

15

Epsilon Nets and Transversals of Hypergraphs

For historical reasons, a finite set-system is often called a hypergraph. More precisely, a *hypergraph* H consists of a finite set $V(H)$ of *vertices* (points) and a family $E(H)$ of subsets of $V(H)$. The elements of $E(H)$ are usually called *hyperedges* (or, in short, *edges*). If the hyperedges of H are r -element sets, then H is said to be an *r -uniform hypergraph*. Using this terminology, a graph is a two-uniform hypergraph. In Chapter 10 we have extended some graph-theoretic results to r -uniform hypergraphs (cf. Theorems 10.11 and 10.12).

The concept of hypergraphs is a very general one, so it is not surprising that hypergraph theory has a large scale of applications in various fields of mathematics, including geometry. Given a hypergraph H , a subset $T \subseteq V(H)$ is called a *transversal* of H if $T \cap E$ is nonempty for every edge $E \in E(H)$. Many extremal problems from combinatorics and geometry can be reformulated as questions of the following type: What is the size of a smallest transversal in a given hypergraph H ? This problem, in general, is known to be computationally intractable (cf. Garey and Johnson, 1979). However, under certain specific conditions on H , one can guarantee the existence of a relatively small transversal. The present chapter focuses on results of this kind. In particular, we shall see how a powerful probabilistic idea of Vapnik and Chervonenkis can be applied to obtain a number of interesting geometric and algorithmic results.

TRANSVERSALS AND FRACTIONAL TRANSVERSALS

Let H be a hypergraph with vertex set $V(H)$ and edge set $E(H)$. Let $\tau(H)$ denote the size of a smallest transversal of H , that is, the smallest number τ such that one can choose τ vertices with the property that any edge of H contains at least one of them. $\tau(H)$ is usually called the *transversal number* (or the *vertex-cover number*) of H .

The *packing number* (or *matching number*) of a hypergraph H is defined as the largest number $\nu = \nu(H)$ such that H has ν pairwise disjoint hyperedges. Obviously, $\nu(H) \leq \tau(H)$ for any hypergraph H . Typically, $\tau(H)$ is strictly larger

than $\nu(H)$. In fact, $\tau(H)$ cannot even be bounded by any function of $\nu(H)$ (see Exercise 15.3).

Let \mathbb{R}^+ denote the set of all nonnegative real numbers. Let us call a function $t: V(H) \rightarrow \mathbb{R}^+$ a *fractional transversal* of H if

$$\sum_{x \in E} t(x) \geq 1 \quad \text{for every hyperedge } E \in E(H). \quad (15.1)$$

The minimum of $\sum_{x \in V(H)} t(x)$ over all fractional transversals of H is called the *fractional transversal number* of H , and is denoted by $\tau^*(H)$. One can associate with each transversal T of H a function $t_T: V(H) \rightarrow \mathbb{R}^+$ defined as

$$t_T(x) = \begin{cases} 1 & \text{if } x \in T, \\ 0 & \text{if } x \notin T. \end{cases}$$

Since this function satisfies (15.1) and $\sum_{x \in V(H)} t_T(x) = |T|$, we have that $\tau^*(H) \leq \tau(H)$.

Similarly, a *fractional packing* of H is a nonnegative function $p: E(H) \rightarrow \mathbb{R}^+$ such that

$$\sum_{x \in E} p(E) \leq 1 \quad \text{for every vertex } x \in V(H).$$

The maximum of $\sum_{E \in E(H)} p(E)$ over all fractional packings of H is called the *fractional packing number* of H , and is denoted by $\nu^*(H)$. As before, we have $\nu^*(H) \geq \nu(H)$.

It is easy to deduce directly from the definition that $\nu^*(H) \leq \tau^*(H)$ (see Exercise 15.1). In fact, these two numbers are always equal to each other. Moreover, the following is true.

Theorem 15.1. *For every hypergraph H ,*

$$\nu(H) \leq \nu^*(H) = \tau^*(H) \leq \tau(H),$$

and the value of $\nu^(H) = \tau^*(H)$ can be determined by linear programming.*

Proof. Let x_i ($1 \leq i \leq n$) and E_j ($1 \leq j \leq m$) be the vertices and the edges of H , respectively. Let $A = (a_{ij})$ be the *incidence matrix* of H , i.e.,

$$a_{ij} = \begin{cases} 1 & \text{if } x_i \in E_j, \\ 0 & \text{if } x_i \notin E_j. \end{cases}$$

Let A^T denote the transpose of A , and let $\mathbf{1}_n$ denote the matrix consisting of one column of length n , all of whose entries are 1's. Given a function $t: V(H) \rightarrow \mathbb{R}$ (and $p: E(H) \rightarrow \mathbb{R}$), let \underline{t} (resp. \underline{p}) denote a matrix consisting of one column whose i th entry is $t(x_i)$ (resp. $p(E_j)$).

Observe that t is a fractional transversal of H if and only if

$$A^T \underline{t} \geq \mathbf{1}_m \quad \text{and} \quad \underline{t} \geq \mathbf{0}.$$

Similarly, p is a fractional packing of H if and only if

$$A \underline{p} \leq \mathbf{1}_n \quad \text{and} \quad \underline{p} \geq \mathbf{0}.$$

Thus,

$$\begin{aligned} \tau^*(H) &= \min \{ \mathbf{1}_n^T \underline{t} \mid A^T \underline{t} \geq \mathbf{1}_m, \underline{t} \geq \mathbf{0} \}, \\ \nu^*(H) &= \max \{ \mathbf{1}_m^T \underline{p} \mid A \underline{p} \leq \mathbf{1}_n, \underline{p} \geq \mathbf{0} \}. \end{aligned}$$

These two linear programming problems are dual to each other, so it follows immediately from the duality theorem of linear programming that their solutions, $\tau^*(H)$ and $\nu^*(H)$, are equal (see, e.g., Papadimitriou and Steiglitz, 1982; Chvátal, 1983; and Grötschel et al., 1987). \square

In general, $\tau^*(H)$ can be much smaller than $\tau(H)$ (see Exercise 15.3). The following theorem of Lovász (1975) shows that this is not the case when every point of H belongs to relatively few hyperedges.

Theorem 15.2 (Lovász). *Let H be a hypergraph whose every vertex is contained in at most D edges. Then*

$$\tau^*(H) \leq \tau(H) \leq (\ln D + 1) \tau^*(H).$$

Proof. We have to prove only the second inequality. Let $t: V(H) \rightarrow \mathbb{R}^+$ be a fractional transversal of H with $\sum_{x \in V(H)} t(x) = \tau^*(H)$.

