CODES ON GRAPH

Kiki Ariyanti Sugeng
CIMPA Research School on **Group Actions in Arithmetic and Geometry**
February 27, 2020

Outline

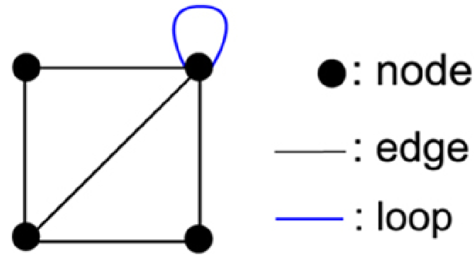
- Introduction on graph theory
- Cycle code and graph code of a graph

INTRODUCTION ON GRAPH THEORY

Graph

- A **graph** Γ is a pair (V, E) where V is a nonempty set and E is a set disjoint from V. The element of V are called vertices/nodes, and members of E are called edges.

- Edges are **incident** to one or two vertices which are called the ends of the edge.
- If an edge is incident with exactly one vertex, then its is called **loop**,



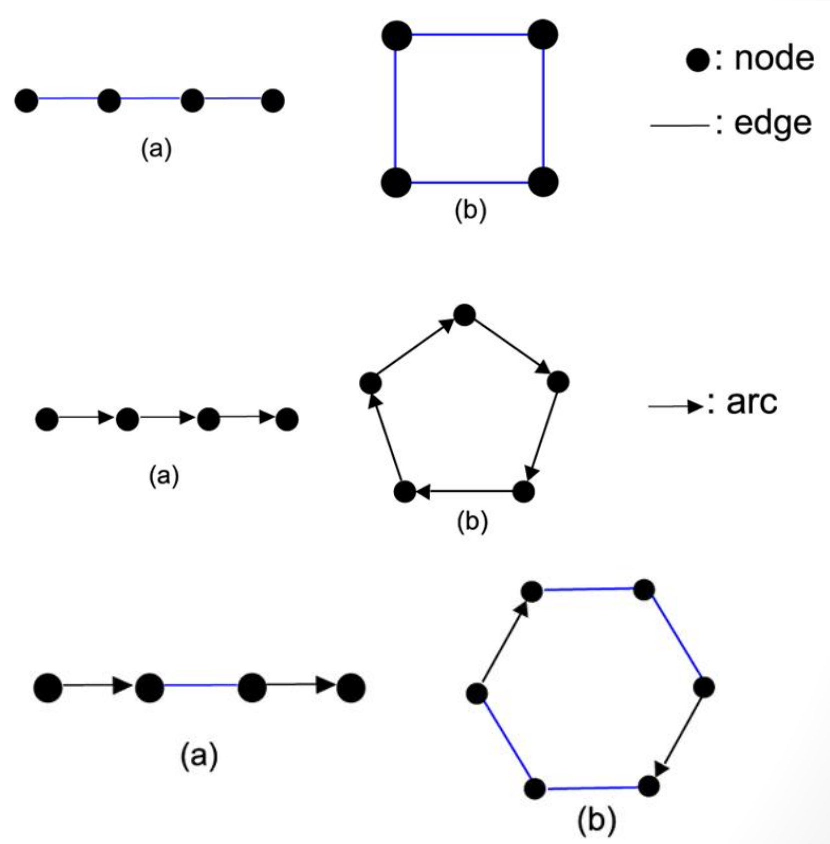
$|V|$ = 4 and $|E|$ = 6

Adjacent Vertices

- If u and v are vertices that are incident with an edge, then they are called **neighbors** or $\subseteq$.
- Two edges are called **parallel** if they are incident with the same vertices.
- The graph is called **simple** if it has no loops and no parallel edges

