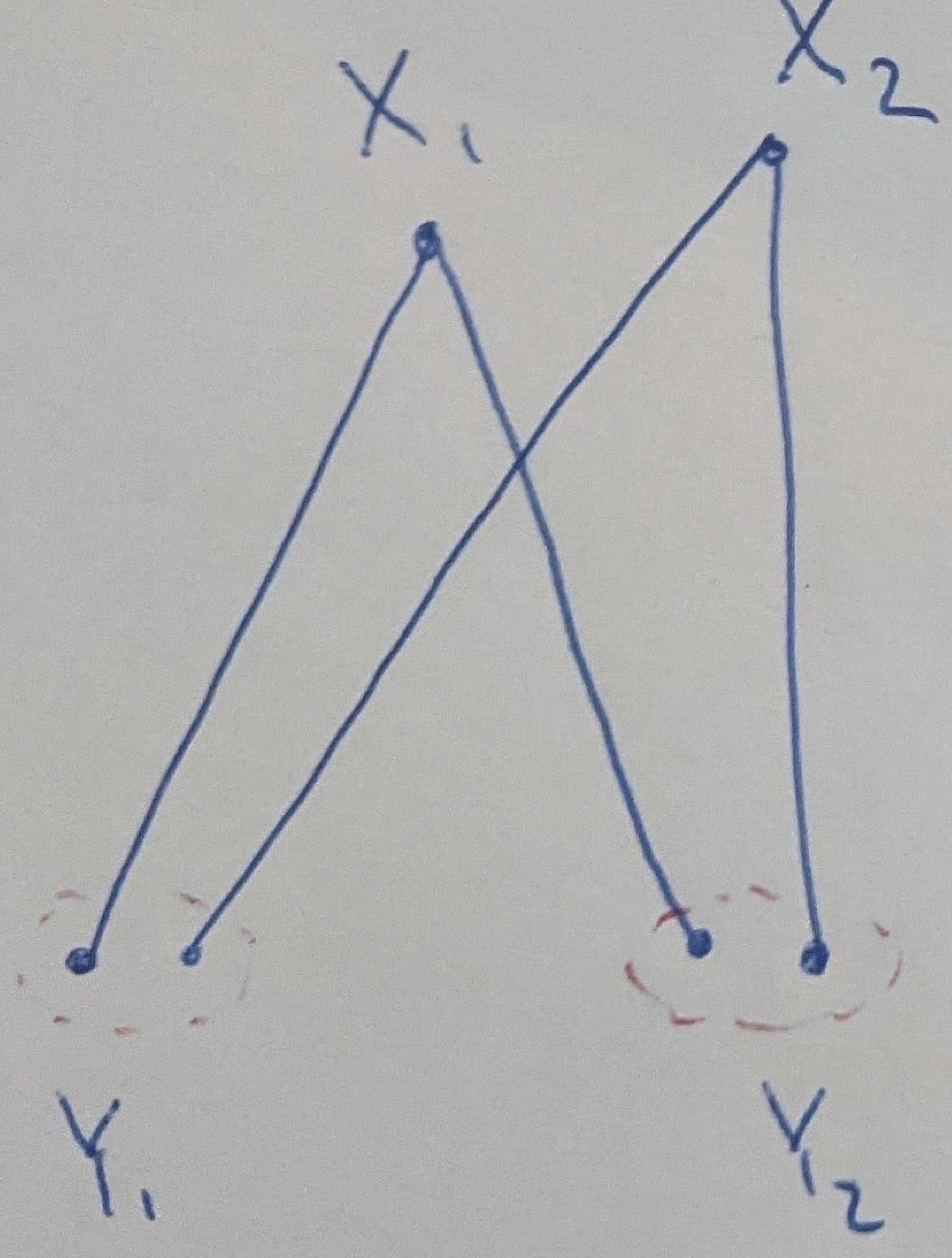
Lecture 18: Sidorenko & Shearer



$$H(X_1, X_2 \mid Y_1, Y_2) = H(X_1 \mid Y_1, Y_2) + H(X_2 \mid Y_1, Y_2).$$

Properties of entropy:

$$H(X) \leq \log_2 |\mathrm{support}(X)|$$

$$H(X) \leq H(X,Y) \leq H(X) + H(Y)$$

equality when X & Y are independent

$$H(X \mid Y, Z) \leq H(X \mid Z)$$

Homomorphism densities:

A homomorphism from H to G is a map $\varphi : V(H) \to V(G)$ which maps edges to edges. (not necessarily injective)

$t(H,G)$ = Probability that a random map $V(H) \to V(G)$ is a homomorphism.

Sidorenko's conjecture:

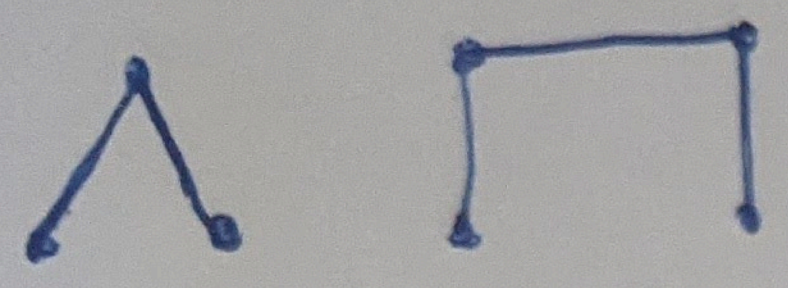
$$t(H,G) \geq \left(t(K_2, G)\right)^{|E(H)|}$$

for any bipartite H.

(if G is a random graph with density p then $t(K_2, G) \sim p$ $\quad t(H,G) \sim p^{|E(H)|}$)

Theorem 9.10: (Blakey & Roy, 1965) Sidorenko's conjecture holds when H is a tree.

Proof:



$H = P_3$ easy exercise

$H = P_4$ $\downarrow$ quite non-trivial.

Let T be a tree with n vertices.
Let $\mathrm{Hom}(T,G)$ denote the set of all homomorphisms T to G.

$$t(T,G) = \frac{|\mathrm{Hom}(T,G)|}{|V(G)|^n} \qquad (t(K_2,G))^{n-1} = \left(\frac{2|E(G)|}{|V(G)|^2}\right)^{n-1}$$

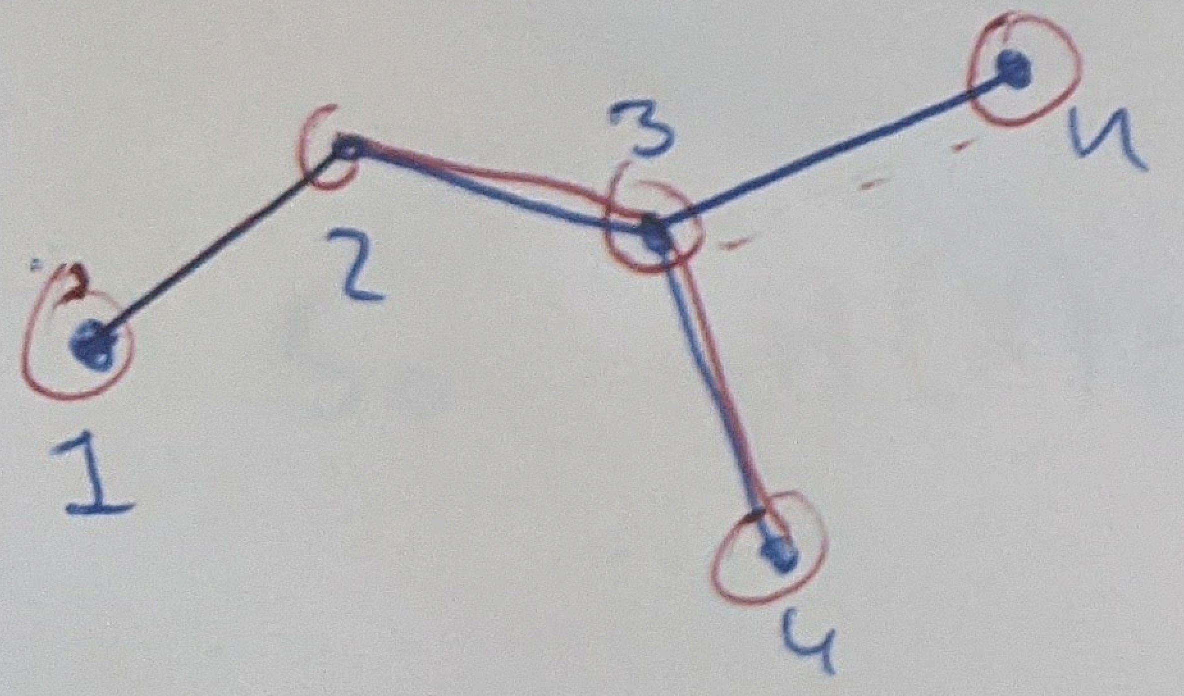
Sidorenko's:

$$|\mathrm{Hom}(T,G)| \geq (2|E(G)|)^{n-1} / |V(G)|^{n-2}$$

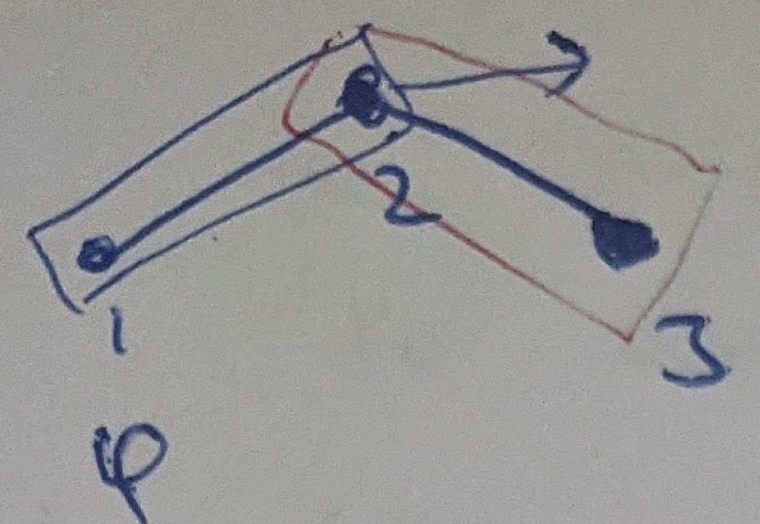
To prove this we will construct distribution X on $\mathrm{Hom}(T,G)$ s.t.

$$\log_2|\mathrm{Hom}(T,G)| \geq \boxed{H(X) \geq (\log_2(2|E(G)|))\cdot(n-1) - (\log_2|V(G)|)\cdot(n-2)}$$

Let $V(T) = \{1,2,\dots,n\}$ s.t. $\{1,2,\dots,i\}$ induce a subtree of T with i a leaf for every i.



We construct distribution X by mapping edge 12 to uniformly random edge of G, and then once vertices $1,2,\dots,i-1$ are mapped mapping i to uniformly random neighbor of the vertex of G it should be adjacent to:



$\Pr[\varphi(1)=v] = \boxed{\dfrac{\deg(v)}{2|E(G)|}} = \Pr[\varphi(2)=v]$

$\Pr[\varphi(3)=v] = \,!$

?

$\Pr[\varphi(i)=v] =$

$\Pr[\varphi(e)=f] = \dfrac{1}{|E(G)|}$

for any $e \in E(T)$
$f \in E(G)$

Let $X_1, X_2, \dots, X_n$ denote the distributions of images of $1, \dots, n$.