We are going to select a set of vertices x_1, x_2, \dots by a *greedy* algorithm. Let x_1 be any vertex of H whose degree (i.e., the number of edges containing it) is maximal. Let D_1 denote the degree of x_1 in H . Set $H_1 = H - x_1$, that is, the hypergraph obtained from H by deleting the vertex x_1 and all edges containing x_1 . If $x_1, \dots, x_i \in V(H)$ have already been selected, then let $H_i = H - x_1 - x_2 - \dots - x_i$. If H_i has no edges, we stop. Otherwise, let x_{i+1} be a vertex of H_i whose degree D_{i+1} is maximal, and so on. Clearly,

$$|E(H_i)| - |E(H_{i+1})| = D_{i+1}. \quad (15.2)$$

By the properties of t ,

$$\begin{aligned}
|E(H_i)| &= \sum_{E \in E(H_i)} 1 \leq \sum_{E \in E(H_i)} \sum_{x \in E} t(x) \\
&= \sum_{x \in V(H_i)} t(x) \sum_{\substack{E \in E(H_i) \\ E \ni x}} 1 \\
&\leq \sum_{x \in V(H_i)} t(x) D_{i+1} \\
&\leq D_{i+1} \tau^*(H).
\end{aligned}$$

Assume now that our procedure terminates in s steps; i.e., H_s is empty. Then, of course, $\tau(H) \leq s$. Put $H_0 = H$. By (15.2), we have

$$\begin{aligned}
s &= \sum_{i=0}^{s-1} 1 = \sum_{i=0}^{s-1} \frac{|E(H_i)| - |E(H_{i+1})|}{D_{i+1}} \\
&= \frac{|E(H)|}{D_1} + \sum_{i=1}^{s-1} |E(H_i)| \left(\frac{1}{D_{i+1}} - \frac{1}{D_i} \right).
\end{aligned}$$

Hence, using the inequality $|E(H_i)| \leq D_{i+1} \tau^*(H)$ ($0 \leq i \leq s$) and the fact $D_s \geq 1$, we obtain

$$\begin{aligned}
s &\leq \tau^*(H) + \sum_{i=1}^{s-1} D_{i+1} \tau^*(H) \left(\frac{1}{D_{i+1}} - \frac{1}{D_i} \right) \\
&= \tau^*(H) \left(1 + \sum_{i=1}^{s-1} \frac{D_i - D_{i+1}}{D_i} \right) \\
&\leq \tau^*(H) \left(1 + \sum_{i=1}^{s-1} \sum_{k=D_{i+1}+1}^{D_i} \frac{1}{k} \right) \\
&= \tau^*(H) \left(1 + \sum_{k=D_s+1}^{D_1} \frac{1}{k} \right) \\
&\leq \tau^*(H) (1 + \ln D_1).
\end{aligned}$$

Thus,

$$\tau(H) \leq s \leq \tau^*(H) (1 + \ln D_1),$$

as desired. □

VAPNIK–CHERVONENKIS DIMENSION

Suppose that for a public opinion poll we want to select a small number of individuals representing all major sections of the society. First, we have to choose certain categories of people and then decide which of these groups are considered “important.” According to our democratic principles, we shall measure the “importance” of a group by its size (in the percentage of the population). Then the important groups will define a hypergraph H with the property that $|E| \geq \varepsilon|V(H)|$ for every edge $E \in E(H)$, where ε is some fixed constant ($0 < \varepsilon < 1$). The smallest number of people representing all important groups is $\tau(H)$.

Clearly, the function $t(x) = 1/(\varepsilon|V(H)|)$, for all $x \in V(H)$, is a fractional transversal of H with $\sum_{x \in V(H)} t(x) = 1/\varepsilon$. Hence, $\tau^*(H) \leq 1/\varepsilon$, and Theorem 15.2 implies that

$$\tau(H) \leq \frac{1}{\varepsilon}(\ln D + 1), \quad (15.3)$$

where D is the maximum degree of the vertices of H . This bound is extremely poor if D is large.

In their seminal paper, Vapnik and Chervonenkis (1971) pointed out that if H satisfies certain natural conditions, the above upper bound can be replaced by a function depending only on ε . To specify these conditions, we need some preparation.

Definition 15.3. Let $H = (V(H), E(H))$ denote a hypergraph. A subset $A \subseteq V(H)$ is called *shattered* if for every $B \subseteq A$ there exists an $E \in E(H)$ such that $E \cap A = B$. The *Vapnik–Chervonenkis dimension* (or *VC dimension*) of H is the cardinality of the largest shattered subset of $V(H)$. It will be denoted by $\text{VC-dim}(H)$.

The following theorem was proved independently by Shelah (1972), Sauer (1972), and Vapnik and Chervonenkis (1971).

Theorem 15.4. *Let H be a hypergraph with n vertices and VC-dimension d . Then*

$$|E(H)| \leq \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{d},$$

and this bound cannot be improved.

First Proof. The assertion is trivial if $d = 0$ or $n \leq d$. Assume that we have already proved it for every hypergraph \bar{H} with $\text{VC-dim}(\bar{H}) < d$, and for every hypergraph \bar{H} with $\text{VC-dim}(\bar{H}) = d$ and $|V(\bar{H})| < n$.

Given a hypergraph H with n vertices and VC-dimension d , let us define two

other hypergraphs, H_1 and H_2 , as follows. Let $V(H_1) = V(H_2) = V(H) - \{x\}$ for some fixed $x \in V(H)$, and set

$$E(H_1) = \{E - \{x\} \mid E \in E(H)\},$$

$$E(H_2) = \{E \in E(H) \mid x \notin E \text{ and } E \cup \{x\} \in E(H)\}.$$

Obviously, $\text{VC-dim}(H_1) \leq d$ and $\text{VC-dim}(H_2) \leq d - 1$.

On the other hand, by the induction hypothesis,

$$\begin{aligned} |E(H)| &= |E(H_1)| + |E(H_2)| \\ &\leq \sum_{i=0}^d \binom{n-1}{i} + \sum_{i=0}^{d-1} \binom{n-1}{i} \\ &= \sum_{i=0}^d \binom{n}{i}. \end{aligned}$$

The tightness of this bound follows from the fact that if $E(H) = \{U \subseteq V \mid |U| \leq d\}$, then $\text{VC-dim}(H) = d$. □

We also include a slightly more complicated proof due to Frankl and Pach (1983), because it is a good illustration of the so-called *linear algebra method* (see, e.g., Babai and Frankl, 1988).

Second Proof. Let $E(H) = \{E_i \mid 1 \leq i \leq m\}$, and let $X_j, 1 \leq j \leq \sum_{i=0}^d \binom{n}{i}$, be a list of all subsets of $V(H)$ of size at most d . Define an $m \times \sum_{i=0}^d \binom{n}{i}$ matrix $\mathbf{A} = (a_{ij})$ by

$$a_{ij} = \begin{cases} 1 & \text{if } E_i \supseteq X_j, \\ 0 & \text{if } E_i \not\supseteq X_j. \end{cases}$$

Suppose, for contradiction, that $m > \sum_{i=0}^d \binom{n}{i}$. Then the rows of \mathbf{A} are linearly dependent over the reals; thus there exists a nonzero function $f: E(H) \rightarrow \mathbb{R}$ such that

$$\sum_{E_i \supseteq X_j} f(E_i) = 0 \quad \text{for every } X_j.$$

Let $A \subseteq V(H)$ be a *minimal* subset for which

$$\sum_{E_i \supseteq A} f(E_i) = \alpha \neq 0.$$

(Sets A with nonzero sums certainly exist, for we get a nonzero sum for any maximal element A of the family $\{A \in E(H) \mid f(A) \neq 0\}$.) Obviously, $|A| \geq d + 1$. Given any $B \subseteq A$, let

$$F(B) = \sum_{E_i \cap A = B} f(E_i).$$

Thus, $F(A) = \alpha$, and setting $B = A - \{a\}$ for any fixed $a \in A$,

$$\begin{aligned} F(B) &= \sum_{E_i \supseteq B} f(E_i) - \sum_{E_i \supseteq A} f(E_i) \\ &= 0 - \alpha = -\alpha \end{aligned}$$

