A point \( x \in \text{conv}(A_1) \cap \text{conv}(A_2) \), where \( A_1 \) and \( A_2 \) are as in the theorem, is called a Radon point of \( A \), and the pair \((A_1, A_2)\) is called a Radon partition of \( A \) (it is easily seen that we can require \( A_1 \cup A_2 = A \)).

Here are two possible cases in the plane:

\[ \text{Proof.} \] Let \( A = \{a_1, a_2, \ldots, a_{d+2}\} \). These \( d+2 \) points are necessarily affinely dependent. That is, there exist real numbers \( \alpha_1, \ldots, \alpha_{d+2} \), not all of them 0, such that \( \sum_{i=1}^{d+2} \alpha_i = 0 \) and \( \sum_{i=1}^{d+2} \alpha_i a_i = 0 \).

Set \( P = \{i: \alpha_i > 0\} \) and \( N = \{i: \alpha_i < 0\} \). Both \( P \) and \( N \) are nonempty. We claim that \( P \) and \( N \) determine the desired subsets. Let us put \( A_1 = \{a_i: i \in P\} \) and \( A_2 = \{a_i: i \in N\} \). We are going to exhibit a point \( x \) that is contained in the convex hulls of both these sets.

Put \( S = \sum_{i \in P} \alpha_i; \) we also have \( S = -\sum_{i \in N} \alpha_i \). Then we define

\[ x = \sum_{i \in P} \frac{\alpha_i}{S} a_i. \] (1.1)

Since \( \sum_{i=1}^{d+2} \alpha_i a_i = 0 = \sum_{i \in P} \alpha_i a_i + \sum_{i \in N} \alpha_i a_i \), we also have

\[ x = \sum_{i \in N} \frac{-\alpha_i}{S} a_i. \] (1.2)

The coefficients of the \( a_i \) in (1.1) are nonnegative and sum to 1, so \( x \) is a convex combination of points of \( A_1 \). Similarly, (1.2) expresses \( x \) as a convex combination of points of \( A_2 \).

Helly's theorem is one of the most famous results of a combinatorial nature about convex sets.

**1.3.2 Theorem (Helly's theorem).** Let \( C_1, C_2, \ldots, C_n \) be convex sets in \( \mathbb{R}^d \), \( n \geq d+1 \). Suppose that the intersection of every \( d+1 \) of these sets is nonempty. Then the intersection of all the \( C_i \) is nonempty.

The first nontrivial case states that if every 3 among 4 convex sets in the plane intersect, then there is a point common to all 4 sets. This can be proved by an elementary geometric argument, perhaps distinguishing a few cases, and the reader may want to try to find a proof before reading further.

In a contrapositive form, Helly's theorem guarantees that whenever \( C_1, C_2, \ldots, C_n \) are convex sets with \( \bigcap_{i=1}^{n} C_i = \emptyset \), then this is witnessed by some at most \( d+1 \) sets with empty intersection among the \( C_i \). In this way, many proofs are greatly simplified, since in planar problems, say, one can deal with 3 convex sets instead of an arbitrary number, as is amply illustrated in the exercises below.
It is very tempting and quite usual to formulate Helly’s theorem as follows: “If every \( d+1 \) among \( n \) convex sets in \( \mathbb{R}^d \) intersect, then all the sets intersect.” But, strictly speaking, this is false, for a trivial reason: For \( d \geq 2 \), the assumption as stated here is met by \( n = 2 \) disjoint convex sets.

**Proof of Helly’s theorem.** (Using Radon’s lemma.) For a fixed \( d \), we proceed by induction on \( n \). The case \( n = d+1 \) is clear, so we suppose that \( n \geq d+2 \) and that the statement of Helly’s theorem holds for smaller \( n \). Actually, \( n = d+2 \) is the crucial case; the result for larger \( n \) follows at once by a simple induction.

