ON SET COVERINGS IN CARTESIAN PRODUCT SPACES

Rudolf Ahlswede

Universität Bielefeld
Fakultät für Mathematik
Postfach 100131
33501 Bielefeld
Germany

Preface

This is a literal copy of a paper from the fall of 1971, which remained unpublished, because its main result had been obtained independently, but somewhat earlier in the paper "Hide and seek, data storage and entropy, Annals Math. Statistics 42, 1706–1716, 1971" by E.C. Posner and Mc Eliece.

Further developments have shown that our paper still has its merits, because of its methods. They immediately yield the Covering Theorem for the case of non–identical factors and a general class of infinite hypergraphs without any topological assumptions.

Furthermore there is a thorough discussion of connections to other solved and unsolved combinatorial problems. The paper under [5] appeared in the Journal of Applied Probability 10(4), 824–836, 1973. It deals with packings of controlled overlaps and led us to investigate coverings.

Only familiarity with a few elementary combinatorial and probabilistic concepts is demanded from the reader.

Abstract

Consider $(X, \mathcal{E})$, where X is a finite set and $\mathcal{E}$ is a system of subsets whose union equals X . For every natural number $n \in \mathbb{N}$ define the cartesian products $X_n = \prod_1^n X$ and $\mathcal{E}_n = \prod_1^n \mathcal{E}$. The following problem is investigated: how many sets of $\mathcal{E}_n$ are needed to cover X_n ? Let this number be denoted by $c(n)$. It is proved that for all $n \in \mathbb{N}$

$$\exp\{C \cdot n\} \leq c(n) \leq \exp\{Cn + \log n + \log\log |X|\} + 1.$$

A formula for C is given. The result generalizes to the case where X and $\mathcal{E}$ are not necessarily finite and also to the case of non–identical factors in the product. As applications one obtains estimates on the minimal size of an externally stable set in cartesian product graphs and also estimates on the minimal number of cliques needed to cover such graphs.

1. A covering theorem

Let X be a non–empty set with finitely many elements and let $\mathcal{E}$ be a set of non–empty subsets of X with the property $\bigcup\limits_{E\in\mathcal{E}} E = X$. (We do not introduce an index set for $\mathcal{E}$ in order to keep the notations simple). For $n \in \mathbb{N}$, the set of natural numbers, we define the cartesian product spaces $X_n = \prod_1^n X$ and $\mathcal{E}_n = \prod_1^n \mathcal{E}$. The elements of $\mathcal{E}_n$ can be viewed as subsets of X_n .

We say that $\mathcal{E}'_n \subset \mathcal{E}_n$ *covers* X_n or is a *covering* of X_n , if $X_n = \bigcup\limits_{E_n\in\mathcal{E}'_n} E_n$. We are interested in obtaining bounds on the numbers $c(n)$ defined by

$$c(n) = \min_{\mathcal{E}'_n \text{ covers } X_n} |\mathcal{E}'_n|, \ n \in \mathbb{N}. \tag{1.1}$$

Clearly, $c(n_1+n_2) \le c(n_1)\cdot c(n_2)$ for $n_1, n_2 \in \mathbb{N}$. Example 1 below shows that equality does not hold in general.

Denote by Q the set of all probability distributions on the finite set $\mathcal{E}$, denote by $1_E(\cdot)$ the indicator function of a set E , and define K by

$$K = \max_{q\in Q} \min_{x\in X} \sum_{E\in\mathcal{E}} 1_E(x) q_E. \tag{1.2}$$

Theorem.

With $C = \log K^{-1}$ *the following estimates hold:*

a) $c(n) \ge \exp\{C\cdot n\},\ n\in\mathbb{N}$.

b) $c(n) \le \exp\{C\cdot n + \log n + \log\log|X|\} + 1,\ n \in \mathbb{N}$.

c) $\lim_{n\to\infty} \frac{1}{n}\log c(n) = C$.

Proof.

c) is a consequence of a) and b). In order to show a) let us assume that $\mathcal{E}^*_{n+1}$ covers X_{n+1} and that $|\mathcal{E}^*_{n+1}| = c(n+1)$.