Graph: Path and Cycle

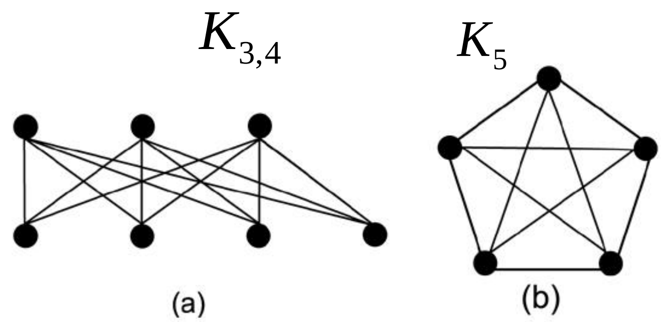
- Undirected:
- Directed:
- Mixed:



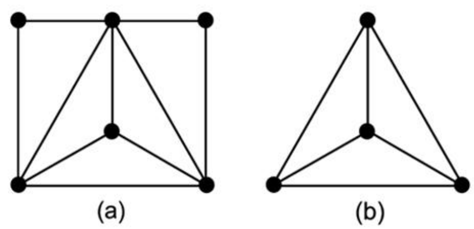
Graph : Complete Graph

(a) Complete bipartite graph

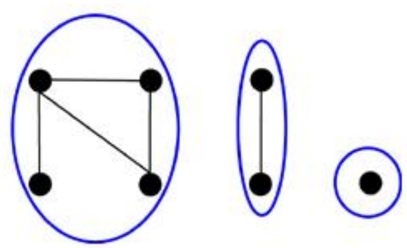
(b) Complete graph



Clique graph :



Connected Components:



Subgraph

Let $\Gamma = (V, E)$ be a graph. Suppose that $\mathrm{V}' \subseteq V$and E’ $\subseteq$ E and all the endpoints of e’ in E’ are in V’. Then $\Gamma^{'} = (V', E')$ is a graph and it is called a **subgraph** of Γ.

Two vertices are connected

- Two vertices u and v are **connected** by a path from u to v if there is a t-tuple of mutually distinct vertices $(v_1, v_2, \dots, v_t)$ with $u = v_1$ and $v = v_t$, and $(t-1)$-tuple of mutually distinct edges $(e_1, e_2, \dots, e_{t-1})$ such that e_i is incident with v_i and v_{i+1} for all $1 \leq i < t$.
- If moreover e_t is an edge that is incident with u and v and distinct from e_i for all $i < t$, then $(e_1, e_2, \dots, e_{t-1}, e_t)$ is called a cycle. The length of the smallest cycles is called the **girth** of the graph and is denoted by $\gamma(\Gamma)$

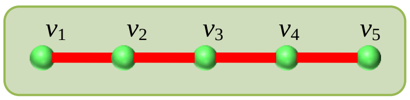
Connected Graph

- The graph is called **connected** if every two vertices are connected by a path.
- A maximal connected subgraph of Γ is called a connected component of Γ.
- If Γ is not connected, then the vertex set V of Γ is a disjoint union of subset V_i and the set of edge is disjoint union of subset E_i such that $\Gamma_i = (V_i, E_i)$ is a connected component of Γ. The number of connected component of Γ is denoted by $c(\Gamma)$

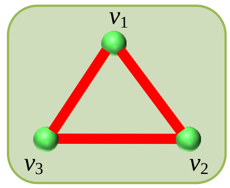
Adjacency Matrix

The Adjacency Matrix of a graph Γ, denoted $A(\Gamma)$, is an $n \times n$ matrix that for each (u, v) contains the number of edges in G between vertex u and vertex v.

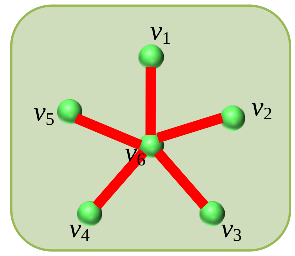
Adjacency matrix : Examples



$$\begin{array}{c c} & \begin{array}{ccccc} v_1 & v_2 & v_3 & v_4 & v_5 \end{array} \\ \begin{array}{c} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{array} & \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \end{array}$$



$$\begin{array}{c c} & \begin{array}{ccc} v_1 & v_2 & v_3 \end{array} \\ \begin{array}{c} v_1 \\ v_2 \\ v_3 \end{array} & \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \end{array}$$



$$\begin{array}{c c} & \begin{array}{cccccc} v_1 & v_2 & v_3 & v_4 & v_5 & v_6 \end{array} \\ \begin{array}{c} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{array} & \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix} \end{array}$$

