Maximal independent sets in the covering graph of the cube

Dwight Duffus [a], Peter Frankl [b], Vojtěch Rödl [a,*]

[a] *Mathematics & Computer Science Department, Emory University, Atlanta, GA 30322, USA*
[b] *Tokyo, Japan*

ARTICLE INFO

ABSTRACT

Several familiar problems in extremal set theory can be cast as questions about the maximum possible size of an independent set defined on a suitable graph, about the total number of independent sets in such graphs, or about enumeration of the maximal independent sets. Here we find bounds on the number of maximal independent sets in the covering graph of a hypercube.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

We consider problems from extremal set theory recast as questions about independent sets in graphs, usually graphs defined on the power set of a finite set. In particular, we are interested in enumerating all independent sets and all maximal [non-extendable] independent sets. The main results provide bounds for the graph C_n on the cube, defined as follows: the vertices of C_n are the 0, 1-vectors of length n, with two adjacent if they differ in exactly one coordinate. Alternatively, we can regard each vector $(a_1, a_2, \ldots, a_n)$ as the subset A of $[n] = \{1, 2, \ldots, n\}$ where $A = \{i \in [n] \mid a_i = 1\}$.

Let X be the set of subsets with an even number and Y, the set of subsets with an odd number of elements. Then these sets form a bipartition of C_n, that is, $C_n = (X, Y; E)$ such that for all $A \in X$ and $B \in Y$,

$$AB \in E \iff |A \Delta B| = 1.$$

If we consider the power set as partially ordered by set containment, then C_n is just the *covering graph* of the power set lattice.

We will consider a few other graphs defined on the power set. For instance, the *Sperner graph* S_n has edges all pairs where one set is properly contained in the other, that is, S_n is the *comparability graph* of the power set lattice. The classical theorem of Sperner [14] can be stated as follows:

if S is an independent set of vertices in S_n then $|S| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$.

Dedekind's problem [2] concerns the same graph:

determine the total number of independent sets in S_n.

A celebrated theorem of Kleitman [8] provides the first solution to Dedekind's problem:

the total number of independent sets in S_n is $2^{\binom{n}{\lfloor \frac{n}{2} \rfloor}(1+o(1))}$.

---

* Corresponding author.
E-mail addresses: dwight@mathcs.emory.edu (D. Duffus), peter.frankl@gmail.com (P. Frankl), rodl@mathcs.emory.edu (V. Rödl).

The above result states that in a sense the average independent set is just a subset of one of the maximum-sized independent sets in S_n, namely, the family of all subsets of $[n]$ of size $\lfloor \frac{n}{2} \rfloor$ or all subsets of size $\lceil \frac{n}{2} \rceil$.

There is a considerable literature on Dedekind's problem, including improvements of the $o(1)$ term [10] and an asymptotic formula for the total number [11,13]. In particular, Kahn [6] has developed entropy arguments that allow enumeration of independent sets in bipartite graphs and of antichains in graded partially ordered sets, from which the results in [10] follow.

The starting point for our work is the problem of finding better upper bounds for the total number of *maximal* independent sets in S_n, that is, better than the bounds from Kleitman's result. First, we introduce some notation that will simplify our presentation. Given a graph G, let tis(G) be the total number of independent sets in G, let $\mathcal{I}(G)$ denote the family of all maximal independent sets in G, and let mis(G) = $|\mathcal{I}(G)|$.

Problem 1. Determine the asymptotic value of the $\log_2$ mis(S_n). In particular, improve the upper bound inherited from Kleitman's result on $\log_2$ tis(S_n).

In a related paper [3] we addressed this problem and obtained some partial results. However, we were unable to obtain a bound of the form $2^{\binom{n}{\lfloor \frac{n}{2} \rfloor}(\alpha + o(1))}$ for some fixed $\alpha < 1$. Here we are concerned with the graph C_n.

Problem 2. Determine the asymptotic value of $\log_2$ mis(C_n).

Our main result is a first step in this direction. We are inclined to believe the lower bound in the following.

Theorem 1. *For all n,*

$$2^{n-2} \leq \log_2 \text{mis}(C_n) \leq 0.78(1 + o(1))2^{n-1},$$

where $o(1) \to 0$ as $n \to \infty$.