Write an element E_{n+1} of $\mathcal{E}^*_{n+1}$ as $E^1E^2\dots E^{n+1}$ and denote by ${}_xX_{n+1}$ the set of all those elements of X_{n+1} which have x as their first component. Finally, define a probability distribution q^* on $\mathcal{E}$ by

$$q^*_E = \big|\{E_{n+1} \mid E_{n+1}\in\mathcal{E}^*_{n+1}, E^1 = E\}\big| c^{-1}(n+1) \text{ for } E\in\mathcal{E}. \tag{1.3}$$

In order to cover the set ${}_xX_{n+1}$ we need at least $c(n)$ elements of $\mathcal{E}^*_{n+1}$. This and the definition of q^* yield

$$c(n+1)\sum_{E\in\mathcal{E}} 1_E(x)q_E^* \geq c(n). \tag{1.4}$$

Since (1.4) holds for all $x \in X$ we obtain

$$c(n+1)\min_{x\in X}\sum_{E\in\mathcal{E}} 1_E(x)q_E^* \geq c(n) \tag{1.5}$$

and therefore also

$$c(n+1)\max_{q\in Q}\min_{x\in X}\sum_{E\in\mathcal{E}} 1_E(x)q_E \geq c(n). \tag{1.6}$$

Inequality a) is an immediate consequence of (1.6).

We prove now b). Let r be an element of Q for which the maximum in (1.2) is assumed. Denote by r_n the probability distribution on $\mathcal{E}_n$, which is defined by

$$r_n(E_n) = \prod_{t=1}^n r_E t,\ E_n = E^1E^2\dots E^n \in \mathcal{E}_n. \tag{1.7}$$

Let N be a number to be specified later. Select now N elements $E_n^{(1)},\dots,E_n^{(N)}$ of $\mathcal{E}_n$ *independently* of each other according to the random experiment $(\mathcal{E}_n, r_n)$. If every $x_n \in X_n$ is covered by $\{E_n^{(1)},\dots,E_n^{(N)}\}$ with positive probability then there exists a covering of X_n with N sets. Let $x_n = (x^1,\dots,x^n)$ be any element of X_n . Define $\mathcal{E}(x_n)$ by

$$\mathcal{E}(x_n) = \{E_n | E_n \in \mathcal{E}_n, x_n \in E_n\}. \tag{1.8}$$

Clearly, $\mathcal{E}(x_n) = \prod_1^n\{E|E\in\mathcal{E}, x^t\in E\}$ and therefore

$$r_n\big(\mathcal{E}(x_n)\big) = \prod_{t=1}^n\left(\sum_E 1_E(x^t)r_E\right). \tag{1.9}$$

Recalling the definitions for r and K we see that $\sum_E\ 1_E(x^t)r_E \geq K$ and that

$$r_n\big(\mathcal{E}(x_n)\big) \geq K^n. \tag{1.10}$$

This implies that x_n is *not* contained in anyone of the N selected sets with a probability smaller than $(1-K^n)^N$ and therefore X_n is not covered by those sets with a probability smaller than $|X|^n(1-K^n)^N$.

Thus there exist coverings of cardinality N for all N satisfying

$$|X|^n(1-K^n)^N < 1. \tag{1.11}$$

Since $(1-K^n)^N \le \exp\{-K^n N\}$ one can choose any N satisfying

$$\begin{aligned}&\exp\{-K^n N\} \le \exp\{-\log|X|n\} \quad \text{or (equivalently)}\\ &N \ge \exp\{\log K^{-1}\cdot n + \log n + \log\log|X|\}.\end{aligned}$$

The proof is complete.

Probabilistic arguments like the one used here have been applied frequently in solving combinatorial problems, especially in the work of Erdös and Renyi. The cleverness of the proofs lies in the choice of the probability distribution assigned to the combinatorial structures. The present product distribution has been used for the first time by Shannon [2] in his proof of the coding theorem of Information Theory. For the packing problem defined in section 5 the present approach will not yield asymptotically optimal results.

<u>Example 1</u>

$X = \{0,1,2,3,4\},\ \mathcal{E} = \big\{\{x, x+1\} \mid x \in X\big\}$.

The addition is understood $\mod 5$. Clearly, $c(1) = 3$. We list the elements of X_2 as follows:

00	22	02	20	43	14	34	41
01	23	03	30	44	24	40	42
10	32	12	21	04			
11	33	13	31				

The elements in every column are contained in a set which is an element of $\mathcal{E}_2$. Therefore $c(2) \le 8 < c(1)^2$. Since in the present case $K^{-1} = \frac{5}{2}$ and since $c(2) \ge c(1)K^{-1} = \frac{15}{2} > 7$ we obtain that actually $c(2) = 8$. Moreover, since $\lim\limits_{n\to\infty} \frac{1}{n}\log c(n) = \log\frac{5}{2}$ there exists infinitely many n with $c(2n) < c^2(n)$.