In general, if B is any $(|A| - k)$ -element subset of A ($0 \leq k \leq |A|$), then

$$F(B) = (-1)^k \alpha \neq 0.$$

This yields, in particular, that there exists at least one hyperedge E_i with $E_i \cap A = B$. Thus, A is shattered, contradicting our assumption that $\text{VC-dim}(H) = d$. \square

Vapnik and Chervonenkis (1971) discovered an ingenious probabilistic (counting) argument based on the above result, which leads to a substantial improvement of the bound (15.3). They showed (in a somewhat different setting) that there exists a function $f(d, \varepsilon)$ such that the transversal number of every hypergraph H of VC-dimension d , all of whose edges have at least $\varepsilon|V(H)|$ elements, is at most $f(d, \varepsilon)$ (see Exercise 15.6). The ideas of Vapnik and Chervonenkis have been adapted by Haussler and Welzl (1987) and Blumer et al. (1989) to obtain various upper bounds on $f(d, \varepsilon)$. These results were sharpened and generalized by Komlós, Pach, and Woeginger (1992), as follows. Given a finite set V , a function $\mu: V \rightarrow \mathbb{R}^+$ is called a *probability measure* if

$$\sum_{x \in V} \mu(x) = 1.$$

The measure of any subset $X \subseteq V$ is defined by $\mu(X) = \sum_{x \in X} \mu(x)$.

Theorem 15.5 (Komlós et al.). *Let H be a hypergraph of VC-dimension d , let $\varepsilon > 0$, and let μ be a probability measure on $V(H)$ such that $\mu(E) \geq \varepsilon$ for every $E \in E(H)$. Then $\tau(H) \leq t(d, \varepsilon)$, where $t(d, \varepsilon)$ denotes the smallest positive integer t satisfying*

$$2 \sum_{i=0}^d \binom{T}{i} \left(1 - \frac{t}{T}\right)^{(T-i)\varepsilon-1} < 1$$

for some integer $T > t$. Consequently, for any $\varepsilon \leq \frac{1}{2}$, we have

$$\tau(H) \leq \frac{d}{\varepsilon} \left(\ln \frac{1}{\varepsilon} + 2 \ln \ln \frac{1}{\varepsilon} + 6 \right)$$

(cf. Exercise 15.9).

Proof. Let us select with possible repetition t random points of $V(H)$, where the selections are done with respect to the probability measure μ . We get a *random sample*

$$x \in [V(H)]^t = \underbrace{V(H) \times \cdots \times V(H)}_{t \text{ times}}.$$

We say that x is a *transversal* of H if every edge $E \in E(H)$ contains at least one point of x . Let $I(E, x)$ denote the number of components of x that belong to E , counting with multiplicity. Then

$$\Pr[x \text{ is not a transversal of } H] = \Pr[\exists E \in E(H) : I(E, x) = 0].$$

Having picked the string x of length t , let us choose randomly another $T - t$ elements from $V(H)$. Let $y \in [V(H)]^{T-t}$ denote this new string, and let $z = xy \in [V(H)]^T$ stand for the full sequence. Furthermore, let $\langle z \rangle = \langle xy \rangle$ denote the *multiset* of all elements occurring in z (i.e., they are counted with multiplicities but their order is irrelevant).

For any $E \in E(H)$, $I(E, y)$ is a random variable having binomial distribution. Let m_E be the *median* of $I(E, y)$,

$$\Pr[I(E, y) > m_E] \leq \frac{1}{2} \leq \Pr[I(E, y) \geq m_E].$$

The following inequality is an immediate consequence of the independence of x and y .

$$\begin{aligned} & \Pr[\exists E \in E(H) : I(E, x) = 0] \\ & \leq \frac{\Pr[\exists E \in E(H) : I(E, x) = 0 \text{ and } I(E, y) \geq m_E]}{\min_{E \in E(H)} \Pr[I(E, y) \geq m_E]} \\ & \leq 2 \Pr[\exists E \in E(H) : I(E, x) \\ & = 0 \text{ and } I(E, y) \geq m_E]. \end{aligned}$$

For a fixed $E \in E(H)$, the conditional probability for given $\langle z \rangle = \langle xy \rangle$

$$\begin{aligned}
& \Pr [I(E, x) = 0 \text{ and } I(E, y) \geq m_E | \langle z \rangle] \\
&= \chi [I(E, z) \geq m_E] \frac{\binom{T-t}{I(E, z)}}{\binom{T}{I(E, z)}} \\
&\leq \chi [I(E, z) \geq m_E] \left(1 - \frac{t}{T}\right)^{I(E, z)} \\
&\leq \chi [I(E, z) \geq m_E] \left(1 - \frac{t}{T}\right)^{m_E}.
\end{aligned}$$

(Here $\chi [A]$ is the characteristic function of A , that is, $\chi [A] = 1$ if A is true, and 0 otherwise.)

By Theorem 15.4, a fixed multiset $\langle z \rangle$ has at most $\sum_{i=0}^d \binom{T}{i}$ different intersections with the edges of H . Thus,

$$\begin{aligned}
& \Pr [\exists E \in E(H) : I(E, x) = 0 \text{ and } I(E, y) \geq m_E | \langle z \rangle] \\
&\leq \left(\sum_{i=0}^d \binom{T}{i}\right) \left(1 - \frac{t}{T}\right)^m,
\end{aligned}$$

where $m = \min_{E \in E(H)} m_E$. Using the known fact that the median of a binomial distribution is within 1 of the mean,

$$m \geq (T - t) \min_{E \in E(H)} \mu(E) - 1 \geq (T - t)\epsilon - 1.$$

Hence, we obtain

$$\Pr [\exists E \in E(H) : I(E, x) = 0] \leq 2 \left(\sum_{i=0}^d \binom{T}{i}\right) \left(1 - \frac{t}{T}\right)^{(T-t)\epsilon-1}.$$

If the last expression is less than 1, then x is a transversal of H , with positive probability. This proves the first statement of the theorem. Choosing

$$\begin{aligned}
t &= \left\lfloor \frac{d}{\epsilon} \left(\ln \frac{1}{\epsilon} + 2 \ln \ln \frac{1}{\epsilon} + 6 \right) \right\rfloor, \\
T &= \left\lfloor \frac{\epsilon}{d} t^2 \right\rfloor,
\end{aligned}$$

we get after some calculations that

$$2 \sum_{i=0}^d \binom{T}{i} \left(1 - \frac{t}{T}\right)^{(T-t)\epsilon-1} < 1,$$

provided that $\epsilon \leq \frac{1}{2}$. □

The above theorem is valid for any probability measure μ defined on the vertex set of H . In particular, one can choose μ to be constant; that is, $\mu(x) = 1/|V(H)|$ for every $x \in V(H)$. We can deduce another interesting result from Theorem 15.5 by applying it to the measure $\mu'(x) = t(x)/\tau^*(H)$, where $t: V(H) \rightarrow \mathbb{R}^+$ is a fractional transversal of H with $\sum_{x \in V(H)} t(x) = \tau^*(H)$. Observe that in this case

$$\begin{aligned} \mu'(E) &= \sum_{x \in E} \mu'(x) \\ &= \sum_{x \in E} \frac{t(x)}{\tau^*(H)} \geq \frac{1}{\tau^*(H)} \end{aligned}$$

holds for every $E \in E(H)$. Thus, choosing $\varepsilon = 1/\tau^*(H)$ in Theorem 15.5, we obtain the following.

