG_n\) and comment only in the final section on the corresponding results for \(TCG_n\), which can be obtained by the same method with less effort.

The first problem concerning the coprime graph and its subgraphs was introduced by Erdős [8] in 1962. Meanwhile interesting features relating number theory and graph theory have been unearthed (for various “graphs on the integers” the reader is referred to [17], Chapter 20, 7.4):

- In 1984 Pomerance and Selfridge [18] proved Newman’s coprime mapping conjecture: If \(I_1 = \{1, 2, \ldots, n\}\) and \(I_2\) is any interval of \(n\) consecutive integers, then there is a perfect coprime matching from \(I_1\) to \(I_2\). Note that the statement is not true if \(I_1\) is also an arbitrary interval of \(n\) consecutive integers. Example: \(I_1 = \{2, 3, 4\}\) and \(I_2 = \{8, 9, 10\}\); any one-to-one correspondence between \(I_1\) and \(I_2\) must have at least one pair of even numbers in the correspondence.
- In a series of papers between 1994 and 1996 Ahlswede and Khachatrian (cf. [2], [3], [4]) and very recently Ahlswede and Blinovsky [1] proved results on extremal sets without coprime elements, extremal sets without \(k + 1\) pairwise coprime elements, and sets of integers with pairwise common divisors. Two edges in the coprime graph are not coprime if they are connected in the com-
plemernary graph $\mathcal{T}CG_n$ of $TCG_n$. Therefore one has to search for maximal complete subgraphs in $\mathcal{T}CG_n$.


Let $A_n = (a_{i,j})_{n \times n}$ be the adjacency matrix of $LCG_n$, i.e.,

$$a_{i,j} = \begin{cases} 
0 & \text{if } \gcd(i, j) > 1 \text{ or } i = j = 1, \\
1 & \text{otherwise}.
\end{cases}$$

(1)

Apparently $LCG_n$ is an undirected loopless graph, and $A_n$ is symmetric.

For several decades spectra and eigenspaces of graphs (cf. [11]), that is, spectra and eigenspaces of their adjacency matrices, have been studied for quite a few different types of graphs (for references see [6], [7] or [10]). For reasons like characterization of graphs or computational advantages it is of particular interest to find so-called simple bases (all entries are $-1, 0, 1$) for eigenspaces, especially for the kernel of a graph. Such bases can be found for trees and forests (see [20], [5]), unicyclic graphs [21] and powers of circuit graphs [22].

Computational experiments provided evidence for the following observations:

- The dimensions of the kernels of the coprime graphs $TCG_n$ and $LCG_n$, respectively, are growing with $n$.

- These kernels always have a simple basis in the above sense.

It is the purpose of this work to clarify the observations made. In fact, we shall prove precise formulae for $\dim \ker TCG_n$ and $\dim \ker LCG_n$ and construct an explicit simple basis for each of them – if one exists. In order to determine those kernels which have no simple basis, results about the Mertens function

$$M(n) := \sum_{k=1}^{n} \mu(k)$$

will be involved, where $\mu(n)$ denotes Möbius’ function.

2. Basic Facts

We denote by $\kappa(m) = \prod_{p \in \mathbb{P}, p|m} p$ the squarefree kernel of a positive integer $m$. For each squarefree integer $k > 1$ the vector $b = (b_1, b_2, \ldots, b_n) \in \mathbb{R}^n$ is called $k$-basic if for some $m$, $k < m \leq n$, satisfying $\kappa(m) = k$,

$$b_j = \begin{cases} 
1 & \text{for } j = k, \\
-1 & \text{for } j = m, \\
0 & \text{otherwise}.
\end{cases}$$
If \( \mathbf{b} \in \mathbb{R}^n \) is \( k \)-basic for some squarefree \( k \), we call it a basic vector. The set of all basic vectors \( \mathbf{b} \in \mathbb{R}^n \) will be denoted by \( \mathcal{B}_n \).

**Lemma 1** The number \( \nu(n) := |\mathcal{B}_n| \) of basic vectors satisfies

\[
\nu(n) = n - \sum_{k \leq n} |\mu(k)|.
\]

**Proof.** Associate with each non-squarefree positive integer \( m \leq n \) the basic vector \( \mathbf{b} \in \mathbb{R}^n \) defined by

\[
b_j = \begin{cases} 
1 & \text{for } j = \kappa(m), \\
-1 & \text{for } j = m, \\
0 & \text{otherwise.}
\end{cases}
\]  

(2)

This correspondence is apparently one-to-one. Now \( \nu(n) \) precisely counts the non-squarefree positive integers \( m \leq n \).