2. Generalizations of the covering theorem

Let $(X^t, \mathcal{E}^t)_{t=1}^{\infty}$ be a sequence of pairs, where X^t is an arbitrary non–empty set and $\mathcal{E}^t$ is an arbitrary system of non–empty subsets of X^t . For every $n \in \mathbb{N}$ set $X_n = \prod\limits_{t=1}^{n} X^t$, $\mathcal{E}_n = \prod\limits_{t=1}^{n}\mathcal{E}^t$, and define $c(n)$ again as the smallest cardinality of a covering of X_n . Define Q^t , $t \in \mathbb{N}$, as the set of all probability distributions on $\mathcal{E}^t$ which are concentrated on a *finite* subset of $\mathcal{E}^t$. Finally, set

$$K^t = \sup_{q^t \in Q^t} \inf_{x^t \in X^t} \sum_{E^t \in \mathcal{E}^t} 1_{E^t}(x^t) q^t_{E^t}$$

and

$$C^t = \log(K^t)^{-1} \quad \text{for} \quad t \in \mathbb{N}.$$

A. The case of identical factors.

Let us assume that $(X^t, \mathcal{E}^t) = (X, \mathcal{E})$ for $t \in \mathbb{N}$. This implies that also $Q^t = Q$, $K^t = K$, and $C^t = C$ for $t \in \mathbb{N}$.

Corollary 1.

a) $c(n) \geq \exp\{C \cdot n\}$, $n \in \mathbb{N}$

b) For every $\delta > 0$ *there exists an* n_δ *such that*

$c(n) \leq \exp\{C \cdot n + \delta n\}$ *for* $n \geq n_\delta$.

c) $\lim_{n \to \infty} \frac{1}{n} \log c(n) = C$.

Proof.

If $c(1) = \infty$, then also $c(n) = \infty$ and a) is obviously true. b) holds in this case, because $K = 0$. If $c(1) < \infty$, then also $c(n) < \infty$. Replacing "max" by "sup" and "min" by "inf" the proof for a) of the theorem carries over verbally to the present situation. We prove now b). Choose r^* such that

$$|\log K^{-1} - \log \left(\inf_{x \in X} \sum_{E \in \mathcal{E}} 1_E(x) r^*_E \right)^{-1} | < \frac{\delta}{2}. \tag{2.1}$$

Let $\mathcal{E}^*$ be the finite support of r^* . We define an equivalence relation on X by

$$x \sim x' \text{ iff } \{E | E \in \mathcal{E}^*, x \in E\} = \{E | E \in \mathcal{E}^*, x' \in E\}. \tag{2.2}$$

Thus we obtain at most $2^{|\mathcal{E}^*|}$ many equivalence classes. Denote the set of equivalence classes by $\overline{X}$ and let $\overline{E}$ be the subset of $\overline{X}$ obtained from E by replacing it's elements by their equivalence classes. Write $\overline{\mathcal{E}} = \{\overline{E} | E \in \mathcal{E}\}$, $\overline{X}_n = \prod_1^n \overline{X}$, and $\overline{\mathcal{E}}_n = \prod_1^n \overline{\mathcal{E}}$. A covering of $\overline{X}_n$ induces a covering of X_n with the same cardinality. If follows from the theorem and from (2.1) that

$$c(n) \le \exp\left\{Cn + \frac{\delta}{2}n + \log n + \log\log 2^{|\mathcal{E}^*|}\right\} + 1. \tag{2.3}$$

This implies b). c) is again a consequence of a) and b).

B. Non–identical factors.

Corollary 2.

Assume that $\max_t |\mathcal{E}^t| \le a < \infty$. *Then for all* $n \in \mathbb{N}$ *:*

a) $c(n) \ge \exp\left\{\sum_{t=1}^{n} C^t\right\}$

b) $c(n) \le \exp\left\{\sum_{t=1}^{n} C^t + \log n + \log\log 2^a\right\} + 1$.

Proof.

Introducing equivalence relations in every X^t as before in X we see that it suffices to consider the case $\max_t |X^t| \le 2^a$. a) is proved as in case of identical factors. We show now b). Since $\mathcal{E}^t$ is finite there exists an r^t for which C^t is assumed. Replace the definition of r_n given in (1.7) by

$$r_n(E_n) = \prod_{t=1}^{r} r^t(E^t) \ \text{ for all } \ E_n = E^1 \dots E^n \in \mathcal{E}_n. \tag{2.4}$$

By the argument which led to (1.11) we obtain now

$$|2^a|^n \left(1 - \prod_{t=1}^{n} K^t\right)^N < 1 \tag{2.5}$$

and therefore b).

