

---

# BAGCHI'S CONJECTURE

---

Amit Rajaraman

Last updated June 18, 2022

## Contents

<b>1</b>	<b>Definitions</b>	<b>1</b>
<b>2</b>	<b>Inamdar's Result</b>	<b>2</b>
2.1	Combinatorial Arguments . . . . .	2
2.2	Linear Algebra . . . . .	3
2.3	Graphs . . . . .	5

## §1. Definitions

**Definition 1.1** (Incidence System). An *incidence system* is a pair  $(\mathcal{P}, \mathcal{L})$ , where  $\mathcal{P}$  is a set and  $\mathcal{L}$  is a set of subsets of  $\mathcal{P}$ . Elements of  $\mathcal{P}$  are called *points* and elements of  $\mathcal{L}$  are called *lines*. A line  $\ell$  is said to be *incident* on a point  $p$  if  $p \in \ell$ .

**Definition 1.2** (Partial Linear Space). An incidence system  $(\mathcal{P}, \mathcal{L})$  is said to be a *partial linear space* if

1. for each  $\ell \in \mathcal{L}$ ,  $|\ell| \geq 2$ .
2. for distinct  $x, y \in \mathcal{P}$ , there is at most one  $\ell \in \mathcal{L}$  such that  $\{x, y\} \subseteq \ell$ .

An incidence system is said to be a *linear space* if in the second condition above, the “at most one” is replaced with “exactly one”.

**Definition 1.3.** Given an incidence system  $\mathcal{X} = (\mathcal{P}, \mathcal{L})$  and a field  $\mathbb{F}$ , we define the linear code  $\mathcal{C}_{\mathbb{F}}(\mathcal{X})$  over  $\mathbb{F}^{\mathcal{P}}$  as follows. Identify each line  $\ell$  with the codeword in  $\mathbb{F}^{\mathcal{P}}$  whose  $x$ th coordinate is 1 if  $x \in \ell$  and 0 otherwise.  $\mathcal{C}_{\mathbb{F}}(\mathcal{X})$  is then the space spanned by the codewords corresponding to the lines in  $\mathcal{L}$ . If  $\mathbb{F} = \mathbb{F}_q$ , we sometimes denote the above as  $\mathcal{C}_q(\mathcal{X})$ .

We also often denote this as  $\mathcal{C}_q(\mathcal{L})$  if the point set is clear from context. The incidence system  $\mathcal{X}$  is said to be *trivial* at  $q$  if  $\mathcal{C}_q(\mathcal{X})$  is all of  $\mathbb{F}_q^{\mathcal{P}}$ .

**Definition 1.4** (Join). Given two partial linear spaces  $(\mathcal{P}_1, \mathcal{L}_1)$  and  $(\mathcal{P}_2, \mathcal{L}_2)$  with  $\mathcal{P}_1 \cap \mathcal{P}_2 = \emptyset$ , one can define the *join* of the two partial linear spaces by  $(\mathcal{P}_1 \cup \mathcal{P}_2, \mathcal{L}_1 \cup \mathcal{L}_2 \cup \mathcal{L}_3)$ , where

$$\mathcal{L}_3 = \{\{x, y\} : x \in \mathcal{P}_1, y \in \mathcal{P}_2\}.$$

## §2. Inamdar's Result

**Theorem 2.1.** If a PLS  $\mathcal{X} = (\mathcal{P}, \mathcal{L})$  is non-trivial at  $p$  and has at least  $p+1$  lines through every point, then  $|\mathcal{P}| \geq 2p$ . Moreover, equality holds iff  $\mathcal{X}$  is the join of two  $p$ -lines.

### 2.1. Combinatorial Arguments

For the rest of this section, assume that  $\mathcal{X} = (\mathcal{P}, \mathcal{L})$  is a partial linear space of the above prescribed format with  $\mathcal{P} \leq 2p$ . We wish to show that  $\mathcal{P} = 2p$ . Let  $\mathcal{C} = \mathcal{C}_p(\mathcal{X})$ .

It may be shown that it can be assumed that

1. Each point is incident on exactly  $p+1$  lines (throw away extra lines).
2.  $\mathcal{C}^\perp$  is one-dimensional. Suppose that it is equal to  $\langle w \rangle$  (restrict to the support of the minimum support word in  $\mathcal{C}^\perp$ ).
3. For any line  $\ell$  and any  $\ell' \subsetneq \ell$ ,  $\langle w, \ell' \rangle \neq 0$  (do the splitting procedure).