2. Preliminaries

We denote a bipartite graph G with vertex sets X and Y and edge set E by $G = (X, Y; E)$. For $X' \subseteq X$, let

$$S(X') = \{y \in Y \,|\, xy \in E \text{ for some } x \in X'\}$$

and call this the *span* of X' in G. Say that $Y' \subseteq Y$ is a *spanned set* or a *span* if $Y' = S(X')$ for some $X' \subseteq X$. Write $S(x)$ in place of $S(\{x\})$. Given $F \subseteq E$, $\cup F$ is the set of vertices of G that belong to edges in F. We say that F is an *induced* matching if the subgraph $G[\cup F]$ of G induced by $\cup F$ is a matching. We also say that F is a *cover* of G if every edge of G contains some member of $\cup F$.

We now partition the parts X and Y of bipartite graph $C_n = (X, Y; E)$. For $i = 1, 2, \ldots, n$, let

$$X_i^0 = \{A \in X \mid i \notin A\}, \qquad X_i^1 = \{A \in X \mid i \in A\}, \tag{1}$$

$$Y_i^0 = \{B \in Y \mid i \notin B\}, \qquad Y_i^1 = \{B \in Y \mid i \in B\}. \tag{2}$$

For subsets $X' \subseteq X$ and $Y' \subseteq Y$, we use $[X', Y']$ to denote the set of edges in the subgraph $C_n[X' \cup Y']$ of C_n induced by the vertex set $X' \cup Y'$.

Each of the following statements is an easy consequence of these definitions.

Fact 1. *Let $C_n = (X, Y; E)$. Then the following hold.*
(a) *E is partitioned by $E = \bigcup_{i=1}^{n} [X_i^0, Y_i^1] \cup [X_i^1, Y_i^0]$.*
(b) *For $i = 1, 2, \ldots, n$, $C_n[X_i^0 \cup Y_i^1]$ and $C_n[X_i^1 \cup Y_i^0]$ are both induced matchings in C_n, each containing 2^{n-2} edges.*
(c) *For $i = 1, 2, \ldots, n$, $C_n[X_i^0 \cup Y_i^0]$ and $C_n[X_i^1 \cup Y_i^1]$ are both $(n-1)$-regular subgraphs with $(n-1)$-factorizations in C_n.*

The following is proved in [3]—we include the straightforward verification here for completeness. As noted by one of the referees, the same proof establishes these statements for all graphs.

Lemma 1. *Let $G = (X, Y; E)$ be a bipartite graph.*
(a) *For any induced matching M of G, $2^{|M|} \leq |\mathcal{I}(G)|$.*
(b) *For any matching M that is a cover of G, $|\mathcal{I}(G)| \leq 3^{|M|}$.*

Proof. (a) There are exactly $2^{|M|}$ subsets of $\cup M$ that contain just one element from each edge in M. Each such set extends to a maximal independent set of G with the addition of vertices from $X \cup Y - (\cup M)$. Consequently, $2^{|M|} \leq |\mathcal{I}(G)|$.
(b) Each independent set in $\cup M$ contains at most one vertex from each edge in M, so there are at most $3^{|M|}$ independent sets contained in $\cup M$. For all $I \in \mathcal{I}(G)$, $I \cap (\cup M)$ is an independent set in $\cup M$ and for all $z \in X \cup Y - (\cup M)$, $z \in I$ if and only if $z \notin S(I \cap (\cup M))$. Thus, each independent set contained in $\cup M$ can be extended to at most one maximal independent set of G. This shows that $|\mathcal{I}(G)| \leq 3^{|M|}$. □

The observation in Lemma 1(a) above was suggested to us by Kahn for the bipartite graph induced by the middle two levels of the Boolean lattice, for n odd.

The lower bound in Theorem 1 is an immediate consequence of Fact 1(b) and Lemma 1(a).

3. The upper bound for maximal independent sets

We first give two observations about matchings in general bipartite graphs, then obtain the upper bound in Theorem 1. The first result records the obvious limiting effect of induced matchings in bipartite graphs on the number of independent subsets.

Fact 2. *Let $G = (X, Y; E)$ have a perfect matching of X onto Y. For any $\widetilde{X} \subseteq X$ and $\widetilde{Y} \subseteq Y$, the number of sets $X_0 \subset X$ such that $X_0 \cap \widetilde{X} = \emptyset$ and $S(X_0) \cap \widetilde{Y} = \emptyset$ is at most $2^{|X| - \max(|\widetilde{X}|, |\widetilde{Y}|)}$.*

The following lemma guarantees expansion in span sets when degree conditions are assumed.