Then $H(X) = H(X_1, \dots, X_n) =$

$$= H(X_1) + \underbrace{H(X_2 | X_1)} + \underbrace{H(X_3 | X_2, X_1)} + \dots + \underbrace{H(X_n | X_1, \dots, X_{n-1})}.$$

What is $\underbrace{H(X_i | X_1, \dots, X_{i-1})}_{\text{if } j \text{ is the unique neighbor of } i \text{ in } \{1,\dots,i-1\}.} = H(X_i | X_j) = H(X_2|X_1) =$

$$= \underline{H(X_2, X_1) - H(X_1)}.$$

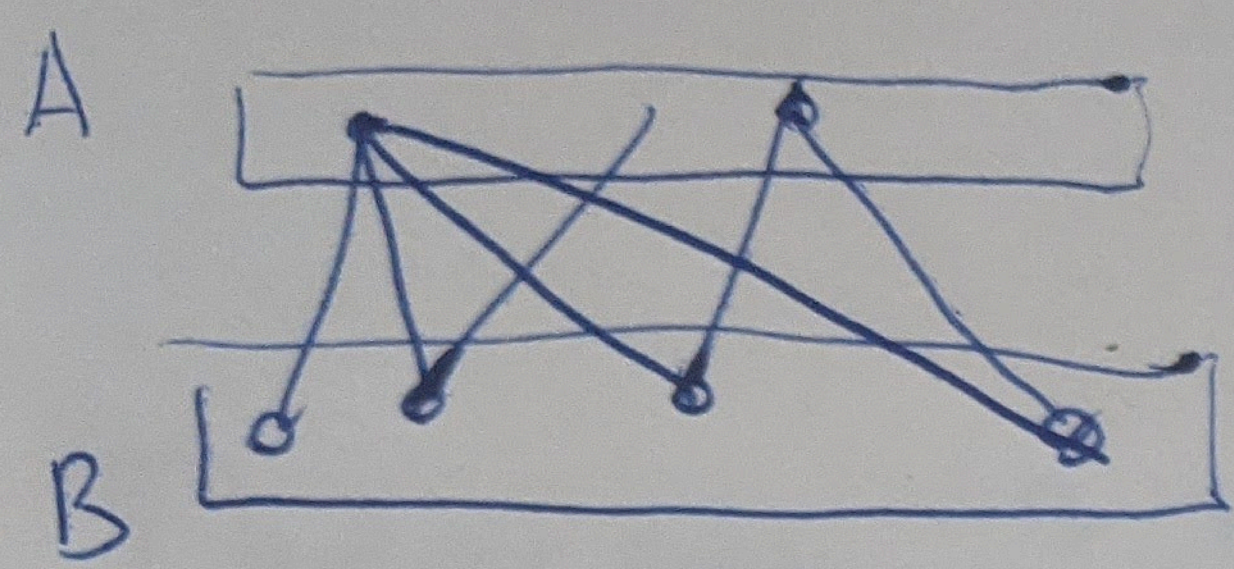
So $H(X) = H(X_1) + (n-1)\left(H(X_2,X_1) - H(X_1)\right)$

$$= (n-1)\underbrace{H(X_2,X_1)}_{\text{uniform over oriented edges of } G} - (n-2)\underbrace{H(X_1)}_{\le}.$$

$$\geq \left(\log_2(2|E(G)|)\right)\cdot(n-1) - (n-2)\log_2|V(G)|. \checkmark$$

Use of entropy for Sidorenko: Li & Szegedy 2011.
presentation is via Yufei Zhao.