Also consider the colouring of  $\mathcal{P}$  wherein each point  $P$  is coloured  $w(P)$ .

**Proposition 2.2.** If  $\mathcal{X}$  has a  $p$ -line, it is equal to a join of two  $p$ -lines.

*Proof.* Let  $P_1 \cdots P_p$  be a  $p$ -line. Since each point has  $p$  lines remaining, there must be at least  $p$  points other than the  $P_i$ , say  $(Q_i)_{i=1}^p$ . Since  $|\mathcal{P}| \leq 2p$ , these constitute all the points. Further, since each  $P_i$  has  $p$  lines to the  $Q_j$ , there must be a 2-line  $P_i Q_j$  for each  $1 \leq i, j \leq p$ . Now, suppose that  $w(P_1) = 1$ . Because  $w(P_1) + w(Q_j) = 0$  for all  $j$  ( $P_1 Q_j$  forms a  $p$ -line),  $w(Q_j) = -1$  for all  $j$ . Each of the  $Q_i$  now has one line not accounted for. This line must be contained within the  $(Q_i)$ . However, due to all of them having the same colour, the size of any such line must be  $p$ , completing the proof.  $\blacksquare$

**Proposition 2.3.** If  $\mathcal{X}$  has no  $p$ -line, the largest line in  $\mathcal{X}$  is of size at most  $2p/3$ .

*Proof.* Let  $\ell$  be a line with  $|\ell| > 2p/3$ . Given a  $P \in \ell$ , let  $x_P$  be the number of 2-lines through  $P$ . We then have that

$$2p - |\ell| \geq \underbrace{x_P}_{\text{points in 2-lines}} + \underbrace{2(p - x_P)}_{\text{points in } \geq 3\text{-lines}},$$

so  $x_P \geq \ell$ . Observe that if  $x_P > (2p - |\ell|)/2$  for all  $P \in \ell$ , it follows by a pigeonhole argument that any two points  $P, Q$  in  $\ell$  have a “2-neighbour” (a point  $u$  such that  $uP$  and  $uQ$  are 2-lines) in common. This is indeed the case because  $x_P \geq |\ell| > (2p/3) > (2p - |\ell|)/2$ . This in turn implies that  $w(P) = w(Q)$ , because  $w(P) = -w(u) = w(Q)$ . Therefore,  $\ell$  is monochromatic, so for any fixed  $P \in \ell$ ,

$$0 = \sum_{P \in \ell} w(P) = |\ell|w(P).$$

As  $w(P) \neq 0$ ,  $|\ell| = p$ , yielding a contradiction.  $\blacksquare$

As a result, we are done if we manage to show that  $\mathcal{X}$  has a line of size greater than  $2p/3$ .

Now, let us restrict ourselves to the case where  $\mathcal{X}$  is a linear space with exactly  $p+1$  lines through each point.

## 2.2. Linear Algebra

Let  $C$  be the  $m \times n$  incidence matrix of  $\mathcal{X}$ , where  $m = |\mathcal{L}|$  and  $n = |\mathcal{P}|$ . Observe that  $(C^t C)_{ij} = pI + J$ , where  $J$  is the all 1s matrix. As a result, its characteristic equation is just  $(X - (p + n))(X - p)^{n-1}$ . Recall that the characteristic matrix of  $CC^t$  and  $C^t C$  only differ by a power of  $X$ . That is, the characteristic equation of  $CC^t$  is  $X^{m-n}(X - (p + n))(X - p)^{n-1}$ . Over  $\mathbb{F}_p$ , this becomes  $X^{m-1}(X - n)$ .

Suppose that  $y$  is an eigenvector with eigenvalue 0. That is,  $CC^t y = 0$ . A question to ask is: is  $C^t y$  always 0? If it is not, we can explicitly get a vector in  $\mathcal{C}^\perp$ .

**Proposition 2.4.** There exists  $y$  in the kernel of  $CC^t$  such that  $C^t y \neq 0$  (in  $\mathbb{F}_p$ ).

*Proof.* Indeed, note that the kernel of  $CC^t$  is of dimension  $m - 1$ , whereas the kernel of  $C^t$  is of dimension at most  $m - n$ . It follows that the kernel of  $C^t$  is a proper subset of  $CC^t$ . ■

**Definition 2.1** (Irreducible Matrix). A matrix  $A = (a_{ij})_{n \times n}$  is said to be *irreducible* if the (directed) graph  $G_A$  with vertex set  $[n]$  and edge from  $i$  to  $j$  iff  $a_{ij} \neq 0$  is strongly connected.