Lemma 2 *([3]). Let $G = (X, Y; E)$ and let c and d be positive integers such that $\delta(X) \geq c$ and $\Delta(Y) \leq d$. Then for all $p \leq c$ there exists a q-element set $\widetilde{X} \subseteq X$ such that*

$$q = \left\lceil \frac{|X|(c - p + 1)}{c + pd - p + 1} \right\rceil, \quad and \quad |S(\widetilde{X})| \geq qp.$$

To verify the upper bound in Theorem 1, let I be a maximal independent set in C_n, fix $i \in [n]$, and let

$$A_0 = I \cap X_i^0, \qquad A_1 = I \cap X_i^1,$$
$$B_0 = I \cap Y_i^0, \qquad B_1 = I \cap Y_i^1.$$

Note that I is determined by $A_0 \cup A_1$ and by $B_0 \cup B_1$. We overestimate the number of maximal independent sets by finding an upper bound on the number of pairs (A_0, A_1) or (B_0, B_1).

We shall use the (binary) entropy function,

$$H(\alpha) = \alpha \log_2(1/\alpha) + (1 - \alpha) \log_2(1/(1 - \alpha)).$$

(See [1] for properties of this function and [6] for its applications to enumeration of independent sets and antichains.)

Suppose that $|B_0| + |B_1| = \beta 2^{n-1}$. For each β, an upper bound on the number of pairs (B_0, B_1), and, hence, an upper bound for $\text{mis}(C_n)$, is given by

$$\text{mis}(C_n) \leq \binom{2^{n-1}}{\beta 2^{n-1}} = 2^{2^{n-1}(1+o(1))H(\beta)}. \tag{3}$$

We need an accompanying bound on the number of pairs (A_0, A_1). For this, we apply Lemma 2 to the bipartite graph $G = (X, Y; E) = C_n[B_1 \cup X_i^1]$. That is, $X = B_1$ and $Y = X_i^1$. Then the degree of vertices in X is $n - 1$ because C_n is n-regular and each vertex of B_1 is adjacent to precisely one vertex in X_i^0, by Fact 1(b), and has all other neighbors in X_i^1. The maximum degree in G of vertices in Y is $n - 1$, as can be seen from Fact 1(b). Thus, Lemma 2 does indeed apply with $c = d = n - 1$. With $p = \sqrt{n}$, we obtain $\widetilde{B}_1 \subseteq B_1$ such that

$$|\widetilde{B}_1| = \left\lceil \frac{|B_1|(n - 1 - \sqrt{n} + 1)}{n - 1 + \sqrt{n}(n - 1) - \sqrt{n} + 1} \right\rceil, \quad and \quad |S(\widetilde{B}_1)| \geq |\widetilde{B}_1|\sqrt{n}.$$

Thus,

$$|\widetilde{B}_1| = \frac{|B_1|}{\sqrt{n}}(1 - o(1)), \quad and \quad |S(\widetilde{B}_1)| \geq |B_1|(1 - o(1)). \tag{4}$$

We now apply Lemma 2 to $G = (X, Y; E) = C_n[X_i^1 \cup Y_i^0]$, which is a perfect matching, by Fact 1(b). We set $\widetilde{X} = S(\widetilde{B}_1)$ and $\widetilde{Y} = B_0$. With (4) and Fact 2, we see that the number of possible sets A_1 is at most

$$2^{2^{n-2} - \max(|B_0|, |S(\widetilde{B}_1)|)} \leq 2^{2^{n-2} - \max(|B_0|, |B_1|(1-o(1)))} \tag{5}$$

once $\widetilde{B}_1$ is determined. (See Fig. 1 for a schematic of these sets.)