The condition on the $\mathcal{E}^t$'s can actually be weakened to the following uniformity condition:

For every $\delta > 0$ there exists an m_δ and r^t's with supports of cardinality smaller then m_δ such that

$$\left|\log(K^t)^{-1} - \log\left(\inf_{x^t \in X^t} \sum_{E^t \in \mathcal{E}^t} 1_{E^t}(x^t) r^t_{E^t}\right)^{-1}\right| \le \delta. \tag{2.6}$$

The upper bound on $c(n)$ which one then obtains is of course only of a sharpness as the one in b) of corollary 1.

Remark 1.

One can assign weights to the elements of $\mathcal{E}_n$ and than ask for coverings with minimal total weight. It may be of some interest to elaborate conditions on the weight function under which the covering theorem still holds. The weight function will of course enter the definition of K .

3. Hypergraphs. Duality

In this and later sections we consider only finite sets and products of finite sets, even though the results obtained can easily be generalized along the lines of section 2 to the infinite case. Thus we have the benefit of notational simplicity.

Let $X = \big\{x(i) | i = 1, \dots, a\big\}$ be a non–empty finite set and let $\mathcal{E} = \big(E(j) | j = 1, \dots, b\big)$ be a *family* of subsets of X . The pair $H = (X, \mathcal{E})$ is called a hypergraph (see [3]), if

$$\bigcup_{j=1}^{b} E(j) = X \quad \text{and} \quad E(j) \neq \varnothing \quad \text{for} \quad j = 1, \dots, b. \tag{3.1}$$

The $x(i)$'s are called vertices and the $E(j)$'s are called edges. A hypergraph is called simple, if $\mathcal{E}$ is a set of subsets of X . For the problems studied in this paper we can limit ourselves without loss of generality to simple hypergraphs and we shall refer to them shortly as hypergraphs. A hypergraph is a graph (without isolated vertices), if $|E(j)| \leq 2$ for $j = 1, \dots, b$. Interpreting $E(1), \dots, E(b)$ as points $e(1), \dots, e(b)$ and $x(1), \dots, x(a)$ as sets $X(1), \dots, X(a)$, where

$$X(j) = \big\{e(i) | i \leq a, x(j) \in E(i)\big\} \tag{3.2}$$

one obtains the dual hypergraph $H^* = (E^*, \mathcal{X}^*)$. A hypergraph is characterized by it's incidence matrix A . The incidence matrix of H^* is the conjugate of A . Let $H^t = (X^t, \mathcal{E}^t)$, $t \in \mathbb{N}$, be hypergraphs. For $n \in \mathbb{N}$ we define cartesian product hypergraphs $H_n = \prod\limits_{t=1}^{n} H^t$ by

$$H_n = (X_n, \mathcal{E}_n). \tag{3.3}$$

The covering theorem can be interpreted as a statement about edge coverings of cartesian product hypergraphs. We are looking now for the dual statement. One easily verifies that

$$H_n^* = (E_n^*, \mathcal{X}_n^*) = \prod_{t=1}^{n} (H^t)^*. \tag{3.4}$$

This means that the dual of the product hypergraph is the product of the dual hypergraphs. A set $T \subset X$ is called a transversal (or support) in $H = (X, \mathcal{E})$ if

$$T \cap E \neq \varnothing \text{ for all } E \in \mathcal{E}. \tag{3.5}$$

Denote the smallest cardinality of transversals in H_n (resp. H_n^*) by $t(n)$ (resp. $t^*(n)$). A transversal in H_n is a covering in H_n^* , and vice versa. Denoting the smallest cardinality of coverings in H_n^* by $c^*(n)$ we thus have

$$t(n) = c^*(n), t^*(n) = c(n), n \in \mathbb{N}. \tag{3.6}$$

Let now P be the set of all probability distributions on X and define K^* by

$$K^* = \max_{p \in P} \min_{E \in \mathcal{E}} \sum_{x \in X} 1_E(x) p_x. \tag{3.7}$$

K^* plays the same role for H_n^* as K does for H_n . The covering theorem implies

Corollary 3.

With $C^* = \log K^{*-1}$ *the following estimates hold for* $n \in \mathbb{N}$ *:*

a) $t(n) = c(n) \geq \exp\{C^* \cdot n\}$

b) $t(n) \leq \exp\{C^* \cdot n + \log + \log \log |\mathcal{E}|\} + 1$.