Corollary 15.6 (Koslós et al.). *Let H be any hypergraph of VC-dimension d .*

(i) *If every edge of H has at least $\varepsilon|V(H)|$ elements for some $\varepsilon \leq \frac{1}{2}$, then*

$$\tau(H) \leq \frac{d}{\varepsilon} \left(\ln \frac{1}{\varepsilon} + 2 \ln \ln \frac{1}{\varepsilon} + 6 \right).$$

(ii) *If $\tau^*(H) \geq 2$, then*

$$\tau(H) \leq d\tau^*(H) (\ln \tau^*(H) + 2 \ln \ln \tau^*(H) + 6).$$

Next we show that for $d \geq 2$, the bound given in Theorem 15.5 is close to being optimal.

Theorem 15.7 (Koslós et al.). *Given any natural number $d \geq 2$ and any real $\gamma < 2/(d+2)$, there exists a constant $\varepsilon_{d,\gamma} > 0$ with the following property.*

For any $\varepsilon \leq \varepsilon_{d,\gamma}$, one can construct a hypergraph H of VC-dimension d , all of whose edges have at least $\varepsilon|V(H)|$ points, and

$$\tau(H) \geq (d-2+\gamma) \frac{1}{\varepsilon} \ln \frac{1}{\varepsilon}.$$

Proof. Again, we use the probabilistic method. Let γ' be a fixed constant, $\gamma < \gamma' < 2/(d+2)$. Given a sufficiently small ε , let $n = (K/\varepsilon) \ln(1/\varepsilon)$, where K is a constant depending only on d , γ , and γ' (but not on ε), which will be specified later. Furthermore, let

$$r = \varepsilon n, \quad p = \frac{\varepsilon^{1-d-\gamma'}}{\binom{n}{r}}, \quad t = (d-2+\gamma) \frac{1}{\varepsilon} \ln \frac{1}{\varepsilon}.$$

We assume that n , r , t are integers, disregarding all roundoff errors.

Let V be a fixed n -element set. Construct a hypergraph H on the vertex set V by randomly selecting some r -element subsets of V , where each r -tuple is chosen independently with probability p . We are going to show that with high probability

(i) $\text{VC-dim}(H) \leq d$, and

(ii) $\tau(H) > t$.

$\Pr[\text{VC-dim}(H) > d]$

$$\begin{aligned}
&\leq \binom{n}{d+1} \Pr[\text{a fixed } (d+1)\text{-element subset } A \subseteq V \text{ is shattered by } H] \\
&= \binom{n}{d+1} \prod_{B \subseteq A} \Pr[\exists E \in E(H) : E \cap A = B] \\
&= \binom{n}{d+1} \prod_{B \subseteq A} \left(1 - (1-p)^{\binom{n-|A|}{|B|}}\right) \\
&= \binom{n}{d+1} \prod_{j=0}^{d+1} \left(1 - (1-p)^{\binom{n-d-1}{r-j}}\right)^{\binom{d+1}{j}} \\
&= \binom{n}{d+1} \prod_{i=0}^{d+1} \left(1 - (1-p)^{\binom{n-d-1}{r-d-1+i}}\right)^{\binom{d+1}{d+1-i}} \\
&= \binom{n}{d+1} \prod_{i=0}^{d+1} \left(1 - (1-p)^{\binom{n-d-1}{r-d-1+i}}\right)^{\binom{d+1}{i}} \\
&\leq \binom{n}{d+1} \prod_{i=0}^{d+1} \left(p \binom{n-d-1}{r-d-1+i}\right)^{\binom{d+1}{i}} \\
&= \binom{n}{d+1} p \binom{n-d-1}{r-d-1} \left(p \binom{n-d-1}{r-d}\right)^{d+1} \\
&\leq n^{d+1} p \binom{n}{r} \left(\frac{r}{n}\right)^{d+1} \left(p \binom{n}{r} \left(\frac{r}{n}\right)^d\right)^{d+1} \\
&= \left(K \ln \frac{1}{\varepsilon}\right)^{d+1} \varepsilon^{2-(d+2)\gamma'},
\end{aligned}$$

which tends to 0 as $\varepsilon \rightarrow 0$. This proves (i).

Next we show that (ii) also holds with high probability.

$$\begin{aligned} \Pr[\tau(H) \leq t] &= \binom{n}{t} (1-p)^{\binom{n-t}{r}} \\ &\leq \binom{n}{t} \exp\left[-p \binom{n-t}{r}\right] \\ &\leq \left(\frac{en}{t}\right)^t \exp\left[-p \binom{n}{r} \left(1 - \frac{r}{n-t+1}\right)^t\right]. \end{aligned}$$

From this, using the inequality $1 - ax > e^{-bx}$ for $b > a$, $0 < x < 1/a - 1/b$, we obtain the upper bound

$$\begin{aligned} &\left(\frac{en}{t}\right)^t \exp\left[-p \binom{n}{r} e^{-K\epsilon t/(K-d)}\right] \\ &= \left(\frac{eK}{d-2+\gamma}\right)^t \exp\left[-\epsilon^{1-d-\gamma'+K(d-2+\gamma)/(K-d)}\right], \end{aligned}$$

which tends to 0 if

$$1 - d - \gamma' + K(d - 2 + \gamma)/(K - d) < -1,$$

i.e., if K is sufficiently large. \square

The condition $d \geq 2$ in Theorem 15.7 is not merely a technical assumption. In fact, it is not hard to characterize all finite hypergraphs H with VC-dimension 1, and one can check that $\tau(H) \leq \lceil 1/\epsilon \rceil - 1$, provided that every edge of H has at least $\epsilon|V(H)|$ points for some $0 < \epsilon < 1$ (see Exercise 15.8).

The following simple assertion will help us in deciding whether a given hypergraph has low VC-dimension.

Lemma 15.8. *Let H be a hypergraph of VC-dimension d , and let $\varphi(E_1, \dots, E_k)$ be a set-theoretic formula of k variables (using $\cup, \cap, -$). If every edge E' of a hypergraph H' can be expressed as*

$$E' = \varphi(E_1, \dots, E_k) \quad \text{for suitable } E_i \in E(H),$$

then

$$\text{VC-dim}(H') \leq 2dk \log(2dk).$$

Proof. Let A be a d' -element subset of $V(H') = V(H)$, which is shattered in H' . By Theorem 15.4,

$$|\{E \cap A \mid E \in E(H)\}| \leq \sum_{i=0}^{d'} \binom{d'}{i}.$$

Using the assumption on H' , this yields

$$2^{d'} = |\{E' \cap A \mid E' \in E(H')\}| \leq \left(\sum_{i=0}^d \binom{d'}{i} \right)^k.$$

Comparing the two sides of this inequality, we obtain that $d' \leq 2dk \log(2dk)$, as required. \square

Ding, Seymour, and Winkler (1994) have introduced another parameter of a hypergraph, closely related to its VC-dimension. They defined $\lambda(H)$ as the largest integer l such that one can choose l edges $E_1, E_2, \dots, E_l \in E(H)$ with the property that for any $1 \leq i < j \leq l$, there is a vertex $x_{ij} \in E_i \cap E_j$ that does not belong to any other E_g ($g \neq i, j$). It is easy to see that $\text{VC-dim}(H) < \binom{\lambda(H) + 1}{2}$ for every hypergraph H . Combining Corollary 15.6 with Ramsey's theorem (Theorem 9.13), one can establish the following result.