**Proposition 2** Let \( n > 1 \) be an arbitrary integer.

(i) If \( \mathbf{b} \in \mathcal{B}_n \) then \( \mathbf{b} \in \ker A_n \).

(ii) \( \mathcal{B}_n \) is linearly independent over \( \mathbb{R} \).

**Proof.** (i) Let \( \mathbf{b} \in \mathcal{B}_n \), i.e., \( \mathbf{b} \) is \( k \)-basic for some squarefree \( k > 1 \). Hence \( \mathbf{b} \) has entries 0 apart from \( b_k = 1 \) and \( b_m = -1 \) for some \( m \) satisfying \( k < m \leq n \) and \( \kappa(m) = k \). Now let \( \mathbf{a}_i \) be the \( i \)-th row vector of \( A_n \). Then we have for the scalar product

\[
\mathbf{a}_i \cdot \mathbf{b} = a_{i,k} - a_{i,m} = 0,
\]

because \( k \) and \( m \) have the same prime factors and therefore \( \gcd(i, k) \) and \( \gcd(i, m) \) are both 1 or both greater than 1. This means that \( \mathbf{b} \) belongs to \( \ker A_n \).

(ii) Let \( m \leq n \) be a non-squarefree positive integer. Then there is precisely one basic vector \( \mathbf{b} \in \mathcal{B}_n \) satisfying \( b_m = -1 \), namely the vector defined in (2). All other vectors \( \mathbf{b}' \in \mathcal{B}_n \) have \( b'_m = 0 \). Thus, \( \mathcal{B}_n \) is linearly independent.

From Proposition 2 we obtain immediately

**Corollary 3** For any integer \( n > 1 \) we have \( \dim_{\mathbb{R}} \ker A_n \geq \nu(n) \).

We shall prove in the sequel that in fact \( \dim_{\mathbb{R}} \ker A_n = \nu(n) \) for most \( n \). This was suggested by numerical calculations. It turns out, however, that there are infinitely many exceptions.

### 3. Truncated Möbius Inversion and Mertens’ Function

In the sequel we make use of the truncated version of the Möbius inversion formula (cf. [13, Chapter 6.4, Theorem 4.1]). Then an important role is played by Mertens’...
well-known function

\[ M(n) := \sum_{k=1}^{n} \mu(k). \]

Trivially \(|M(n)| \leq n\) for all \(n\). The relevance of this function becomes immediately

\[ \text{clear from the facts that } M(n) = o(n) \text{ is equivalent to the prime number theorem and } M(n) = O(n^{\frac{1}{2}+\varepsilon}) \text{ is equivalent to the Riemann hypothesis.} \]

The famous Mertens conjecture from 1897 saying that \(|M(n)| < \sqrt{n}\) for all \(x > 1\) was disproved by Odlyzko and te Riele [15] in 1985. For our purpose it is essential to know something about the value distribution of \(M(n)\) (see Remark 6(ii)).

**Proposition 4** Let \( n > 1 \) be an arbitrary integer. A vector \( b = (b_1, \ldots, b_n) \in \mathbb{R}^n \) lies in \( \text{Ker}A_n \) if and only if

\[ (M(n) - 1) b_1 = 0 \]

and for \( 2 \leq k \leq n, \mu(k) \neq 0 \)

\[ \sum_{\substack{j=1 \\
 j \equiv 0 \mod k}}^{n} b_j - b_1 = 0. \]

**Proof.** It is well-known that the summatory function \( \varepsilon(n) = \sum_{d|n} \mu(d) \) of the Möbius function satisfies

\[ \varepsilon(n) = \begin{cases} 1 & \text{for } n = 1, \\ 0 & \text{for } n > 1. \end{cases} \]

This implies

\[ \varepsilon(\gcd(i, j)) = \sum_{\substack{d|i \\
 d|j}} \mu(d) = \begin{cases} 1 & \text{for } \gcd(i, j) = 1, \\ 0 & \text{for } \gcd(i, j) > 1, \end{cases} \]

and therefore we have for the entries \( a_{i,j} \) of the adjacency matrix of \( LCG_n \) (see (1))

\[ a_{i,j} = \varepsilon(\gcd(i, j)) - \gamma_{ij} \]

for all \( 1 \leq i, j \leq n \), where \( \gamma_{ij} \) equals 1 for \( i = j = 1 \) and 0 otherwise.