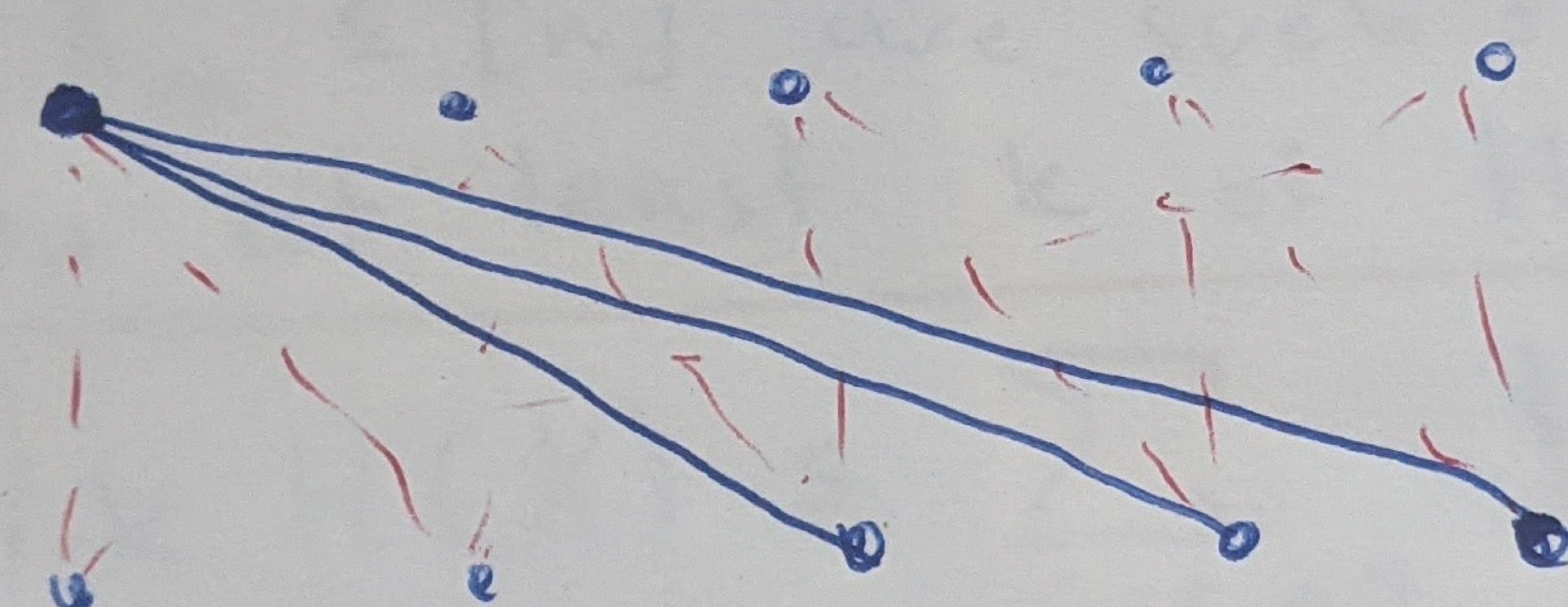
In Zhao's notes you can see adaptation of this proof to complete bipartite graphs & more generally to any graph with bipartition (A, B) s.t. some vertex of A is complete to B.



Hamed Hatami proved the conjecture for 1-skeletons of hypercubes

(can also be done using entropy).

The smallest open case is $\underline{K_{5,5} - \text{edge set of } C_{10}}$

$H =$ 

~~Thex~~

Shearer's lemma

$A_i = \{i\}$ $k=1$

(*) $H(X_1, X_2, \ldots, X_n) \le H(X_1) + H(X_2) + \ldots + H(X_n)$.

$A_i = \{1,2,3\} - \{i\}$
$k = 2$.

$$2H(X_1, X_2, X_3) \overset{!}{\le} H(X_1, X_2) + H(X_2, X_3) + H(X_3, X_1)$$

$$2H(X_1,X_2,X_3) \le 2(H(X_1) + H(X_2) + H(X_3)) \ge H(X_1, X_2) + H(X_2, X_3) + H(X_3, X_1)$$

Theorem 9.11 (Shearer's lemma): Let $X = (X_1, X_2, \ldots, X_n)$. For $A \subseteq [n]$ let $X_A = (X_i)_{i \in A}$.