Consider sets \( C_1, C_2, \ldots, C_n \) satisfying the assumptions. If we leave out any one of these sets, the remaining sets have a nonempty intersection by the inductive assumption. Let us fix a point \( a_i \in \bigcap_{j \neq i} C_j \) and consider the points \( a_1, a_2, \ldots, a_{d+2} \). By Radon’s lemma, there exist disjoint index sets \( I_1, I_2 \subset \{1, 2, \ldots, d+2\} \) such that
\[
\text{conv}(\{a_i: i \in I_1\}) \cap \text{conv}(\{a_i: i \in I_2\}) \neq \emptyset.
\]
We pick a point \( x \) in this intersection. The following picture illustrates the case \( d = 2 \) and \( n = 4 \):

![Diagram illustrating Helly's theorem](Image)

We claim that \( x \) lies in the intersection of all the \( C_i \). Consider some \( i \in \{1, 2, \ldots, n\} \); then \( i \notin I_1 \) or \( i \notin I_2 \). In the former case, each \( a_j \) with \( j \in I_1 \) lies in \( C_i \), and so \( x \in \text{conv}(\{a_j: j \in I_1\}) \subseteq C_i \). For \( i \notin I_2 \) we similarly conclude that \( x \in \text{conv}(\{a_j: j \in I_2\}) \subseteq C_i \). Therefore, \( x \in \bigcap_{i=1}^{n} C_i \).

**An infinite version of Helly’s theorem.** If we have an infinite collection of convex sets in \( \mathbb{R}^d \) such that any \( d+1 \) of them have a common point, the entire collection still need not have a common point. Two examples in \( \mathbb{R}^1 \) are the families of intervals \( \{(0, 1/n): n = 1, 2, \ldots\} \) and \( \{[n, \infty): n = 1, 2, \ldots\} \). The sets in the first example are not closed, and the second example uses unbounded sets. For *compact* (i.e., closed and bounded) sets, the theorem holds:

**1.3.3 Theorem (Infinite version of Helly’s theorem).** Let \( C \) be an arbitrary infinite family of compact convex sets in \( \mathbb{R}^d \) such that any \( d+1 \) of the sets have a nonempty intersection. Then all the sets of \( C \) have a nonempty intersection.
Proof. By Helly’s theorem, any finite subfamily of $\mathcal{C}$ has a nonempty intersection. By a basic property of compactness, if we have an arbitrary family of compact sets such that each of its finite subfamilies has a nonempty intersection, then the entire family has a nonempty intersection.

Several nice applications of Helly’s theorem are indicated in the exercises below, and we will meet a few more later in this book.

Bibliography and remarks. Helly proved Theorem 1.3.2 in 1913 and communicated it to Radon, who published a proof in [Rad21]. This proof uses Radon’s lemma, although the statement wasn’t explicitly formulated in Radon’s paper. References to many other proofs and generalizations can be found in the already mentioned surveys [Eck93] and [DGK63].

Helly’s theorem inspired a whole industry of Helly-type theorems. A family $\mathcal{B}$ of sets is said to have Helly number $h$ if the following holds: Whenever a finite subfamily $\mathcal{F} \subseteq \mathcal{B}$ is such that every $h$ or fewer sets of $\mathcal{F}$ have a common point, then $\bigcap \mathcal{F} \neq \emptyset$. So Helly’s theorem says that the family of all convex sets in $\mathbb{R}^d$ has Helly number $d+1$. More generally, let $P$ be some property of families of sets that is hereditary, meaning that if $\mathcal{F}$ has property $P$ and $\mathcal{F}' \subseteq \mathcal{F}$, then $\mathcal{F}'$ has $P$ as well. A family $\mathcal{B}$ is said to have Helly number $h$ with respect to $P$ if for every finite $\mathcal{F} \subseteq \mathcal{B}$, all subfamilies of $\mathcal{F}$ of size at most $h$ having $P$ implies $\mathcal{F}$ having $P$. That is, the absence of $P$ is always witnessed by some at most $h$ sets, so it is a “local” property.

Exercises

1. Prove Carathéodory’s theorem (you may use Radon’s lemma).

2. Let $K \subseteq \mathbb{R}^d$ be a convex set and let $C_1, \ldots, C_n \subseteq \mathbb{R}^d$, $n \geq d+1$, be convex sets such that the intersection of every $d+1$ of them contains a translated copy of $K$. Prove that then the intersection of all the sets $C_i$ also contains a translated copy of $K$.

This result was noted by Vincensini [Vin39] and by Klee [Kle53].