Of course the dual results to Corollaries 1, 2 also hold. There is generally no simple relationship between K and K^* . By choosing $\mathcal{E}$ as $\{\{x\} \mid x \in X\} \cup \{X\}$ we obtain $c(n) = 1, t(n) = |X|^n$, and therefore $K^* < K$ in this case. $K > K^*$ occurs for the dual problem. It may be interesting (and not too hard) to characterize hypergraphs for which $K^* = K$. We show now that K (resp. K^*) can be expressed as a function of P (resp. Q).

Lemma 1.

a) $K = \max\limits_{q \in Q} \min\limits_{x \in X} \sum\limits_{E \in \mathcal{E}} 1_E(x) q_E = \min\limits_{p \in P} \max\limits_{E \in \mathcal{E}} \sum\limits_{x \in X} 1_E(x) p_x = \overline{K}$,

b) $K^* = \max\limits_{p \in P} \min\limits_{E \in \mathcal{E}} \sum\limits_{x \in X} 1_E(x) p_x = \min\limits_{q \in Q} \max\limits_{x \in X} \sum\limits_{E \in \mathcal{E}} 1_E(x) q_E$.

Proof.

We have to show a) only since b) follows by dualization. P and Q are convex and compact in the supremum norm topology. The function $f(p,q) = \sum_{x \in X} \sum_{E \in \mathcal{E}} 1_E(x) p_x q_E$ is linear and continuous in both variables p and q. Therefore von Neumann's Minimax Theorem ([4]) is applicable and yields

$$\max_q \min_p \sum_x \sum_E 1_E(x) p_x q_E = \min_p \max_q \sum_x \sum_E 1_E(x) p_x q_E = M, \text{ say.} \tag{3.8}$$

Write K as $\max_q \min_{\delta_{x_0}} \sum_x \sum_E 1_E(x) \delta(x, x_0) q_E$, where δ_{x_0} is the probability distribution concentrated on x_0 and $\delta(\cdot,\cdot)$ is Kronecker's symbol. We see that $K \geq M$ and similarly that $M \geq \overline{K}$. For all p and q we have

$$\max_E \sum_x 1_E(x) p_x \geq \sum_E q_E \sum_x 1_E(x) p_x = \sum_x p_x \sum_E 1_E(x) q_E \geq \min_x \sum_E 1_E(x) q_E. \tag{3.9}$$

This implies $\overline{K} \geq K$ and thus $\overline{K} = K$. In studying infinite hypergraphs one could make use of more general Minimax Theorems, which have been proved by Kakutani, Wald, Nikaido, and others.

4. Applications to graphs

Let $G = (X, U)$ be a non–oriented graph without multiple edges. Define Γx by

$$\Gamma x = \{ y | y \in X, (x,y) \in U \}, x \in X. \tag{4.1}$$

Γx is the set of vertices connected with x by an edge. The graph G is completely described by X and Γ and we therefore also write $G = (X, \Gamma)$. Given a sequence of graphs $(G^t)_{t=1}^\infty$ then we define for every $n \in \mathbb{N}$ the cartesian product graphs $G_n = (X_n, \Gamma_n) = \prod_{t=1}^n G^t$ by

$$X_n = \prod_{t=1}^n X^t, \Gamma_n x_n = \prod_{t=1}^n \Gamma^t x^t \tag{4.2}$$

for all $x_n = (x^1, \dots, x^n) \in X_n$. (This product has also been called the cardinal product in the literature). Two vertices $x_n = (x^1, \dots, x^n)$ and $y_n = (y^1, \dots, y^n)$ of G_n are connected by an edge if and only if they are connected component–wise.

In the sequal we shall show that the covering theorem leads to estimates for some fundamental graphic parameters in case of product graphs.

A. The coefficient of external stability.

Given a graph $G=(X,\Gamma)$, a set $S, S\subset X$, is said to be externally stable if

$$\Gamma x\cap S\neq\varnothing \text{ for all } x\in S^c \tag{4.3}$$

or (equivalently) if

$$\bigcup_{x\in S}\big(\Gamma x\cup\{x\}\big)=X. \tag{4.4}$$

The coefficient of external stability $s(G)$ of a graph G is defined by

$$s(G)=\min_{S\text{ ext. stable}}|S|. \tag{4.5}$$

Finally, denote by $Q(X,\Gamma)$ the set of all probability distributions on $\{\Gamma y|y\in X\}$.

Corollary 4.

Let $G=(X,\Gamma)$ be a finite graph with all loops included, that is $x\in\Gamma x$ for all $x\in X$. With $\overline{C}=\log\left(\max_{q\in Q(X,\Gamma)}\min_{x\in X}\sum_{y\in X}1_{\Gamma y}(x)q_{\Gamma y}\right)^{-1}$ and $s(n)=s\left(\prod_1^n G\right)$ the following estimates hold for $n\in\mathbb{N}$:

a) $s(n)\geq\exp\{\overline{C}n\}$

b) $s(n)\leq\exp\{\overline{C}n+\log n+\log\log|X|\}+1$.