Theorem 15.9 (Ding et al., 1994). *For any hypergraph H ,*

$$\tau(H) \leq 6\lambda^2(H) (\lambda(H) + \nu(H)) \binom{\lambda(H) + \nu(H)}{\lambda(H)}^2.$$

At the beginning of this chapter we pointed out that in general it is impossible to bound τ from above by any function of ν . Gyárfás and Lehel (1983, 1985) initiated the investigation of certain classes of hypergraphs for which such functions exist. Theorem 15.9 provides a sufficient condition for a family of hypergraphs to have this property. It implies that if there exists a constant K such that $\lambda(H) \leq K$ for all members of a family, then τ can be bounded from above by a polynomial of ν . For various geometric consequences of this fact, see Pach (1995).

RANGE SPACES AND ε -NETS

Haussler and Welzl (1987) were the first to recognize the relevance of the above machinery to geometric problems, and in fact they formulated and proved the first version of Theorem 15.5, too. It seems to capture the essence of the so-called *random* (or *probabilistic*) method in a large variety of geometric applications. This ready-to-use kit will save us a lot of time (and space) in situations where otherwise we would go through lengthy but routine calculations. However, the main significance of these ideas is that they shed some light on the general transversal problem. The transversal number is a *global* parameter of a set system. The results in the preceding section show that in any measure space of total measure 1, any system of large measurable sets admits a relatively small transversal, provided that its *local* behavior is nice (i.e., its VC-dimension is bounded).

In recent years, the above concepts and methods have found many interesting applications in discrete and computational geometry, learning theory, etc. (see Mulmuley, 1994; Anthony and Biggs, 1992). However, most of the literature in these fields is written in a different terminology. This language is so widely used that it is worth learning the basic definitions.

Definition 15.10. A *range space* is a (finite or infinite) hypergraph, i.e., a pair $\Sigma = (X, \mathcal{R})$, where X is an underlying set and \mathcal{R} is a family of subsets of X . (That is, $V(\Sigma) = X$ and $E(\Sigma) = \mathcal{R}$.) The elements of \mathcal{R} (i.e., the hyperedges of Σ) are called *ranges*.

Given any $Y \subseteq X$, we define the *subspace* of Σ induced by Y as

$$\Sigma_Y = (Y, \{R \cap Y \mid R \in \mathcal{R}\}).$$

If $\Sigma = (X, \mathcal{R})$ is a finite range space and $\varepsilon > 0$, then let

$$\Sigma^\varepsilon = (X, \{R \in \mathcal{R} \mid |R| \geq \varepsilon|X|\}).$$

A subset $T \subseteq X$ is called an ε -*net* for Σ if it intersects every element of $\mathcal{R}^\varepsilon = \{R \in \mathcal{R} \mid |R| \geq \varepsilon|X|\}$, that is, every “large” range of Σ .

The *Vapnik–Chervonenkis dimension* (VC-dimension) of a range space is defined in exactly the same way as for hypergraphs. That is,

$$\text{VC-dim}(\Sigma) = \max \{|A| \mid A \subseteq X \text{ and } \forall B \subseteq A \exists R \in \mathcal{R} \text{ such that } R \cap A = B\}.$$

Of course, $\text{VC-dim}(\Sigma_Y) \leq \text{VC-dim}(\Sigma)$ for any subspace Σ_Y of Σ .

Let $f(d, \varepsilon)$ denote the maximum size of the smallest ε -net of Σ over all finite range spaces Σ of VC-dimension at most d .

Theorems 15.5 to 15.7 can now be summarized as follows.

$$d - 2 + \frac{2}{d + 2} \leq \liminf_{\varepsilon \rightarrow 0} \frac{f(d, \varepsilon)}{\frac{1}{\varepsilon} \ln \frac{1}{\varepsilon}} \leq \limsup_{\varepsilon \rightarrow 0} \frac{f(d, \varepsilon)}{\frac{1}{\varepsilon} \ln \frac{1}{\varepsilon}} \leq d.$$

Next we give some important examples of range spaces whose VC-dimensions are finite (see also Exercise 15.11).

Example 15.11

- (a) $X = \mathbb{R}^d$, $\mathcal{R} = \{\text{all half-spaces in } \mathbb{R}^d\}$;
- (b) $X = \mathbb{R}^d$, $\mathcal{R} = \{\text{all open (unit) balls in } \mathbb{R}^d\}$;
- (c) $X = \{\text{all hyperplanes in } \mathbb{R}^d\}$. Given any open segment s in \mathbb{R}^d , let R_s be the collection of all hyperplanes intersecting s . Set $\mathcal{R} = \{R_s \mid \text{for all open segments } s\}$.

It is easy to check that the VC-dimension of each of the above range spaces is finite. We demonstrate this for (c). Assume, in order to obtain a contradiction,

that for any large n we can find an n -element subset $A_n \subset X$ which is shattered, i.e.,

$$|\{R_s \cap A_n \mid R_s \in \mathcal{R}\}| = 2^n.$$

The hyperplanes belonging to A_n define a cell complex consisting of at most $O(n^d)$ cells. Observe that if $s = p_1 p_2$ and $t = q_1 q_2$ are two segments such that p_i, q_i belong to the same cell ($i = 1, 2$), then $R_s \cap A_n = R_t \cap A_n$. Since there are $O(n^{2d})$ possibilities for selecting two cells for the endpoints of s , we obtain

$$|\{R_s \cap A_n \mid R_s \in \mathcal{R}\}| = O(n^{2d}),$$

a contradiction.

Lemma 15.8 allows us to construct from (a), (b), (c) a number of additional examples of range spaces of finite VC-dimension. It implies that all range spaces whose ranges can be obtained from the ranges of a space Σ of finite dimension by using $\cap, \cup, -$ a bounded number of times, have finite VC-dimensions. In particular, this is true for the following range spaces.

- (d) $X = \mathbb{R}^d$, $\mathcal{R} = \{\text{all open polyhedra in } \mathbb{R}^d \text{ with at most } k \text{ vertices}\}$;
- (e) $X = \mathbb{R}^d$, $\mathcal{R} = \{\text{all annuli in } \mathbb{R}^d\}$;
- (f) $X = \{\text{all hyperplanes in } \mathbb{R}^d\}$. Given any k -dimensional simplex $S_k \subset \mathbb{R}^d$, let R_{S_k} be the collection of all hyperplanes intersecting the interior of S_k . Set $\mathcal{R} = \{R_{S_k} \mid \text{for all simplices } S_k\}$.

A typical example of a geometric range space whose VC-dimension is not finite is the following.

- (g) $X = \mathbb{R}^2$, $\mathcal{R} = \{\text{all open convex sets in } \mathbb{R}^2\}$

(see Exercise 15.18(i)).