If $A_1, A_2, \ldots, A_m \subseteq [n]$ are such that every element of $[n]$ belongs to at least k of them, then

$$\boxed{k\,H(X) \le \sum_{i \in [m]} H(X_{A_i})}.$$

Proof: By induction on k.

Base case ($k=1$): Recall that $H(X, Y) \ge H(X)$
so $H(X_A) \ge H(X_B)$
whenever $B \subseteq A$

So reducing sizes of A_i we may assume that A_i are disjoint, in which case we get (*) above.

Induction step: If $A_i = [n]$ for some i then inequality follows from induction hypothesis by cancelling $H(X) = H(X_{A_i})$.

We will reduce the proof to this case.

Claim: $H(Y,Z,W) + H(Z) \leq H(Y,Z) + H(Z,W)$ for any r.v. Y, Z, W.

$$\Big\downarrow$$

$$\underbrace{H(Y,Z,W) - H(Y,Z)}_{=\,H(W \mid Y,Z)} \leq \underbrace{H(Z,W) - H(Z)}_{=\,H(W \mid Z)}$$

$$H(W \mid Y,Z) \leq H(W \mid Z) \checkmark$$

dropping conditioning.

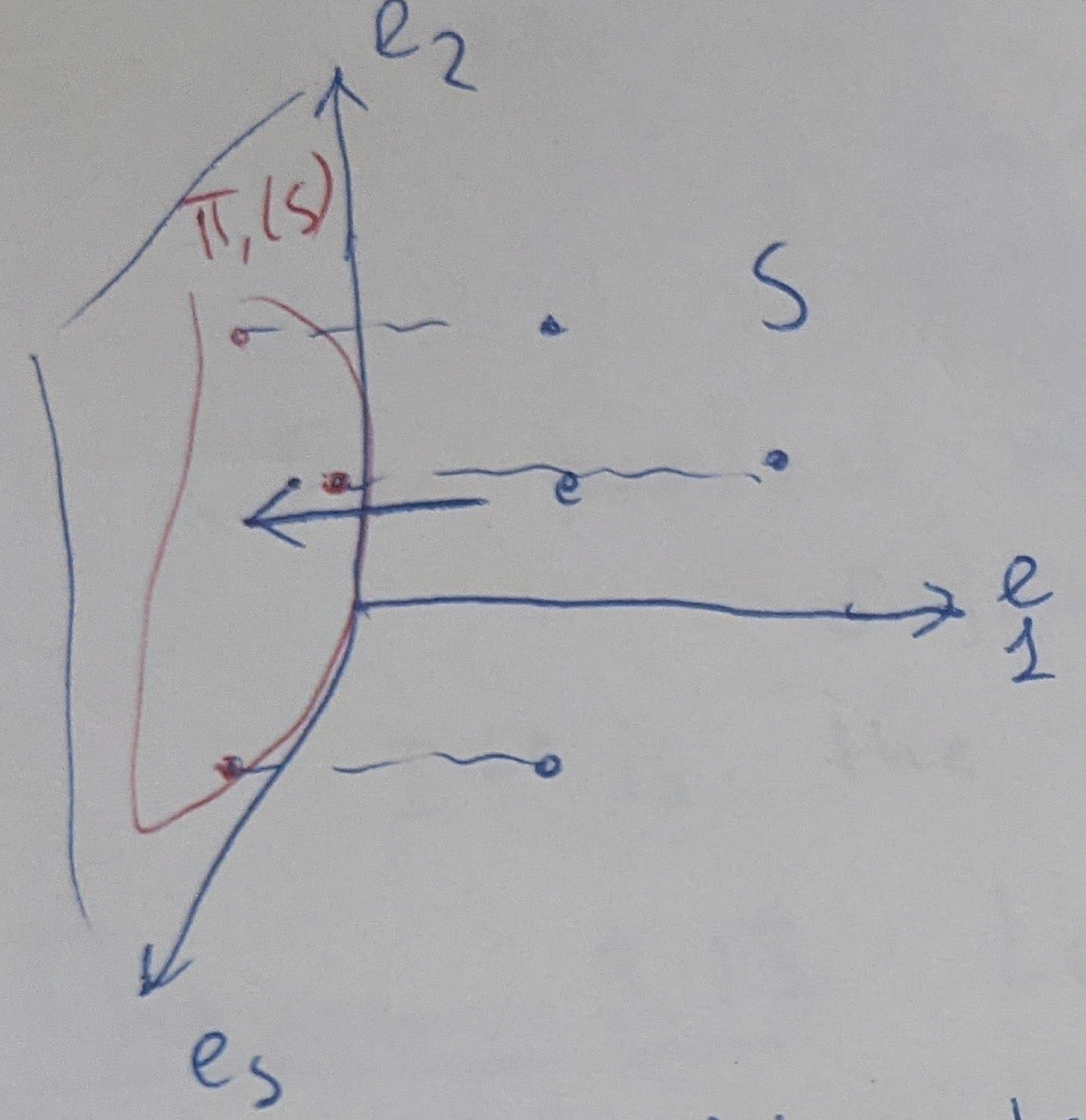
By the claim, $H(X_{A \cup B}) + H(X_{A \cap B}) \leq H(X_A) + H(X_B)$