Proof.

Since $x\in\Gamma x$ by assumption we also have that $x_n\in\Gamma_n x_n=\prod_{t=1}^n\Gamma x^t$.

According to (4.4) $S_n\subset X_n$ is externally stable if and only if $\bigcup_{x_n\in S_n}\Gamma_n x_n=X_n$. Consider the hypergraph $H=(X,\mathcal{E})$, where $\mathcal{E}=\{\Gamma x|x\in X\}$, and it's product $H_n=(X_n,\mathcal{E}_n)$. An externally stable set S_n corresponds to a covering of X_n by edges of H_n , and vice versa. The corollary follows therefore from the covering theorem.

B. Clique coverings.

We recall that a clique in G is simply a complete subgraph of G . A clique is maximal if it is not properly contained in another clique.

Lemma 2.

Given $G_n = \prod_1^n G$ *, where* G *is a graph with an edge set containing all loops. The maximal cliques* M_n *in* G_n *are exactly those cliques which can be written as* $M_n = \prod_{t=1}^n M^t$ *, where the* M^t*'s are maximal cliques in* G *.*

Proof.

Products of maximal cliques are a maximal clique in the product graph. It remains to show the converse. Define

$B^t = \{x^t \mid \exists y_n = (y^1, \ldots, y^t, \ldots, y^n) \in M_n \text{ with } y^t = x^t\}; t = 1, 2, \ldots, n$.

The B^t's are cliques and therefore $B_n = \prod_{t=1}^n B^t$ is a clique in G_n containing M_n . Since M_n is maximal we have that $M_n = B_n$ and also that the B^t's are maximal. The system of cliques $\{M_n^{(i)} \mid i = 1, \ldots, m\}$ covers G_n if $\bigcup_{i=1}^m M_n^{(i)} = X_n$. We denote by $m(n)$ the smallest number of cliques needed to cover G_n . Define $\mathcal{M}$ as the set of all maximal cliques in G and define $Q(\mathcal{M})$ as set of all probability distributions on $\mathcal{M}$.

Corollary 5.

Let G *be a finite graph with all loops in the edge set.*

With $L = \log\left(\max_{q \in Q(\mathcal{M})} \min_{x \in X} \sum_{M \in \mathcal{M}} 1_M(x) q_M\right)^{-1}$ *the following estimates hold for* $n \in \mathbb{N}$ *:*

a) $m(n) \geq \exp\{Ln\}$

b) $m(n) \leq \exp\{Ln + \log n + \log\log |X|\} + 1$.

Proof.

It follows from Lemma 2 that clique coverings for G_n are simply edge coverings of the hypergraph $H_n = \prod_1^n H$, where $H = (X, \mathcal{M})$. The corollary is a consequence of the covering theroem.

Remark 2.

A clique covering of G_n can be interpreted as a colouring of the dual graph G_n^c . This graph can be written as $\prod_1^n G^c$, where the product is to be understood as follows: two vertices $x_n = (x^1, \ldots, x^n)$, $y_n = (y^1, \ldots, y^n) \in X_n$ are joined by an edge if for at least one t , $1 \le t \le n$, x^t and y^t are joined. Thus the corollary 5 gives estimates for minimal colorings of *–product graphs. The result of the present section generalize of course to the case of non–identical factor and also to the so called strong product.

5. A packing problem and it's equivalence to a problem by Shannon

A. The problem.

Instead of asking how many edges are needed to cover the set of all vertices of the hypergraph $H_n = (X_n, \mathcal{E}_n)$ one may ask how many non–intersecting edges can one pack into X_n . Formally, $\mathcal{E}'_n \subset \mathcal{E}_n$ is called a packing in H_n if $E_n \cap E'_n = \varnothing$ for all $E_{n'} E'_n \in \mathcal{E}'_n$. Define the maximal packing number $\pi(n)$ by

$$\pi(n) = \max_{\mathcal{E}'_n \text{ is packing in } H_n} |\mathcal{E}'_n|, \; n \in \mathbb{N}. \tag{5.1}$$

Using the argument which led to (1.6) one obtains

$$\pi(n+1) \le \pi(n) \left(\min_{q \in Q} \max_{n} \sum_{E \in \mathcal{E}} 1_E(x) q_E \right)^{-1} . \tag{5.2}$$

The inequality goes in the other direction and the roles of "max" and "min" are exchanged, because we are dealing with packings rather than with coverings. We know from Lemma 1 that $\min\limits_{q \in Q} \max\limits_{x} \sum\limits_{E \in \mathcal{E}} 1_E(x) q_E = K^*$. Since obviously $\pi(n) \le t(n)$ inequality (5.2) becomes trivial. Equality does not hold in general.