SPANNING TREES OF LOW STABBING NUMBER

We are going to illustrate the power of the concepts and results described above by proving a beautiful theorem of Welzl (1988) (see also Chazelle and Welzl, 1989; and Welzl, 1992), and by indicating its algorithmic consequences.

Let S be a set of n points in the plane. A tree T whose vertices are the points of S and whose edges are straight-line segments is called a *spanning tree* of S . The *stabbing number* of a spanning tree T , denoted by $\sigma(T)$, is defined as the maximum number of segments in T that can be intersected by a single line.

Theorem 15.12 (Welzl). *For any set S of n points in \mathbb{R}^2 in general position, there is a spanning tree T whose stabbing number is at most $c\sqrt{n} \log n$, where c is an appropriate constant.*

The proof is based on the following selection lemma.

Lemma 15.13. *Given a set S of $n \geq 2$ points in \mathbb{R}^2 in general position, and a multiset L of m lines avoiding these points, one can always find a pair $x, y \in S$ such that the segment connecting them intersects at most $c'(m/\sqrt{n}) \log n$ lines of L , where c' is a suitable constant.*

Proof. One can assume without loss of generality that L is a set (i.e., it contains no repeated elements); otherwise, one can slightly change the position of some lines. Let us consider the subspace of the range space in Example 15.11(c) restricted to L ($d = 2$). That is, let

$$\Sigma_L = (L, \{R_s \cap L \mid \text{for all open segments } s\}),$$

where $R_s \cap L$ is the set of all lines of L intersected by a given segment s .

As we have pointed out before, Σ_L has finite VC-dimension D . Hence, by Corollary 15.6(i), there exists a constant $c_D \geq 1$ such that Σ_L has an ε -net of size at most $c_D(1/\varepsilon) \log(1/\varepsilon)$ for any $0 < \varepsilon < 1$.

Setting $\varepsilon = c_D(\log n/\sqrt{n})$, we find that there exists a subset $L' \subseteq L$ with

$$|L'| \leq c_D \frac{1}{\varepsilon} \log \frac{1}{\varepsilon} < \frac{\sqrt{n}}{2},$$

and every open segment crossing at least $\varepsilon m = c_D(m/\sqrt{n}) \log n$ elements of L intersects at least one line belonging to L' .

The lines of L' divide the plane into fewer than n cells. By the pigeonhole principle, at least one of them contains two points $x, y \in S$. The segment connecting them avoids every line of L' ; therefore, it cannot intersect more than $c_D(m/\sqrt{n}) \log n$ elements of L . \square

Proof of Theorem 15.12. Fix a set $S = S_0$ of $n = n_0$ points in the plane in general position and a set L_0 consisting of $m_0 = \binom{n}{2} + 1$ lines avoiding these points such that they represent every possible partition of S into two parts by a line (see Exercise 15.10(iii)).

Applying Lemma 15.13 to the pair (S_0, L_0) , we can find two points $x_0, y_0 \in S_0$ with the property that the segment connecting them intersects at most $c'(m_0/\sqrt{n_0}) \log n_0$ elements of L_0 . Let $S_1 = S_0 - \{x_0\}$, $n_1 = |S_1| = n - 1$, and L_1 be the multiset of lines obtained from L_0 by duplicating each line intersected by the segment $x_0 y_0$. Set $m_1 = |L_1|$. By Lemma 15.13, now there exists a pair of points $x_1, y_1 \in S_1$ such that the segment $x_1 y_1$ intersects at most $c'(m_1/\sqrt{n_1}) \log n_1$ lines of L_1 .

Let $S_2 = S_1 - \{x_1\}$, $n_2 = |S_2| = n - 2$, and let L_2 be the multiset of lines

obtained from L_1 by duplicating all lines intersected by the segment $x_1 y_1$. Set $m_2 = |L_2|$.

Proceeding in this way, for every $i \leq n - 1$, we can recursively define an $n_i = (n - i)$ -element set $S_i \subseteq S$ and an m_i -element multiset of lines L_i . It follows from the construction that

$$m_{i+1} \leq m_i \left(1 + c' \frac{\log n_i}{\sqrt{n_i}} \right) \quad \text{for every } 0 \leq i \leq n - 2.$$

Hence,

$$\begin{aligned} m_{n-1} &\leq m_0 \prod_{i=0}^{n-2} \left(1 + c' \frac{\log n_i}{\sqrt{n_i}} \right) \\ &\leq \left(\binom{n}{2} + 1 \right) \exp \left(\sum_{i=0}^{n-2} c' \frac{\log n_i}{\sqrt{n_i}} \right) \\ &\leq \left(\binom{n}{2} + 1 \right) \exp \left(c' \log n \sum_{i=0}^{n-2} \frac{1}{\sqrt{n-i}} \right) \\ &\leq 2^c \sqrt{n} \log n, \end{aligned}$$

for a suitable constant $c > 0$.

Observe that the segments $x_0 y_0, \dots, x_{n-2} y_{n-2}$ form a spanning tree T of S . Assume that a line $\ell \in L_0$ intersects t edges of T . Then L_{n-1} contains 2^t copies of ℓ , and it follows from the last inequality that $t \leq c \sqrt{n} \log n$.

In view of the fact that the elements of L_0 exhaust all possible ways in which S can be bisected by a straight line, we can conclude that no line can cross more than $c \sqrt{n} \log n$ edges of T . \square

The statement of Theorem 15.12 can be strengthened slightly as follows.

Corollary 15.14. *For any set S of n points in the plane in general position, there is a spanning path P whose stabbing number is at most $C \sqrt{n} \log n$, where C is a suitable constant. Moreover, we can assume that P is a simple (non-self-intersecting) polygon.*

Proof. Let T be a spanning tree of S such that $\sigma(T) \leq c \sqrt{n} \log n$. Let $x_1, x_2, \dots, x_{2n-2} = x_1$ be a sequence of points of S (with repetitions) such that

- (i) $\{x_i, x_{i+1}\}$ is an edge of T , for $i < 2n - 2$,
- (ii) for every edge e of T , $e = \{x_i, x_{i+1}\}$ for exactly two different values of i .

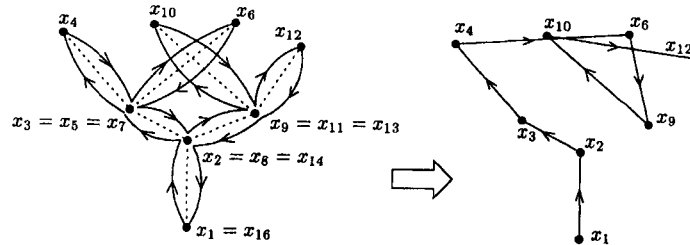


Figure 15.1. Converting a spanning tree to a spanning path.

Erase from this sequence every element x_j such that $x_j = x_i$ for some $i < j$ (see Figure 15.1). The resulting subsequence $x_{i_1} = x_1, x_{i_2}, \dots, x_{i_n}$ contains every point of S exactly once, and it is easy to check that it induces a path P whose stabbing number is at most $2\sigma(T)$.