3. Find an example of 4 convex sets in the plane such that the intersection of each 3 of them contains a segment of length 1, but the intersection of all 4 contains no segment of length 1.

4. A strip of width $w$ is a part of the plane bounded by two parallel lines at distance $w$. The width of a set $X \subseteq \mathbb{R}^2$ is the smallest width of a strip containing $X$.

(a) Prove that a compact convex set of width 1 contains a segment of length 1 of every direction.

(b) Let $\{C_1, C_2, \ldots, C_n\}$ be closed convex sets in the plane, $n \geq 3$, such that the intersection of every 3 of them has width at least 1. Prove that $\bigcap_{i=1}^{n} C_i$ has width at least 1.
The result as in (b), for arbitrary dimension \(d\), was proved by Sallee [Sal75], and a simple argument using Helly’s theorem was noted by Buchman and Valentine [BV82].

5. Statement: Each set \(X \subseteq \mathbb{R}^2\) of diameter at most 1 (i.e., any 2 points have distance at most 1) is contained in some disc of radius \(1/\sqrt{3}\).
   (a) Prove the statement for 3-element sets \(X\).
   (b) Prove the statement for all finite sets \(X\).
   (c) Generalize the statement to \(\mathbb{R}^d\): determine the smallest \(r = r(d)\) such that every set of diameter 1 in \(\mathbb{R}^d\) is contained in a ball of radius \(r\) (prove your claim).

The result as in (c) is due to Jung; see [DGK63].

6. Let \(C \subseteq \mathbb{R}^d\) be a compact convex set. Prove that the mirror image of \(C\) can be covered by a suitable translate of \(C\) blown up by the factor of \(d\); that is, there is an \(x \in \mathbb{R}^d\) with \(-C \subseteq x + dC\).

7. (a) Prove that if the intersection of each 4 or fewer among convex sets \(C_1, \ldots, C_n \subseteq \mathbb{R}^2\) contains a ray then \(\bigcap_{i=1}^{n} C_i\) also contains a ray.
   (b) Show that the number 4 in (a) cannot be replaced by 3. This result, and an analogous one in \(\mathbb{R}^d\) with the Helly number \(2d\), are due to Katchalski [Kat78].

8. For a set \(X \subseteq \mathbb{R}^2\) and a point \(x \in X\), let us denote by \(V(x)\) the set of all points \(y \in X\) that can “see” \(x\), i.e., points such that the segment \(xy\) is contained in \(X\). The kernel of \(X\) is defined as the set of all points \(x \in X\) such that \(V(x) = X\). A set with a nonempty kernel is called star-shaped.
   (a) Prove that the kernel of any set is convex.
   (b) Prove that if \(V(x) \cap V(y) \cap V(z) \neq \emptyset\) for every \(x, y, z \in X\) and \(X\) is compact, then \(X\) is star-shaped. That is, if every 3 paintings in a (planar) art gallery can be seen at the same time from some location (possibly different for different triples of paintings), then all paintings can be seen simultaneously from somewhere. If it helps, assume that \(X\) is a polygon.
   (c) Construct a nonempty set \(X \subseteq \mathbb{R}^2\) such that each of its finite subsets can be seen from some point of \(X\) but \(X\) is not star-shaped.

The result in (b), as well as the \(d\)-dimensional generalization (with every \(d+1\) regions \(V(x)\) intersecting), is called Krasnosel’skii’s theorem; see [Eck93] for references and related results.

9. In the situation of Radon’s lemma (\(A\) is a \((d+2)\)-point set in \(\mathbb{R}^d\)), call a point \(x \in \mathbb{R}^d\) a Radon point of \(A\) if it is contained in convex hulls of two disjoint subsets of \(A\). Prove that if \(A\) is in general position (no \(d+1\) points affinely dependent), then its Radon point is unique.

10. (a) Let \(X, Y \subseteq \mathbb{R}^2\) be finite point sets, and suppose that for every subset \(S \subseteq X \cup Y\) of at most 4 points, \(S \cap X\) can be separated (strictly) by a line from \(S \cap Y\). Prove that \(X\) and \(Y\) are line-separable.
    (b) Extend (a) to sets \(X, Y \subseteq \mathbb{R}^d\), with \(|S| \leq d+2\). The result (b) is called Kirchberger’s theorem [Kir03].
1.4 Centerpoint and Ham Sandwich

We prove an interesting result as an application of Helly’s theorem.