Example 2

$X = \{0,1,2\}$, $E(j) = \{j, j+1\}$ for $j = 0,1,2$. The addition is understood mod 3 . In this case $K^* = \frac{2}{3}$ and therefore $t(n) \ge \left(\frac{3}{2}\right)^n$. However, $\pi(n) = 1$ for all $n \in \mathbb{N}$.

B. The dual problem. Independent sets of vertices.

The packing problem for the dual hypergraph means the following for the original hypergraph: How many vertices can we select from X such that no two of them are contained in an edge? We are simply asking for the largest cardinality of a strongly independent set of vertices. We recall that $I \subset X$ is called a strongly independent set if and only if

$$|I \cap E| \le 1 \;\text{ for all }\; E \in \mathcal{E}. \tag{5.3}$$

$W \subset X$ is called a weakly independent set if and only if

$$|W \cap E| < |E| \ \text{ for all } \ E \in \mathcal{E}. \tag{5.4}$$

One easily verifies that a strongly independent set is also weakly independent provided that $|E| \geq 2$ for all $E \in \mathcal{E}$. (Loops are excluded.) If $H = H(G)$ is the hypergraph of a graph G without loops, then the two concepts are the same. A weakly independent set for $H(G)$ is simply an internally stable set for $G = (X, \Gamma)$, and conversely. $V \subset X$ is said to be an internally stable set of G if $V \cap \Gamma V = \varnothing$. This implies that no element of V has a loop. We would like to call a set $J \subset X$ with no 2 vertices joined by an edge a Shannon stable or briefly S–stable set of a graph, because this concept has been used by Shannon in [1] and because the difference between the two notions of stability seems not to have been emphasized enough in the literature even though it is significant for product graphs. In an S–stable set elements with loops are permitted. An internally stable set is S–stable. The converse is not necessarily true. $T \subset X$ is a transversal in G if every edge has at least one vertex in T . The complement of an internally stable set in G is a transversal in G , and vice versa. The same relationship holds for weakly independent sets and transversals in hypergraphs. Let $v(G_n)$ be the coefficient of internal stability of G_n , that is, the largest cardinality which can be obtained by an internally stable set in G_n and let $t(G_n)$ be the smallest cardinality for a transversal in G_n . We have

$$t(G_n) = |X|^n - v(G_n), n \in \mathbb{N}. \tag{5.5}$$

Denoting by $w(G_n)$ the largest cardinality of a weakly independent set in H_n and writing $t(H_n) = t(n)$ we also obtain

$$t(H_n) = |X|^n - w(H_n), n \in \mathbb{N}. \tag{5.6}$$

Our estimates for $t(H_n)$ (see section 3) can be translated into estimates for $w(H_n)$. However, those hypergraph results have no implications for $t(G_n)$ and $v(G_n)$. This is due to the fact that $H(G_n) \neq \prod_1^n H(G)$ in general. Actually, $v(G_n)$ is not a very interesting function of n . If $G = (X, U)$ is such that U contains all loops then also G_n contains all loops and $v(G_n) = 0$ for all $n \in \mathbb{N}$. If there exists an element $x \in X$ without a loop, then ${}_xX_{n-1}$ is internally stable in G_n and therefore $\lim_{n\to\infty} \frac{1}{n} \log v(G_n) = \log |X|$. This is also true in this case for $j(G_n)$, the largest cardinality of an S–stable set in G_n . Similarly one can show that $w(H_n) \equiv 0$ if H contains all loops and $\lim_{n\to\infty} \frac{1}{n} \log w(H_n) = \log |X|$ otherwise. In summarizing our discussion we can say that the following problems are unsolved:

1.) The transversal–problem for graphs not containing all loops in the edge set $\big(t(G_n)\big)$.

2.) The S–stability–problem for graphs with all loops in the edge set $\big(j(G_n)\big)$.

3.) The strong independence–problem for hypergraphs $\big(i(H_n)\big)$.

4.) The packing problem for hypergraphs $\big(\pi(n)\big)$.