To prove the second statement, observe that if two edges of P cross each other, they are the diagonals of a convex quadrilateral Q . Replacing these two edges of P by two opposite sides of Q , we can obtain another spanning path P' with $\sigma(P') \leq \sigma(P)$. Since the total length of the edges of P' is strictly shorter than that of P , this algorithm will terminate in a finite number of steps. \square

The bounds obtained in Theorem 15.12 and Corollary 15.14 can be improved by a factor of $\log n$ using a variant of Theorem 11.6 (see Exercise 15.13).

RANGE SEARCHING

The last result has some important algorithmic consequences. For example, it can be used to design an efficient algorithm for the *half-plane range searching* problem. In what follows, we briefly describe this problem and outline a solution.

Let S be a fixed set of n points in \mathbb{R}^2 in general position. We wish to build a linear-size *data structure* that allows us to answer efficiently questions (*queries*) of the following type: Given an open half-plane h , what is the number of points of S lying in h ?

Let $P = x_1x_2 \cdots x_n$ be a spanning path of S with stabbing number $\sigma(P) = O(\sqrt{n} \log n)$. Define a rooted binary tree T_P , as follows. Let the root r of T_P be associated with the entire path $x_1 \cdots x_n$. Assume that we have already defined a vertex v of T_P , and it was associated with the subpath $P(v) = x_i \cdots x_j$. If $j = i$, then v is a leaf of T_P . Otherwise, let v have two children v_1 and v_2 associated with the subpaths

$$P(v_1) = x_i \cdots x_{\lfloor \frac{i+j}{2} \rfloor}, P(v_2) = x_{\lfloor \frac{i+j}{2} \rfloor + 1} \cdots x_j.$$

Thus, T_P has exactly n leaves.

After some processing, we can store at each vertex $v \in T_P$

- (i) $|P(v)|$, the size of $P(v)$, and

- (ii) the (coordinates of the) vertices of the convex hull of $P(v)$ in circular order.

Given an open query half-plane h , we can determine $|S \cap h|$ as follows. We visit the root r of T_P , and check whether the boundary of h intersects $P(r)$. If it does, we pass down the question to the children of r . In general, when we visit a vertex v of T_P , we decide whether the boundary of h intersects $P(v)$. If it does not, we stop; otherwise, we visit the children of v . The vertices of T_P that have been visited during this process form a subtree $T'_P \subseteq T_P$. At each leaf v of T'_P we can decide whether $P(v) \subset h$ or $P(v) \cap h = \emptyset$. Starting at the leaves of T'_P and processing backward, we can easily compute the sum of $|P(v)|$ over all leaves for which $P(v) \subset h$. Obviously, this sum gives the number of points of S lying in h .

Notice that at each level of T_P , the number of vertices v for which h intersects $P(v)$ is at most the stabbing number of P . Since the number of levels in T_P is at most $\lceil \log_2 n \rceil + 1$, it is clear that for any query half-plane the number of vertices visited is

$$|V(T'_P)| \leq 3 \lceil \log_2 n \rceil \sigma(P) = O(\sqrt{n} \log^2 n).$$

On the other hand, at each vertex $v \in T'_P$, the time spent is proportional to the time required to detect an intersection between a line ℓ and P_v . However, it is easy to see that ℓ intersects P_v if and only if ℓ intersects the convex hull of P_v , and that the intersection between a line and convex n -gon can be detected in $O(\log n)$ time (see Exercise 15.14).

Summarizing, we have obtained the following:

Corollary 15.15. *Given a set S of n points in the plane in general position, there exists a data structure using $O(n \log n)$ space that allows us to compute $|S \cap h|$ for any query half-plane in time $O(\sqrt{n} \log^3 n)$.*

Spanning trees with low stabbing number have been used to obtain fast algorithms for several problems in computational geometry. See Agarwal (1992), Edelsbrunner et al. (1989), and Welzl (1992) for related work.

Finally, we note that the following statement, which is a slightly weaker version of Theorem 11.6, can also be established using ε -nets (see Exercise 15.20).

Corollary 15.16. *Given a set L of n lines in general position and an integer $r \geq 3$, the plane can be decomposed into at most r^2 triangles (some of which are unbounded) such that the interior of each triangle intersects at most $c(n/r) \log r$ lines of L , where c is a suitable constant.*

EXERCISES

- 15.1 Prove (without using Theorem 15.1 or the duality theorem of linear programming) that $\nu^*(H) \leq \tau^*(H)$ for every hypergraph H .
- 15.2 (König, 1936) Prove that if G is a bipartite graph, then $\nu(G) = \tau(G)$.
- 15.3 Prove that there exists no function $f: \mathbb{R} \rightarrow \mathbb{R}$ such that $\tau(H) \leq f(\tau^*(H))$ for every hypergraph H .
- 15.4 (Erdős, 1964b; Lovász, 1975)
- (i) Prove that every r -uniform hypergraph H with at most 2^{r-1} edges is two-colorable; i.e., one can color its vertices by two colors so that every edge of H contains points of both colors.
 - (ii) Prove that there exists a constant $c > 0$ such that for every r , one can find a non-two-colorable r -uniform hypergraph with at most $cr^2 2^r$ edges.
- 15.5 (Frankl and Pach, 1984) Let H be an r -uniform hypergraph with n vertices and VC-dimension d . Show that $|E(H)| \leq \binom{n}{d}$.
- 15.6 (Vapnik and Chervonenkis, 1971) Given $\varepsilon > 0$ and a hypergraph H with vertex set $V = V(H)$ we say that a subset $A \subseteq V$ is an ε -approximation of H if

$$\left| \frac{|A \cap E|}{|A|} - \frac{|E|}{|V|} \right| \leq \varepsilon \quad \text{for every } E \in E(H);$$

i.e., the relative number of points of A in any edge approximates the relative size of this edge with accuracy ε .

- (i)* Show that there is an absolute constant C such that for any $0 < \varepsilon < 1$, any hypergraph of VC-dimension d has an ε -approximation of at most $C(d/\varepsilon^2) \log(d/\varepsilon)$ elements.
 - (ii) Show that if every edge of a hypergraph H has more than $\varepsilon|V(H)|$ points (for some $\varepsilon > 0$), then every ε -approximation A of H intersects all edges of H ; i.e., A is a transversal of H .
- 15.7* Decide whether the following statement is true. Let H be a hypergraph of VC-dimension d , all of whose edges have at least $\varepsilon|V(H)|$ points for some $\varepsilon > d/|V(H)|$. Then there exists an $\lceil \varepsilon|V(H)| \rceil$ -uniform hypergraph H' of VC-dimension d , with the property that for any $E \in E(H)$ one can find $E' \in E(H)$ such that $E' \subseteq E$.
- 15.8 (Pach and Woeginger, 1990; Komlós et al., 1992) Let H be a hypergraph of VC-dimension 1, all of whose edges have at least $\varepsilon|V(H)|$

points for some $0 < \varepsilon < 1$. Show that

$$\tau(H) \leq \max \{2, \lceil 1/\varepsilon \rceil - 1\},$$

and this bound cannot be improved.