Incidence Matrix

- Let $\Gamma = (V, E)$ be a finite graph. Suppose that V consists of m elements enumerated by $v_1, v_2, \dots, v_m$. Supposet that E consists of m elements enumerated by $e_1, \dots, e_n$. The **incidence matrix** $I(\Gamma)$ is an $m \times\ n$ matrix with entries a_{ij} defined by

- $a_{ij} = \begin{cases} 1, & \text{if } e_j \text{ is incident with } v_i \text{ and } v_k \text{ for some } i < k \\ -1, & \text{if } e_j \text{ is incident with } v_i \text{ and } v_k \text{ for some } i > k \\ 0, \text{otherwise} \end{cases}$

CYCLE CODE AND GRAPH CODE OF A GRAPH

Graph Code

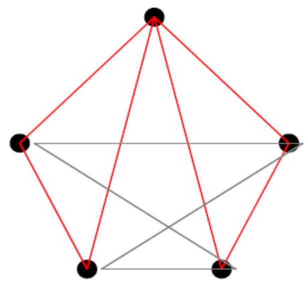
- The **graph code** C_Γ of Γ over F_q is the F_q linear code that is generated by the rows of the incidence matrix $I(\Gamma)$.
- The **cycle code** of Γ is the dual of the graph code of Γ.

Notes

- Cycle code is also referred to as a graphic code, and its dual as a cographic code

Cycle codes of graphs

- $\Gamma = (V, E)$ undirected connected graph with no loops, no multiple edges.
- $E = \{e_1, e_2, \ldots, e_m\}$
- Subgraph $H \subset \Gamma \leftrightarrow$ Characteristic vectors in $\{0,1\}^m$: $h_i = 1_H(e_i)$.
- Cycle $c \in \{0,1\}^m$: Subgraph with all vertices incident with an even number of edges.
- Cycle space of Γ : the vector space over F_2 of all cycles. It has
- dimension $m - n + 1$ (cyclomotic number)



$c = (1, 1, 1, 1, 0, 1, 1, 0, 0, 0)$

Cycle Code of Graphs

Cycle code of Γ : the linear binary code $[n, k, d]$ defined by the cycle space of Γ with

(i) length m (number of edges)

(ii) dimension $k = m - n + 1$ (cyclotomic number)

(iii) minimum distance $d = girth\ of \Gamma$(length of smallest cycle).

(iv) incidence matrix of $\Gamma \longleftrightarrow$ parity-check matrix of the code

(low–density parity-check code)

- Binary Code of length $N = ab$ → Array $a \times b$

$$(100010001) \longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- Code on $GF(2^b)$ of length $N' = a$ → Code on GF(2)

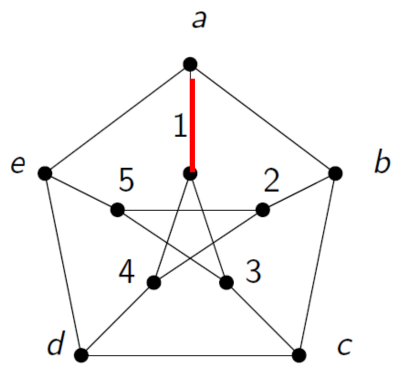
$$(1, x, x^2) \longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- Distance between codewords → Distance between columns: correction of column errors or column erasures
- Array codes are used to address bursts of errors (as opposite to random errors).
- The are implemented in the standards of CD technology (by using Reed–Solomon codes).