A solution of 3.) for all hypergraphs is equivalent to a solution of 4.) for all hypergraphs, because the problems are dual to each other. Moreover, we notice that 2.) is a special case of 3.). Suppose that G is a graph with all loops in the edge set and that $H(G)$ is the hypergraph associated with G , then an S–stable set in G_n is a strongly independent set in $H(G)_n$, and conversely. We show that 4.) is a special case of 2.) and therefore that all three problems are equivalent. Let $H=(X,\mathcal{E})$ be a hypergraph. Define a graph $G(H)$ as follows:

Choose $\mathcal{E}$ as set of vertices and join $E,E'\in\mathcal{E}$ by an edge if and only if $E\cap E'\neq\varnothing$. $G(H)$ is a graph with all loops in the edge set and the packings of H_n are in one to one correspondence to the S–stable sets of $G(H)_n$.

<u>C. Shannon's zero error capacity.</u>

Problem 2.) is due to Shannon [1]. It is a graph theoretic formulation of the information theoretic problem of determining the maximal number of messages which can be transmitted over a memoryless noisy channel with error probability zero. $\lim\limits_{n\to\infty}\dfrac{1}{n}\log j(G_n)$ was called in [1] the zero error capacity C_o , say. Using our standard argument (see (1.6) and (5.2)) one can show that for $G=(X,U)$, where U contains all loops,

$$j(G_{n+1})\leq j(G_n)\left(\min_{p\in P}\ \max_{E\in U}\ \sum_{x\in X}1_E(x)p_x\right)^{-1}. \tag{5.7}$$

This implies that

$$C_o\leq\log\left(\min_{p\in P}\ \max_{E\in U}\ \sum_{x\in X}1_E(x)p_x\right)^{-1}. \tag{5.8}$$

It has been shown in [6] that for bipartite graphs $j(G_n)=\big(j(G)\big)^n$ for all $n\in\mathbb{N}$ and hence that $C_o=\log j(G)$ in this case. The proof uses the marriage theorem. The simplest non–bipartite graph for which C_o is unknown is the pentagon graph.

It was shown in [1] that in this case

$$\frac{1}{2}\log 5\leq C_o\leq\log\frac{5}{2}. \tag{5.9}$$

The lower bound is an immediate consequence of the equation $j(G_2)=5$. The upper bound follows also from (5.8). No improvement has been made until now on any of those bounds. We have been able to prove that

$$j(G_3) = 10,\ j(G_4) = 25,\ j(G_5) = 50, \quad \text{and} \quad j(G_6) = 125. \tag{5.10}$$

We conjecture that

$$j(G_{2n}) = 5^n,\ j(G_{2n+1}) = 2.5^n \quad \text{for all} \quad n \in \mathbb{N}, \tag{5.11}$$

but so far we have no proof for $n > 6$. (5.11) would imply $C_o = \frac{1}{2} \log 5$. The result announced in (5.10) and results which go beyond this (including colouring problems) will appear elsewhere. We would like to mention that we came to the covering problem by trying to understand the results of [5] from a purely combinatorial point of view. Those results can be understood as statements about "packings with small overlapping and an additional weight assignment". It seems to us that the methods of [5] allow refinements which may be helpful for the *construction* of minimal coverings. We expect that the covering theorem has applications in Approximation Theory, in particular for problems involving ε–entropy (see [7]).

It might also be of some interest to compare our estimates with known results (see [8]) on coverings with convex sets in higher dimensional spaces.

References

[1] Shannon, C.E., The zero–error capacity of a noisy channel. IRE Trans. Inform. Theory, IT–2, 8–19, 1956.

[2] Shannon, C.E., Certain results in Coding Theory for noisy channels. Inform. and Control, 1, 6–25, 1957.

[3] Berge, C., Graphes et hypergraphes. Monographies universitaires de mathematiques. Dunod Paris, 1970.

[4] Neumaun, von J.; Morgenstern, O., Theory of Games and Economic Behaviour. Princeton, Princeton University Press, 1944.

[5] Ahlswede, R., Channel capacities for list codes. To appear in Journal f. Applied Probability.

[6] Markosian, A.G., The number of internal stability in a cartesian product graph and it's application to information theory (In Russian). Presented at the Second International Symposium on Information Theory, Sept. 2–8, 1971, Tsahkadsor, Armenian SSR.

[7] Kolmogoroff, A.N.; Tichomirow, W.M., ε–Entropie and ε–Kapazität von Mengen in Funktionalräumen. Arbeiten zur Informationstheorie III, Mathematische Forschungsberichte, VEB Deutscher Verlag der Wissenschaften, Berlin 1960. (Translation from the Russian).

[8] Rogers, C.A., Packing and Covering. Cambridge Tracts in Mathematics and Mathematical Physics, University Press, Cambridge, 1964.