Similarly, we can apply Lemma 2 to the bipartite graph $G = (X, Y; E) = C_n[B_0 \cup X_i^0]$. This time $X = B_0$ and $Y = X_i^0$. Now, once $\widetilde{B}_0$ is specified, the number of choices of A_0 is at most

$$2^{2^{n-2} - \max(|B_1|, |S(\widetilde{B}_0)|)} \leq 2^{2^{n-2} - \max(|B_1|, |B_0|(1-o(1)))}. \tag{6}$$

Thus, given $\widetilde{B}_0$ and $\widetilde{B}_1$, we multiply the right-hand sides of (5) and (6) to see that the number of choices of (A_0, A_1) is at most

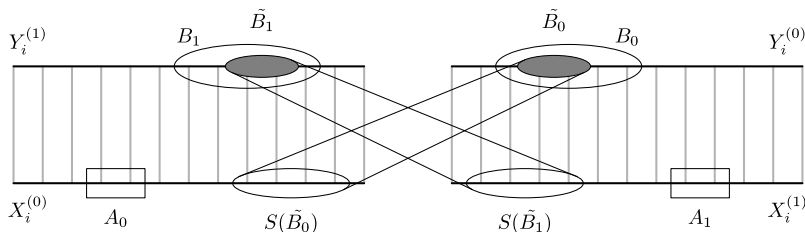$$2^{2^{n-1} - \max(|B_0|, |B_1|(1-o(1))) - \max(|B_1|, |B_0|(1-o(1)))} \leq 2^{2^{n-1}(1-\beta)}.$$

Fig. 1. An independent set $I = (A_0 \cup A_1) \cup (B_0 \cup B_1)$ in C_n.

Since I is determined by $A_0 \cup A_1$, we simply multiply this bound by an upper bound on the number of choices of the pair $(\widetilde{B}_0, \widetilde{B}_1)$. For each choice of β, this gives the following bound for $\mathrm{mis}(C_n)$:

$$\mathrm{mis}(C_n) \leq \left(\sum_{i=0}^{\frac{2^{n-2}}{\sqrt{n}}(1-o(1))} \binom{2^{n-2}}{i} \right)^2 \cdot 2^{2^{n-1}(1-\beta)}$$

$$= 2^{o(2^{n-1})} \cdot 2^{2^{n-1}(1-\beta)}$$

$$= 2^{2^{n-1}(1+o(1))(1-\beta)}. \tag{7}$$

Now consider the bounds on $\mathrm{mis}(C_n)$ provided by (3) and (7). The root β_0 of $H(\beta) = 1-\beta$ satisfies $1-\beta_0 \geq \min(H(\beta), 1-\beta)$ for all $0 \leq \beta \leq 1$. We find that $\beta_0 = 0.2271$.

We can finally count the total number of maximal independent sets, depending upon $|B_0| + |B_1| = \beta 2^{n-1}$. If $\beta \leq \beta_0$ then we use the bound in (3); if $\beta > \beta_0$ then we use the bound in (7). Since there are 2^{n-1} possibilities for $|B_0| + |B_1|$, we infer that

$$\mathrm{mis}(C_n) \leq 2^{n-1} \cdot 2^{0.7729(1+o(1))2^{n-1}}$$

$$\leq 2^{0.7729(1+o(1))2^{n-1}},$$

which completes the proof of Theorem 1.

4. Further problems

In the Introduction we noted that many results of extremal set theory can be formulated as the determination or estimation of the maximum size of an independent set in an appropriately defined graph on the power set of a finite set. Here we are concerned with estimating $\log_2 \mathrm{tis}(G)$ and $\log_2 \mathrm{mis}(G)$ for such G.

We began the paper with the example of the Sperner graph S_n. The Sperner theorem provides the maximum size of an independent set and a complete description of those sets. Kleitman's theorem settles Dedekind's problem, as far as the asymptotics of $\log_2 \mathrm{tis}(S_n)$ is concerned [8]. We have made some progress on the asymptotics of $\log_2 \mathrm{mis}(S_n)$ [3].

Here is a second class of graphs. Given $0 < s < n$, recall that a family $\mathcal{F}$ of sets is s-intersecting if for all $A, B \in \mathcal{F}$, $|A \cap B| \geq s$. Now define the graph $I_{n,s}$ whose vertex set is the power set of $[n]$ and with A and B adjacent if $|A \cap B| < s$. Then independent sets corresponds to s-intersecting families. The maximum size of an independent set is given by the Erdős–Ko–Rado Theorem in the case $s = 1$ [5]. A well-known result of Katona [7] provides the maximum size $k(n, s)$ of an s-intersecting family:

$$k(n, s) = \begin{cases} \displaystyle\sum_{i=\frac{n+s}{2}}^{n} \binom{n}{i}, & n + s \text{ even}; \\ \displaystyle\sum_{i=\frac{n+s-1}{2}}^{n-1} \binom{n-1}{i}, & n + s \text{ odd.} \end{cases} \tag{8}$$

A complete description of the maximum-sized families is also given.