1.4.1 Definition (Centerpoint). Let $X$ be an $n$-point set in $\mathbb{R}^d$. A point $x \in \mathbb{R}^d$ is called a centerpoint of $X$ if each closed half-space containing $x$ contains at least $\frac{n}{d+1}$ points of $X$.

Let us stress that one set may generally have many centerpoints, and a centerpoint need not belong to $X$.

The notion of centerpoint can be viewed as a generalization of the median of one-dimensional data. Suppose that $x_1, \ldots, x_n \in \mathbb{R}$ are results of measurements of an unknown real parameter $x$. How do we estimate $x$ from the $x_i$? We can use the arithmetic mean, but if one of the measurements is completely wrong (say, 100 times larger than the others), we may get quite a bad estimate. A more “robust” estimate is a median, i.e., a point $x$ such that at least $\frac{n}{2}$ of the $x_i$ lie in the interval $(-\infty, x]$ and at least $\frac{n}{2}$ of them lie in $[x, \infty)$. The centerpoint can be regarded as a generalization of the median for higher-dimensional data.

In the definition of centerpoint we could replace the fraction $\frac{1}{d+1}$ by some other parameter $\alpha \in (0, 1)$. For $\alpha > \frac{1}{d+1}$, such an “$\alpha$-centerpoint” need not always exist: Take $d+1$ points in general position for $X$. With $\alpha = \frac{1}{d+1}$ as in the definition above, a centerpoint always exists, as we prove next.

Centerpoints are important, for example, in some algorithms of divide-and-conquer type, where they help divide the considered problem into smaller subproblems. Since no really efficient algorithms are known for finding “exact” centerpoints, the algorithms often use $\alpha$-centerpoints with a suitable $\alpha < \frac{1}{d+1}$, which are easier to find.

1.4.2 Theorem (Centerpoint theorem). Each finite point set in $\mathbb{R}^d$ has at least one centerpoint.

Proof. First we note an equivalent definition of a centerpoint: $x$ is a centerpoint of $X$ if and only if it lies in each open half-space $\gamma$ such that $|X \cap \gamma| > \frac{d}{d+1} n$.

We would like to apply Helly’s theorem to conclude that all these open half-spaces intersect. But we cannot proceed directly, since we have infinitely many half-spaces and they are open and unbounded. Instead of such an open half-space $\gamma$, we thus consider the compact convex set $\text{conv}(X \cap \gamma) \subset \gamma$. 
Letting $\gamma$ run through all open half-spaces $\gamma$ with $|X \cap \gamma| > \frac{d}{d+1} n$, we obtain a family $C$ of compact convex sets. Each of them contains more than $\frac{d}{d+1} n$ points of $X$, and so the intersection of any $d+1$ of them contains at least one point of $X$. The family $C$ consists of finitely many distinct sets (since $X$ has finitely many distinct subsets), and so $\bigcap C \neq \emptyset$ by Helly's theorem. Each point in this intersection is a centerpoint.

In the definition of a centerpoint we can regard the finite set $X$ as defining a distribution of mass in $\mathbb{R}^d$. The centerpoint theorem asserts that for some point $x$, any half-space containing $x$ encloses at least $\frac{1}{d+1}$ of the total mass. It is not difficult to show that this remains valid for continuous mass distributions, or even for arbitrary Borel probability measures on $\mathbb{R}^d$ (Exercise 1).

**Ham-sandwich theorem and its relatives.** Here is another important result, not much related to convexity but with a flavor resembling the centerpoint theorem.

**1.4.3 Theorem (Ham-sandwich theorem).** Every $d$ finite sets in $\mathbb{R}^d$ can be simultaneously bisected by a hyperplane. A hyperplane $h$ bisects a finite set $A$ if each of the open half-spaces defined by $h$ contains at most $\frac{|A|}{2}$ points of $A$.

This theorem is usually proved via continuous mass distributions using a tool from algebraic topology: the Borsuk–Ulam theorem. Here we omit a proof.