Observe that  $CC^t$  is irreducible! Indeed, given any lines  $\ell, \ell'$ , there is a line  $\ell''$  through some (any)  $x \in \ell$  and  $y \in \ell'$ .

**Theorem 2.5** (Perron-Frobenius Theorem). Let  $A = (a_{ij})_{n \times n}$  be a non-negative irreducible matrix. Then,

1. There exists a positive eigenvalue  $\lambda$  such that any other eigenvalue is of absolute value at most  $\lambda$ . Further, there is a number  $h$  called the *period* such that  $\omega_h^k \lambda$  is an eigenvalue for every  $0 \leq k < h$ . These are the only eigenvalues with absolute value  $\lambda$ .
2.  $\lambda$  is a simple root of the characteristic polynomial. As a result, its corresponding eigenspace is one-dimensional.
3.  $\lambda$  has an eigenvector  $v$  with all positive entries. Any eigenvector with all positive entries is in the span of  $v$ .

Since  $CC^t$  is irreducible and non-negative with maximum eigenvalue  $n + p$ , there is an eigenvector  $v$  of  $n + p$  with all positive entries.

For example, we now have that

$$\begin{aligned}
 m(n + p) &= (n + p) \mathbf{1}^\top \mathbf{1} \\
 &\geq \mathbf{1}^\top CC^t \mathbf{1} \\
 &= \sum_{\ell, \ell' \in \mathcal{L}} |\ell \cap \ell'| \\
 &= \sum_{\ell \in \mathcal{L}} (p + 1) |\ell| && \text{(each point on } \ell \text{ has } p + 1 \text{ lines through it)} \\
 &= (p + 1) \cdot n(p + 1) = n(p + 1)^2
 \end{aligned}$$

That is,

$$m \geq \frac{n(p + 1)^2}{n + p}.$$

The fact that  $p + 2 < n \leq 2p$  further gives that

$$m \geq \frac{(p + 1)^2}{2}$$

The system has many lines! Can we use this to show that either there is a “large” line somewhere or the code is trivial?

Can we draw stronger conclusions by using another vector in place of  $\mathbf{1}$ ? It seems that the ideal choice is to assign a larger weight to larger lines. What about the vector defined by  $v_\ell = |\ell|$ ?

Recall that

$$\sum_{\ell \in \mathcal{L}} |\ell| = n(p+1).$$

**Proposition 2.6.**

$$\sum_{\ell \in \mathcal{L}} |\ell|^2 = n(n+p).$$

*Proof.* Consider the vector  $v$  with  $v_\ell = |\ell|$ . Then,

$$\begin{aligned} \sum_{\ell \in \mathcal{L}} |\ell|^2 &= v^t v \\ &\geq \frac{1}{n+p} v^t C C^t v \\ &= \frac{1}{n+p} \left( \sum_{\ell \in \mathcal{L}} \underbrace{v_\ell^2 |\ell|}_{|\ell|^3} + \sum_{\substack{\ell, \ell' \in \mathcal{L} \\ \ell \neq \ell' \\ \ell \cap \ell' \neq \emptyset}} v_\ell v_{\ell'} \right). \end{aligned}$$

Let us compute the second quantity.

$$\begin{aligned} \sum_{\substack{\ell, \ell' \in \mathcal{L} \\ \ell \neq \ell' \\ \ell \cap \ell' \neq \emptyset}} v_\ell v_{\ell'} &= \sum_{P \in \mathcal{P}} \sum_{\substack{\ell, \ell' \ni P \\ \ell \neq \ell'}} |\ell| |\ell'| \\ &= \sum_{P \in \mathcal{P}} \left( \sum_{\ell, \ell' \ni P} |\ell| |\ell'| - \sum_{\ell \ni P} |\ell|^2 \right) \\ &= \sum_{P \in \mathcal{P}} \left( \left( \sum_{\ell \ni P} |\ell| \right)^2 - \sum_{\ell \ni P} |\ell|^2 \right). \end{aligned}$$