We can obtain the asymptotic value of $\log_2 \mathrm{tis}(I_{n,s})$ in case s is fixed and n is large.

Theorem 2. *For s a fixed integer and $n \to \infty$,*

$$\mathrm{tis}(I_{n,s}) = 2^{k(n,s)(1+o(1))},$$

where $o(1) \to 0$ as $n \to \infty$.

Proof. Since there is an independent set of size $k(n, s)$ in $I_{n,s}$, it is immediate that $2^{k(n,s)} \leq \mathrm{tis}(I_{n,s})$.

To see the upper bound, note that every independent set is contained in a maximal one. Each maximal independent set has size at most $k(n, s)$. And every maximal independent set in $I_{n,s}$ is an order filter (is closed under taking supersets) and

so is determined by the antichain of its minimal elements. Therefore,

$$
\begin{aligned}
\text{tis}(I_{n,s}) &\leq \text{mis}(I_{n,s}) \cdot 2^{k(n,s)} \\
&\leq \text{tis}(S_n) \cdot 2^{k(n,s)} \\
&\leq 2^{\binom{n}{\lfloor n/2 \rfloor}(1+o(1))} \cdot 2^{k(n,s)}.
\end{aligned}
$$

The last inequality is Kleitman's theorem. From Stirling's formula and Katona's result (8), $k(n, s) = 2^{n-1}(1 + o(1))$, as long as $s = o(\sqrt{n})$. Since we take s to be fixed and $n \to \infty$, we have

$$
\binom{n}{\lfloor n/2 \rfloor} (1 + o(1)) + k(n, s) = k(n, s)(1 + o(1)).
$$

This completes the proof. □

Now let us consider estimating $\log_2 \text{mis}(I_{n,s})$. This has been done for $s = 1$, that is, for intersecting families, by Erdős and Hindman [4]. Let $I_n = I_{n,1}$. They proved that

$$
\log_2(\text{mis}(I_n)) \sim \binom{n-1}{\lfloor n/2 \rfloor}. \tag{9}
$$

(They also say somewhat more about the lower bound for $\text{mis}(I_n)$.) The upper bound in (9) follows from the observation that a maximal intersecting family is determined by (the antichain of) those of its minimal elements that do not contain some fixed element. Their proof of the lower bound is quite short and straightforward. It can also be obtained by observing that for $n = 2t$ ($n = 2t + 1$), the collection of all t-element subsets of $[n]$ (respectively, all t-element subsets containing a fixed element and their complements) induces a matching with $\binom{2t}{t}/2$ (respectively, $\binom{2t}{t-1}$) edges. In the review of the Erdős–Hindman paper [4], Kleitman [9] states that a sharper asymptotic formula for the number of maximal intersecting families (as opposed to the asymptotics of the log of this number) might be derived from Korshunov's description of the antichains in the power set lattice [11].

For $s > 1$, one can easily obtain a lower bound for $\text{mis}(I_{n,s})$ as follows. Let $\mathcal{F}$ be a maximal independent set in $I_{n-(s-1)}$, that is, a maximal intersecting family on the set $[n - (s - 1)]$. Let $X = \{n - s + 2, n - s + 3, \ldots, n\}$ and let $\mathcal{F}'$ be the family of subsets of $[n]$ defined by $\mathcal{F}' = \{A \cup X \mid A \in \mathcal{F}\}$. Then $\mathcal{F}'$ is s-intersecting and so is an independent set in $I_{n-(s-1)}$. Extend $\mathcal{F}'$ to a maximal independent set $\mathcal{F}''$. It is obvious that if $\mathcal{F}_0$ and $\mathcal{F}_1$ are distinct maximal independent sets in $I_{n-(s-1)}$ then $\mathcal{F}'_0 \neq \mathcal{F}'_1$. Without loss of generality, let $A \in \mathcal{F}_0 - \mathcal{F}_1$ and $A \cup X \in \mathcal{F}'_0 - \mathcal{F}'_1$. Now suppose that $\mathcal{F}''_0 = \mathcal{F}''_1$. Then $A \cup X \in \mathcal{F}''_1$. Since $\mathcal{F}''_1$ is s-intersecting, $|(A \cup X) \cap (B \cup X)| \geq s$ for all $B \in \mathcal{F}_1$. Hence, $|A \cap B| \geq 1$ for all $B \in \mathcal{F}_1$. Since $\mathcal{F}_1$ is a maximal intersecting family, we must have that $A \in \mathcal{F}_1$. This is a contradiction. Therefore, the mapping $\mathcal{F} \to \mathcal{F}''$ is a 1–1 map of the family of maximal independent sets of $I_{n-(s-1)}$ into the family of maximal independent sets in $I_{n,s}$. It follows that