For a given vector \( b = (b_1, b_2, \ldots, b_n) \in \mathbb{R}^n \) let \( f : \{1, 2, \ldots, n\} \to \mathbb{R} \) be defined by

\[ f(k) := \mu(k) \sum_{\substack{j=1 \\
 j \equiv 0 \mod k}}^{n} b_j. \]
By (5) and (6) it follows that the summatory function \( g \) of \( f \) satisfies, for \( 1 \leq i \leq n \),

\[
g(i) := \sum_{d|i} f(d) = \sum_{d|i} \mu(d) \sum_{j=\lfloor d/i \rfloor}^{n} b_j
\]

\[
= \sum_{j=1}^{n} b_j \sum_{d|i} \mu(d) = \sum_{j=1}^{n} b_j \varepsilon(\gcd(i,j))
\]

\[
= \sum_{j=1}^{n} (a_{i,j}b_j + \gamma_{ij}b_j) = \sum_{j=1}^{n} a_{i,j}b_j + \gamma_{i1}b_1.
\]

A vector \( \mathbf{b} = (b_1, \ldots, b_n) \in \mathbb{R}^n \) lies in \( \text{Ker}A_n \) if and only if \( \sum_{j=1}^{n} a_{i,j}b_j = 0 \) for \( 1 \leq i \leq n \). By (7) this is equivalent to \( g(i) = \gamma_{i1}b_1 \) for \( 1 \leq i \leq n \). By the truncated version of the Möbius inversion formula (cf. [13], Chapt. 6.4, Theor. 4.1) this means that, for \( 1 \leq k \leq n \),

\[
f(k) = \sum_{d|k} \mu(d)g\left(\frac{k}{d}\right) = \mu(k)g(1) = \mu(k)b_1,
\]

and hence by the definition of \( f \),

\[
\mu(k) \sum_{j=1}^{n} b_j = \mu(k)b_1.
\]

So far we have shown that \( \mathbf{b} \in \text{Ker}A_n \) if and only if

\[
\sum_{j=1}^{n} b_j = b_1 \quad (1 \leq k \leq n, \ \mu(k) \neq 0).
\]

We have

\[
\sum_{k=2}^{n} \mu(k) \sum_{\mu(k) \neq 0} b_j = \sum_{k=2}^{n} \mu(k) \sum_{j=0 \mod k} b_j = \sum_{j=2}^{n} b_j \sum_{k=2}^{n} \mu(k) = \sum_{j=2}^{n} b_j \sum_{k=2}^{n} \mu(k)
\]

\[
= \sum_{j=2}^{n} b_j (\varepsilon(j) - 1) = -\sum_{j=2}^{n} b_j.
\]
and by adding the corresponding equations for \( k = 2, \ldots, n \) with \( \mu(k) \neq 0 \) in (8) we obtain

\[
- \sum_{j=2}^{n} b_j = \sum_{k=2}^{n} \sum_{j=1}^{n} \mu(k) b_j = b_1 \sum_{k=2}^{n} \mu(k) = b_1(M(n) - 1).
\]

The addition of this to the equation for \( k = 1 \) in (8) gives

\[
b_1 = \sum_{j=1}^{n} b_j - \sum_{j=2}^{n} b_j = b_1 + b_1(M(n) - 1),
\]

and replacing the equation for \( k = 1 \) in (8) by this one does not change the set of solutions. This completes the proof of the proposition. \( \Box \)

4. Main Results

**Theorem 5** For any integer \( n > 1 \) we have

\[
\dim_{\mathbb{R}} \text{Ker}LCG_n = \begin{cases} 
\nu(n) & \text{for } M(n) \neq 1, \\
\nu(n) + 1 & \text{for } M(n) = 1,
\end{cases}
\]

(10)

where \( \nu(n) \) is defined in Lemma 1. Consequently

\[
\dim_{\mathbb{R}} \text{Ker}LCG_n = \left( 1 - \frac{6}{\pi^2} \right) n + O(\sqrt{n}).
\]

(11)

**Proof.** By Proposition 4 a vector \( \mathbf{b} = (b_1, \ldots, b_n) \in \mathbb{R}^n \) lies in \( \text{Ker}LCG_n = \text{Ker}A_n \) if and only if \( \mathbf{b} \) satisfies the homogeneous system consisting of the linear equations (4) for \( 2 \leq k \leq n, \mu(k) \neq 0 \), and, in addition, equation (3). Therefore we obtain

\[
\begin{pmatrix}
  b_2 & +b_4 & +b_6 & +b_8 & +b_{10} & \cdots & -b_1 & = 0 \\
  b_3 & +b_6 & +b_9 & +b_{10} & \cdots & -b_1 & = 0 \\
  b_5 & +b_9 & +b_{10} & \cdots & -b_1 & = 0 \\
  b_6 & \vdots & \vdots & \vdots & \vdots & \vdots \\
  b_7 & \vdots & \vdots & \vdots & \vdots & \vdots \\
  \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
  b_{10} & \cdots & -b_1 & = 0 \\
  \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
  \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
  \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
  \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
  (M(n) - 1)b_1 & = 0
\end{pmatrix}
\]