Note that if $A_i$ has an odd number of points, then every $h$ bisecting $A_i$ passes through a point of $A_i$. Thus if $A_1, \ldots, A_d$ all have odd sizes and their union is in general position, then every hyperplane simultaneously bisecting them is determined by $d$ points, one of each $A_i$. In particular, there are only finitely many such hyperplanes.

Again, an analogous ham-sandwich theorem holds for arbitrary $d$ Borel probability measures in $\mathbb{R}^d$.

**Center transversal theorem.** There can be beautiful new things to discover even in well-studied areas of mathematics. A good example is the following recent result, which “interpolates” between the centerpoint theorem and the ham-sandwich theorem.

**1.4.4 Theorem (Center transversal theorem).** Let $1 \leq k \leq d$ and let $A_1, A_2, \ldots, A_k$ be finite point sets in $\mathbb{R}^d$. Then there exists a $(k-1)$-flat $f$ such that for every hyperplane $h$ containing $f$, both the closed half-spaces defined by $h$ contain at least $\frac{1}{d-k+2} |A_i|$ points of $A_i$, $i = 1, 2, \ldots, k$.

The ham-sandwich theorem is obtained for $k = d$ and the centerpoint theorem for $k = 1$. The proof, which we again have to omit, is based on a result of algebraic topology, too, but it uses a considerably more advanced machinery than the ham-sandwich theorem. However, the weaker result with $\frac{1}{d+1}$ instead of $\frac{1}{d-k+2}$ is easy to prove; see Exercise 2.
Bibliography and remarks. The centerpoint theorem was established by Rado [Rad47]. According to Steinlein's survey [Ste85], the ham-sandwich theorem was conjectured by Steinhaus (who also invented the popular 3-dimensional interpretation, namely, that the ham, the cheese, and the bread in any ham sandwich can be simultaneously bisected by a single straight motion of the knife) and proved by Banach. The center transversal theorem was found by Dol'nikov [Dol'92] and, independently, by Živaljević and Vrećica [ŽV90].

Significant effort has been devoted to efficient algorithms for finding (approximate) centerpoints and ham-sandwich cuts (i.e., hyperplanes as in the ham-sandwich theorem). In the plane, a ham-sandwich cut for two \( n \)-point sets can be computed in linear time (Lo, Matoušek, and Steiger [LMS94]). In a higher but fixed dimension, the complexity of the best exact algorithms is currently slightly better than \( O(n^{d-1}) \). A centerpoint in the plane, too, can be found in linear time (Jadhav and Mukhopadhyay [JM94]). Both approximate ham-sandwich cuts (in the ratio \( 1 : 1+\varepsilon \) for a fixed \( \varepsilon > 0 \)) and approximate centerpoints (\( (\frac{1}{d+1} - \varepsilon) \)-centerpoints) can be computed in time \( O(n) \) for every fixed dimension \( d \) and every fixed \( \varepsilon > 0 \), but the constant depends exponentially on \( d \), and the algorithms are impractical if the dimension is not quite small. A practically efficient randomized algorithm for computing approximate centerpoints in high dimensions (\( \alpha \)-centerpoints with \( \alpha \approx 1/d^2 \)) was given by Clarkson, Eppstein, Miller, Sturtivant, and Teng [CEM+96].

Exercises

1. (Centerpoints for general mass distributions)
   (a) Let \( \mu \) be a Borel probability measure on \( \mathbb{R}^d \); that is, \( \mu(\mathbb{R}^d) = 1 \) and each open set is measurable. Show that for each open half-space \( \gamma \) with \( \mu(\gamma) > t \) there exists a compact set \( C \subset \gamma \) with \( \mu(C) > t \). \( \Box \)
   (b) Prove that each Borel probability measure in \( \mathbb{R}^d \) has a centerpoint (use (a) and the infinite Helly's theorem). \( \Box \)

2. Prove that for any \( k \) finite sets \( A_1, \ldots, A_k \subset \mathbb{R}^d \), where \( 1 \leq k \leq d \), there exists a \( (k-1) \)-flat such that every hyperplane containing it has at least \( \frac{1}{d+1} |A_i| \) points of \( A_i \) in both of its closed half-spaces for all \( i = 1, 2, \ldots, k \).