$$
\text{mis}(I_{n,s}) \geq \text{mis}(I_{n-(s-1)}) \geq 2^{\binom{n-s}{\lfloor (n-s)/2 \rfloor}}.
$$

At this time, this is all we can say about $\text{mis}(I_{n,s})$ and we pose this

Problem 3. Determine the asymptotics of $\log_2(\text{mis}(I_{n,s}))$ for s fixed and $n \to \infty$.

We close with a very natural graph theoretic question that can be formulated in similar language. Indeed, this extremal question may occur to readers, so it is worth noting that it has been completely solved! Erdős and Moser asked for the maximum number $f(n)$ of maximal independent sets that an n-vertex graph could have. With our notation, they asked to determine

$$
f(n) = \max_G \text{mis}(G),
$$

where the maximum is taken over all n-vertex graphs G. Moon and Moser [12] determined $f(n)$ exactly and gave a complete description of the extremal graphs: for all $n \geq 2$

$$
f(n) = \begin{cases} 3^{n/3}, & \text{if } n \equiv 0 \pmod 3; \\ 4 \cdot 3^{\lfloor n/3 \rfloor - 1}, & \text{if } n \equiv 1 \pmod 3; \\ 2 \cdot 3^{\lfloor n/3 \rfloor}, & \text{if } n \equiv 2 \pmod 3. \end{cases} \tag{10}
$$

The extremal graphs realizing the values in (10) are disjoint unions of complete graphs, as many that are triangles as possible and the remaining ones either edges or complete graphs on 4 vertices.

Acknowledgement

References

 [1] N. Alon, J. Spencer, The Probabilistic Method, 2nd ed., John Wiley & Sons, New York, 2000.
 [2] R. Dedekind, Uber Zerlegungen von Zahlen durch ihre grössten gemeinsamen Teiler, Festschrift Hoch. Braunschweig w. ges. Werke II (1897) 103–148.
 [3] D. Duffus, P. Frankl, V. Rödl, Maximal independent sets in Boolean graphs, The Eur. J. Comb. 32 (2011) 1–9.
 [4] P. Erdős, N. Hindman, Enumeration of intersecting families, Discrete Math. 48 (1984) 61–65.
 [5] P. Erdős, C. Ko, R. Rado, Intersection theorems for systems of finite sets, Quart. J. Math. Oxford (2) 12 (1961) 313–320.
 [6] J. Kahn, Entropy, independent sets and antichains: a new approach to Dedekind's problem, Proc. Amer. Math. Soc. 130 (2001) 371–378.
 [7] G.O.H. Katona, Intersection theorems for systems of finite sets, Acta Math. Acad. Sci. Hung. 15 (1964) 327–337.
 [8] D. Kleitman, On Dedekind's problem: the number of isotone Boolean functions, Proc. Amer. Math. Soc. 21 (1969) 677–682.
 [9] D. Kleitman, Review of [4], Math. Reviews, MR732200 (85h:05007).
[10] D. Kleitman, G. Markowsky, On Dedekind's problem: the number of isotone Boolean functions II, Trans. Amer. Math. Soc. 213 (1975) 373–390.
[11] A.D. Korshunov, The number of monotone Boolean functions, Problemy Kibernet. 38 (1981) 5–108 (in Russian).
[12] J. Moon, L. Moser, On cliques in graphs, Israeli J. Math. 3 (1965) 23–28.
[13] A.A. Sapozhenko, The number of antichains in multilayered ranked sets, Diskret. Mat. 1 (1989) 110–128 (in Russian). Translated in Discrete Math. Appl. 1 (1991) 149–169.
[14] E. Sperner, Ein Satz über Untermengen einer endlichen Mengen, Math. Z. 27 (1928) 544–548.