(12)
Apparently (12) is a homogeneous system in row-echelon form with \( n \) variables. Hence the rank of the coefficient matrix \( B_n \) obviously satisfies

\[
\text{rank } B_n = \begin{cases} 
\sum_{k=1}^{n} |\mu(n)| & \text{for } M(n) \neq 1, \\
\sum_{k=1}^{n} |\mu(n)| - 1 & \text{for } M(n) = 1.
\end{cases}
\]

Consequently \( \dim \mathcal{Ker} LCG_n = \dim \mathcal{Ker} A_n = n - \text{rank } B_n \), and by Lemma 1 this proves (10).

It is well-known that

\[
\sum_{k=1}^{n} |\mu(k)| = \frac{1}{\zeta(2)} n + O(\sqrt{n}) = \frac{6}{\pi^2} n + O(\sqrt{n})
\]

(cf. [12], p. 270). Now Lemma 1 and (10) imply (11). \( \square \)

Remarks 6

(i) The proof of Theorem 5 showed that

\[
b \in \mathcal{Ker} LCG_n \iff B_n \tilde{b} = 0,
\]

where \( B_n \) is the coefficient matrix of (12) and \( \tilde{b} := (b_2, b_3, \ldots, b_n, b_1) \).

(ii) Apparently \( \dim \mathcal{Ker} LCG_n \) depends on the value of \( M(n) \), more precisely on whether \( M(n) = 1 \) or not. Results of Pintz and others (cf. [16]) show that \( M(n) \) oscillates between \( \pm \sqrt{n} \), and since \( |M(n+1) - M(n)| \leq 1 \), each value between these bounds is attained infinitely many times. In particular \( |\{ n : M(n) = 1 \}| = \infty \). The smallest integers \( n > 1 \) with \( M(n) = 1 \) are

\[n = 94, 97, 98, 99, 100, 146, 147, 148, \ldots \]

Theorem 7 If \( n \) is an integer satisfying \( M(n) \neq 1 \), we have the following:

(i) \( B_n \) is a basis of \( \mathcal{Ker} LCG_n \).

(ii) \( \mathcal{Ker} LCG_n \) has a simple basis, i.e., the components of all basis vectors are 0, 1 or \(-1\).

(iii) Let \( \iota : \mathbb{R}^n \to \mathbb{R}^{n+1} \) with \( \iota(b_1, \ldots, b_n) := (b_1, \ldots, b_n, 0) \) be the canonical injection. Then we have \( \iota(\mathcal{Ker} LCG_n) \subseteq \mathcal{Ker} LCG_{n+1} \).

Proof. The assertion (i) follows from Theorem 5, Lemma 1 and Proposition 2. This immediately implies (ii).

It remains to show (iii). Note that putting a zero at the end of a basic vector of \( B_n \) turns it into a basic vector of \( B_{n+1} \), so that \( \iota(B_n) \subseteq B_{n+1} \). The desired result now follows from part (i). \( \square \)
Theorems 5 and 7 imply that in case \( M(n) = 1 \) the linearly independent set \( \mathcal{B}_n \) of basic vectors needs a single additional vector \( \tilde{b} \in \mathbb{R}^n \), say, to obtain a basis of \( \ker LCG_n \). Such a vector \( \tilde{b} = (\tilde{b}_1, \ldots, \tilde{b}_n) \) can easily be defined recursively in the following fashion: First put \( \tilde{b}_1 = 1 \). Since all vectors in \( \mathcal{B}_n \) have first entry 0, \( \mathcal{B}_n \cup \{\tilde{b}\} \) is linearly independent. Now the other coefficients of \( \tilde{b} \) are defined working down the subscripts according to (12) (where the ultimate equation disappears). Set \( \tilde{b}_n = 1 \) if \( n \) is squarefree, and 0 otherwise. If the coefficients \( \tilde{b}_n, \tilde{b}_{n-1}, \ldots, \tilde{b}_{k+1} \) with \( k \geq 2 \) have been chosen, let

\[
\tilde{b}_k := 1 - \sum_{2 \leq j \leq k} \tilde{b}_{j-k}.
\]

Obviously, \( \tilde{b} \) satisfies (12) and hence lies in \( \ker LCG_n \) by Proposition 4.