Array Cycle Code

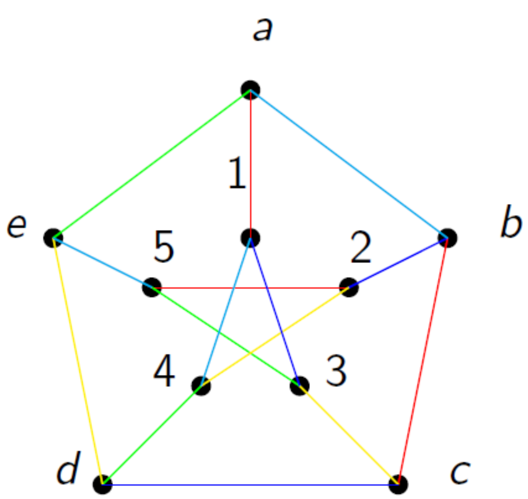
- Graph $G = (V, E)$ with $m = ab$ edges
- Partition the edges in columns

$$\begin{pmatrix} a1 & b2 & c3 & d4 & e5 \\ bc & cd & de & ea & ab \\ 52 & 13 & 24 & 35 & 41 \end{pmatrix}$$



- Partition the edges in columns $\longrightarrow$ edge-coloring of the graph.

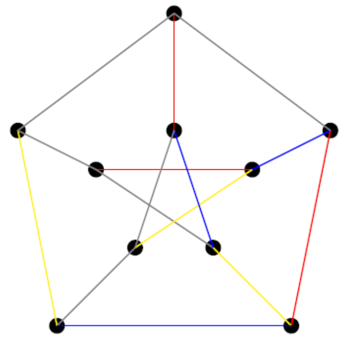
$$\begin{pmatrix} a1 & b2 & c3 & d4 & e5 \\ bc & cd & de & ea & ab \\ 52 & 13 & 24 & 35 & 41 \end{pmatrix}$$



- Array Cycle code: The graph cycle code turned into an array code.

Minimum Distance

- The minimum distance of the array cycle code is the minimum number of colors in a cycle.
- In the example, the code has $|C| = 215 - 10 + 1$ and $D = 4 = 5 - \log 8\,|C| + 1$.
- It is an MDS (Maximum Distance Separating) code.
- This means that every three colors span a spanning tree.

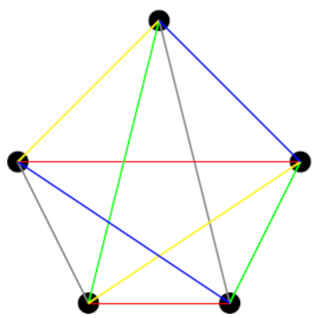


Proposition

- Let Γ be a finite graph. Then the cycle code of Γ is a code with parameters $[n, k, d]$, where $n = |E|, k = \ |E| - |V| + c(\Gamma)$, and $d = \gamma(\Gamma)$. These parameters are independent of the choice of the field F_q

B–codes

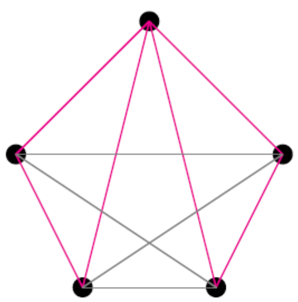
- MDS Array cycle codes with $D = 3$.
  Edge colored graph such that every two colors make a spanning tree
- Largest length $\longleftrightarrow$ maximum number of edges.
  Complete graphs
- Every two colors make an acyclic graph $\longleftrightarrow$ every color is a matching (K_n has triangles).
- They provide MDS array cycle codes with $a = (n-1)/2$, $b = n$, $\mathbb{F}_2$–dimension $n - 2a$ and $D = 3$.
  Known in the literature as B–codes.