- 15.9** (Komlós et al., 1992) Let H be a hypergraph of VC-dimension d , let $\varepsilon > 0$, and let μ be a probability measure on $V(H)$ such that $\mu(E) \geq \varepsilon$ for every $E \in E(H)$. Furthermore, let

$$\pi_H(T) = \max_{\substack{A \subseteq H \\ |A|=T}} |\{E \cap A \mid E \in E(H)\}|.$$

- (i) Show that $\tau(H) \leq t(d, \varepsilon)$, where $t(d, \varepsilon)$ denotes the smallest positive integer t satisfying

$$2\pi_H(T) \left(1 - \frac{t}{T}\right)^{(T-t)\varepsilon-1} < 1 \quad \text{for some integer } T > t.$$

- (ii) Show that for any $C > 2$,

$$\tau(H) \leq \frac{d}{\varepsilon} \left(\ln \frac{1}{\varepsilon} + 2 \ln \ln \frac{1}{\varepsilon} + C \right),$$

provided that $\varepsilon < \varepsilon_C$. In particular, if $d \geq 2$ and $C = 5$ (resp. 4 or 3), the assertion is true with $\varepsilon_C = 1/3$ (resp. $1/6$ or $1/50$).

- 15.10** (Schläfli, 1901)

- (i) Show that any set of n hyperplanes in general position divides \mathbb{R}^d into exactly $\sum_{i=0}^d \binom{n}{i}$ regions.
(ii) Show that any set of n hyperplanes in general position divides the projective d -space into exactly $\sum_{i=0}^{\lfloor d/2 \rfloor} \binom{n}{d-2i}$ regions.
(iii) Show that any set of n points in general position in \mathbb{R}^d can be bisected by a hyperplane in exactly $\sum_{i=0}^{\lfloor d/2 \rfloor} \binom{n}{d-2i}$ different ways.

- 15.11** Determine the VC-dimension of the following range spaces:

- (i) $X = \mathbb{R}^d$, $\mathcal{R} = \{\text{all open half-spaces in } \mathbb{R}^d\}$.
(ii) $X = \mathbb{R}^d$, $\mathcal{R} = \{\text{all open balls in } \mathbb{R}^d\}$.

- 15.12** (Naiman and Wynn, 1994) Given an n -element point set $S \subseteq \mathbb{R}^d$ in general position, show that

$$|\{S \cap B^d \mid B^d \text{ is a ball}\}| = \sum_{i=0}^{d+1} \binom{n}{i},$$

provided that $n \geq d + 3$.

- 15.13** Given a set S of n points in \mathbb{R}^2 in general position, show that there exists a spanning path whose stabbing number is $c\sqrt{n}$, where c is a suitable constant.
- 15.14** Show that for any convex polygon with n sides, one can test in $O(\log n)$ time whether it intersects a given line.
- 15.15** Given a set S of n points in \mathbb{R}^2 in general position, prove that there exists a data structure using only $O(n)$ space, which allows to calculate $|S \cap h|$ for any half-plane h in time $O(\sqrt{n} \log^3 n)$.
- 15.16** (i) Given a set of $4n$ points in \mathbb{R}^2 in general position and any line ℓ with the property that there are exactly $2n$ points in both open half-planes determined by ℓ , show that there is a line ℓ' crossing ℓ such that each quadrant formed by these two lines contains exactly n points.
- (ii) (Willard, 1982) Let S be a set of n points in \mathbb{R}^2 in general position. Use part (i) to construct a data structure requiring linear space that allows us to calculate $|S \cap h|$ for any half-plane h in $O(n^{\log 3 / \log 4}) = O(n^{0.792\dots})$ time.
- (iii) (Edelsbrunner and Welzl, 1986b) Improve the exponent of the query time in part (ii) to $\log_2(\sqrt{5} + 1)/2 \approx 0.695$.
- 15.17*** (Matoušek, 1992) Let S be a set of n points in \mathbb{R}^2 in general position. A *geometric partitioning* of S is a collection $\{(S_1, \Delta_1), (S_2, \Delta_2), \dots, (S_m, \Delta_m)\}$, where S_1, \dots, S_m is a partition of S and Δ_i is a triangle containing S_i . Show that there is a constant c such that for any r ($1 \leq r \leq n$),
- (i) there exist a subset $S' \subseteq S$ of at least $n/2$ points and a geometric partitioning $\{(S'_1, \Delta_1), (S'_2, \Delta_2), \dots, (S'_m, \Delta_m)\}$ of S' such that $|S'_i| = \lfloor n/r \rfloor$ ($1 \leq i \leq m$) and every line intersects at most $c\sqrt{r}$ triangles;
- (ii) there exists a geometric partitioning of S of size at most cr such that $|S_i| \leq n/r$ ($1 \leq i \leq m$) and every line intersects at most \sqrt{r} triangles.
- 15.18** Given any set S of n points in \mathbb{R}^2 in general position, $0 < \varepsilon < 1$, let \mathcal{R}^ε be the family of all convex sets containing at least εn elements of S . A set of points $T \subseteq \mathbb{R}^2$ is said to be a “weak ε -net” for $\Sigma = (S, \mathcal{R}^\varepsilon)$ if every member of \mathcal{R}^ε contains at least one element of T . (Note that the points of T do not necessarily belong to S .) Show that:
- (i) $\max_{|S|=n} \text{VC-dim}(\Sigma) = \lfloor n(1 - \varepsilon) \rfloor$.
- (ii) There exists a constant c (independent of n , S , and ε) such that for every $0 < \varepsilon < 1$, Σ always has a weak ε -net consisting of at most $c/\varepsilon^{\log 4 / \log(4/3)}$ points.

- (iii) (Alon et al., 1992) The bound in part (ii) can be improved to c/ε^2 .
- (iv)* (Chazelle et al., 1993) If the points of S are the vertices of a regular n -gon, then Σ has a weak ε -net whose size is at most c/ε , where c is a suitable constant.
- (v) (Chazelle et al., 1993) If the points of S are the vertices of a convex n -gon, then Σ has a weak ε -net whose size is at most $(c/\varepsilon) \log^{\log^3} (1/\varepsilon)$, where c is a suitable constant.
- 15.19** (Aronov et al., 1994; Pach, 1991) We call two segments “independent” if neither of the lines supporting one of them crosses the other. Prove that n points in \mathbb{R}^2 in general position always determine at least $c\sqrt{n}$ mutually independent segments.
- 15.20** Use Corollary 15.6(i) to prove the following statement. Given a set L of n lines in general position and an integer $r \geq 3$, the plane can be decomposed into at most r^2 triangles (some of which are unbounded) such that the interior of each triangle intersects at most $c(n/r) \log r$ lines of L , where c is some fixed constant.
- 15.21** (Bárány and Lehel, 1987; Pach, 1994) Given two points $p, q \in \mathbb{R}^d$, let $\text{Box}(p, q)$ denote the smallest box parallel to the axes which contains p and q . Show that from any finite set of points $S \subseteq \mathbb{R}^d$ one can choose $2^{2^{d+2}}$ or fewer points p_i ($1 \leq i \leq 2^{2^{d+2}}$) such that

$$\bigcup_{1 \leq i < j \leq 2^{2^{d+2}}} \text{Box}(p_i, p_j) \supseteq S.$$

- 15.22** (Ding et al., 1994) Solve Exercise 14.11 by using Theorem 15.9.
- 15.23** (Gyárfás and Lehel, 1969) Show that for any positive integers k and ν , there exists a number $f(k, \nu)$ with the following property. Let H be a finite family of subsets of \mathbb{R} , each a union of at most k intervals. If H has no $\nu + 1$ pairwise disjoint members, then one can find at most $f(k, \nu)$ points such that every member of H contains one of them.