Apparently, for sufficiently large \( n \) the vector \( \tilde{b} \) is not simple, hence \( \mathcal{B}_n \cup \tilde{b}_n \) is not a simple basis of \( \ker LCG_n \). In fact, we have

**Theorem 8** For any integer \( n \) satisfying \( M(n) = 1 \), \( \ker LCG_n \) does not have a simple basis.

**Proof.** In 1952, Nagura [14] gave a rather short proof for the fact that, given \( x \geq 25 \), there is always a prime \( p \) in the interval \( x < p \leq \frac{6}{5}x \). Setting \( x = \frac{n}{6} \) we obtain that for every integer \( n \geq 150 \) there is a prime \( p \) satisfying

\[
\frac{n}{6} < p \leq \frac{n}{5}.
\]  

The primes 17 and 29, respectively, show that (13) is also valid if \( 94 \leq n \leq 100 \) or \( 146 \leq n \leq 149 \). By Remark 6(ii) we thus can find a prime \( p \) in the interval (13) for each integer \( n > 1 \) satisfying \( M(n) = 1 \). Alternatively, this follows from the more complicated estimates given later by Rosser and Schoenfeld [19].

By Proposition 4 the basis vectors of \( \ker LCG_n \) are described by (12). Since \( M(n) = 1 \), the last equation of (12) disappears. Hence there is at least one basis vector \( b \), say, with \( b_1 \neq 0 \). We shall prove that \( b \) cannot be simple.

From the equations \( b_{3p} - b_1 = 0 \) and \( b_{5p} - b_1 = 0 \) of (12), we get \( b_{3p} = b_{5p} = b_1 \). The equation \( b_{2p} + b_{4p} - b_1 = 0 \) implies \( b_{2p} + b_{4p} = b_1 \). By inserting these into the equation \( b_p + b_{2p} + b_{3p} + b_{4p} + b_{5p} - b_1 = 0 \), we finally obtain \( b_p = -2b_1 \). So \( b \) has the entries \( b_1 \neq 0 \) and \( b_p = -2b_1 \), thus it is not simple.

\[ \Box \]

5. The Traditional Coprime Graph

Let us finally consider the traditional coprime graph \( TCG_n \) having a loop at the vertex 1, i.e., its adjacency matrix \( \tilde{A}_n = (\tilde{a}_{i,j})_{n \times n} \) is defined as

\[
\tilde{a}_{i,j} = \begin{cases} 
0 & \text{if } \gcd(i,j) > 1, \\
1 & \text{otherwise.}
\end{cases}
\]
Then the analogue of Proposition 4 reads

**Proposition 9** Let $n > 1$ be an arbitrary integer. A vector $b = (b_1, \ldots, b_n) \in \mathbb{R}^n$ lies in $\text{Ker}\tilde{\mathcal{A}}_n$ if and only if

$$b_1 = 0$$

and for $2 \leq k \leq n$, $\mu(k) \neq 0$

$$\sum_{j \equiv 0 \mod k}^{n} b_j = 0.$$ 

The proof of Proposition 9 as well as those of the subsequent main results are easily obtained by adjusting the proofs of Proposition 4 and Theorems 5 and 7 accordingly.

**Theorem 10** For any integer $n > 1$ we have $\dim_{\mathbb{R}} \text{Ker}T\mathcal{C}G_n = \nu(n)$, and

$$\dim_{\mathbb{R}} \text{Ker}T\mathcal{C}G_n = (1 - \frac{6}{n^2})n + O(\sqrt{n}).$$

**Theorem 11** For each positive integer $n$, we have

(i) $\mathcal{B}_n$ is a basis of $\text{Ker}T\mathcal{C}G_n$.

(ii) $\text{Ker}T\mathcal{C}G_n$ has a simple basis, i.e., the components of all basis vectors are 0, 1 or $-1$.

(iii) Let $\iota : \mathbb{R}^n \to \mathbb{R}^{n+1}$ with $\iota(b_1, \ldots, b_n) := (b_1, \ldots, b_n, 0)$ be the canonical injection. Then we have $\iota(\text{Ker}T\mathcal{C}G_n) \subseteq \text{Ker}T\mathcal{C}G_{n+1}$.

**Acknowledgement** The authors are grateful to the referee for his/her careful revision of our paper, including a short proof of Theorem 8.

**References**