$$\underbrace{X_{A \cap B} = Z, \quad X_{A \setminus B} = Y, \quad X_{B \setminus A} = W}$$

So for any A_i, A_j we can replace them by $A_i \cup A_j$ & $A_i \cap A_j$ in the lemma without loss of generality.

By doing this repeatedly we get to the case $A_i = [n]$ for some i.

Lemma 9.12: Let S be a finite set of points in $\mathbb{R}^d$. Let $\pi_i(S)$ be the projection of S onto a hyperplane orthogonal to basis vector e_i.



Then
$$|S|^{d-1} \leq \prod_{i=1}^{d} |\pi_i(S)|.$$

Proof: Apply 9.11 to the uniform distribution X on S then

$$(d-1)\log_2|S| = (d-1)H(X) \leq \sum_{i=1}^{d} H(\underbrace{X_1, X_2, \dots, \cancel{X_i}, X_{i+1})}_{= H(X_{[d]-i})}$$

X_i is i^{th} coordinate

$H(X_{[d]-i})$: a distribution on $\pi_i(S)$

$$\leq \sum_{i=1}^{d} \log_2|\pi_i(S)|. \checkmark$$

Corollary 9.13: For any $S \subseteq \mathbb{R}^d$ (not finite anymore)

$$\text{vol}^{d-1}(S) \leq \prod_{i=1}^{d} \text{vol}(\pi_i(S))$$

vol: d-dimensional. $\quad$ vol of $\pi_i(S)$: $(d-1)$ dimensional

Counting independent sets

(?) what is the maximum number of independent sets in a $\underbrace{\text{graph}}_{d\text{-regular}}$ with n vertices

d-regular $\rightarrow$ every vertex has degree d.?

(For perfect matchings the extremal example was disjoint union of $n/2d$ copies of $K_{d,d}$.)

It is the same for independent sets!

Theorem 9.13 (Kahn 2001, Zhao 2010): Let ind(G) denote the number of independent sets in G.

Then if G is a d-regular graph on n vertices, we have

$$\operatorname{ind}(G) \le \Big(\underbrace{\operatorname{ind}(K_{d,d})}_{?}\Big)^{n/2d}$$

$2^{d+1}-1$

Proof:

Kahn for bipartite graphs

It bootstraps to general graphs by considering $G \times K_2$.