The second summation is just

$$\sum_{P \in \mathcal{P}} \sum_{\ell \ni P} |\ell|^2 = \sum_{\ell \in \mathcal{L}} \sum_{P \in \ell} |\ell|^2 = \sum_{\ell \in \mathcal{L}} |\ell|^3.$$

The  $\sum |\ell|^3$  terms cancel out! As a result,

$$\sum_{\ell \in \mathcal{L}} |\ell|^2 \geq \frac{1}{n+p} \sum_{P \in \mathcal{P}} \left( \sum_{\ell \ni P} |\ell| \right)^2.$$

On one hand, linearity and regularity imply that for any  $P$ ,  $\sum_{\ell \ni P} |\ell| = (n-1) + (p+1)$ . Indeed, there is precisely one line through  $P$  and any point  $Q \neq P$ , and there are  $p+1$  lines in all (in which  $P$  is present). Therefore,

$$\sum_{\ell \in \mathcal{L}} |\ell|^2 \geq \frac{1}{n+p} \sum_{P \in \mathcal{P}} (n+p)^2 = n(n+p).$$

Using the Cauchy-Schwarz inequality on the other hand,

$$\begin{aligned}
\sum_{\ell \in \mathcal{L}} |\ell|^2 &\geq \frac{1}{n+p} \sum_{P \in \mathcal{P}} \left( \sum_{\ell \ni P} |\ell| \right)^2 \\
&\geq \frac{1}{n(n+p)} \left( \sum_{P \in \mathcal{P}} \sum_{\ell \ni P} |\ell| \right)^2 \\
&= \frac{1}{n(n+p)} \left( \sum_{\ell \in \mathcal{L}} |\ell|^2 \right)^2 \\
n(n+p) &\geq \sum_{\ell \in \mathcal{L}} |\ell|^2,
\end{aligned}$$

completing the proof. ■

In fact, one similarly gets that for any  $m$ ,

$$\sum_{\ell \in \mathcal{L}} |\ell|^{2m} \geq \frac{1}{n+p} \sum_{P \in \mathcal{P}} \left( \sum_{\ell \ni P} |\ell|^m \right)^2.$$

The right is easily computable for  $m = 0$  (because of regularity) and  $m = 1$  (because of linearity and regularity). Using the Cauchy-Schwarz inequality as we did in the previous proof, one further obtains that

$$\sum_{\ell \in \mathcal{L}} |\ell|^{2m} \geq \frac{1}{n(n+p)} \left( \sum_{\ell \in \mathcal{L}} |\ell|^{m+1} \right)^2.$$

**Porism 2.7.** The eigenvector (up to scalar multiplication) of  $CC^t$  corresponding to the eigenvalue  $n+p$  is the vector  $v_{\mathcal{L}} : \mathcal{L} \rightarrow \mathbb{R}$  with  $v_{\mathcal{L}}(\ell) = |\ell|$ .

Thus, for any vector  $x \in v_{\mathcal{L}}^{\perp}$ ,

$$\left\| C^{\top} x \right\|_2^2 \leq p \|x\|_2^2.$$

### 2.3. Graphs

Consider the graph  $G$  with vertex set  $\mathcal{L}$ , and  $u \leftrightarrow v$  iff  $u \cap v \neq \emptyset$ . What do our conditions on the linear space correspond to?

1. Each point has exactly  $p+1$  lines through it: any maximal clique in  $G$  is of size  $p+1$ .
2. There is a line passing through any two points (linearity): any two maximal cliques have a single point in common.

We wish to count the number of (maximal) cliques in the graph.

Note that each vertex  $\ell$  has degree  $|\ell|$ . As a result, the graph has  $\frac{1}{2}n(p+1)$  edges. Consider the random variable  $X$  that chooses a vertex with probability proportional to its degree (the probability of choosing  $\ell$  is  $|\ell|/n(p+1)$ ). Note that this corresponds to the stationary distribution of a standard random walk on the graph!

Further, we have that

$$\mathbb{E}d(v) = \frac{p+n}{p+1}.$$

Let  $D$  be the  $m \times m$  diagonal matrix with  $D_{ii} = |\ell_i|$ . Let  $A_G$  and  $L_G$  be the adjacency matrix and ??? of  $G$  respectively. Then, observe that

$$\begin{aligned}
L_G &= pD - A_G \text{ and} \\
CC^t &= D + A_G.
\end{aligned}$$