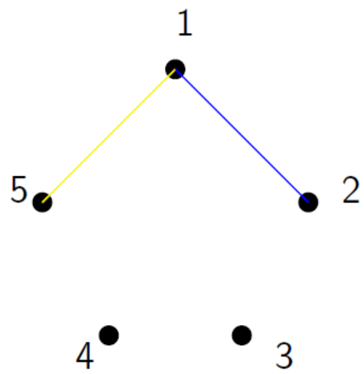


Correction algorithms: Column erasure correction

The sent codeword is $\begin{pmatrix} 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}$

- Two columns have been erased.
  $\begin{pmatrix} 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}$
- Parity check of endvertex 3: edge e_9 is not in the word.
- Parity check of endvertex 4: edge e_8 is not in the word.
- Parity check of endvertex 2: edge e_2 is in the word.
- Parity check of endvertex 5: edge e_1 is in the word.
- The algorithm is linear in n.

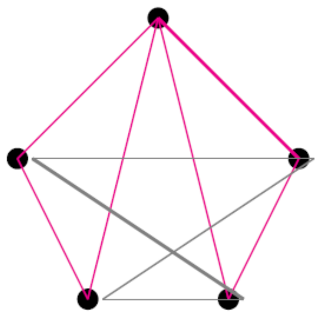


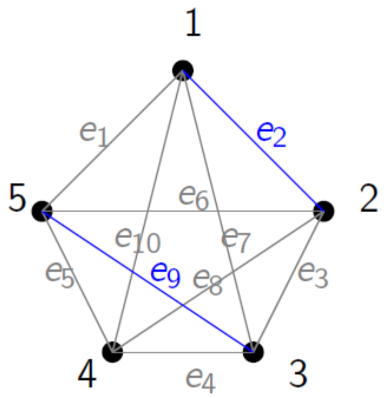


Correction algorithms: Errors in one column

The received codeword is $\begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix}$

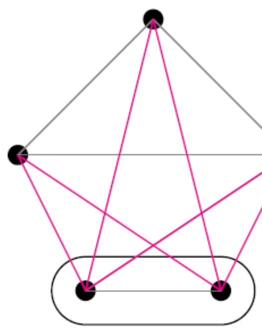
- All errors are located in a single column.
- Find the vertices with unsatisfied parity check vertices.
- Test the color which covers the selected vertices.
- Exchange the bit of the edges covering these selected vertices.
- The algorithm is linear in n.





The dual B–codes

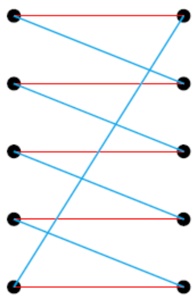
- The dual of a MDS code is again MDS.
- The dual of a B-code is the array cocycle code of K_n. Codewords are sets of edges joining a set with its complement.
- The minimum distance is $D = n - 1$ (the number of matchings).

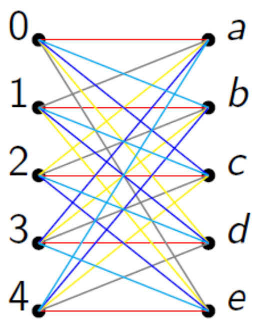


More examples of MDS array cycle codes

For a graph G the above correction algorithms work if G admits an edge–coloring such that every $D-1$ colors make a spanning tree.

- The Petersen graph provides an MDS array cycle code with $D = 4$.
  Unfortunately $D = 4$ implies $n \leq 10$.
- For $D = 3$ the complete bipartite graphs $K_{n-1,n}$ are conjectured to admit such an edge–coloring.
  Equivalently $K_{n,n}$ admits a coloring such that two colors span a Hamiltonian cycle.
  Known to be the case for $n = p$, $n = 2p-1$, $n = p^2$, and small values of n.





$$\begin{bmatrix} 0a & 2d & 4b & 1e & 3c \\ 3d & 0b & 2e & 4c & 1a \\ 1b & 3e & 0c & 2a & 4d \\ 4e & 1c & 3a & 0d & 2b \\ 2c & 4a & 1d & 3b & 0e \end{bmatrix}$$

Reference

- R. Pellikaan, X-W Wu, S. Bulygin and R. Jurrius, Codes, cryptography and curves with computer algebra, Cambridge University Press, 2018
- O. Serra, An application to coding theory and cryptography, presented at CIMPA-Indonesia School, February 2 – 13, 2009

THANK YOU