23 Appendices

Chapter summary

In this chapter we present some basic results from different areas of mathematics required for various proofs in the book. In Section 23.1 we state and prove several fixed point theorems. The main and best known is Brouwer's Fixed Point Theorem, which states that every continuous function from a compact and convex subset of a Euclidean space to itself has a fixed point. This theorem is used in Chapter 5 to prove the existence of a Nash equilibrium in mixed strategies. Using Brouwer's Fixed Point Theorem we prove Kakutani's Fixed Point Theorem, which states that every upper semi-continuous convex-valued correspondence from a compact and convex subset of a Euclidean space to itself has a fixed point. This result provides a shorter proof for the existence of a Nash equilibrium in mixed strategies in strategic-form games. We then prove the KKM theorem, which is used to prove the nonemptiness of the bargaining set (Theorem 19.19, page 790). The main tool for proving both Brouwer's Fixed Point Theorem and the KKM Theorem is Sperner's Lemma, which is stated and proved first.

In Section 23.2 we prove the Separating Hyperplane Theorem, which states that for every convex set in a Euclidean space and a point not in the set there is a hyperplane separating the set and the point. This theorem is used in Chapter 14 to prove that every B-set is an approachable set.

Section 23.3 presents the formulation and the central result in linear programs, namely, the Duality Theorem of linear programming. This result is used to prove the Bondareva–Shapley Theorem that characterizes coalitional games with nonempty cores (Theorem 17.19, page 701) and to characterize balanced collections of coalitions (Theorem 20.25, page 817).

23.1 Fixed point theorems

In this section we formulate and prove three theorems: Brouwer's Fixed Point Theorem, Kakutani's Fixed Point Theorem, and the KKM Theorem. The reader interested in additional fixed point theorems used in game theory, and in a discussion of the connections between them, is directed to Border [1989].

We begin by presenting the statement and proof of Sperner's Lemma, which is a vital element in our proofs of fixed point theorems.

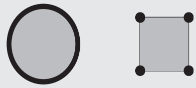


Figure 23.1 The extreme points of a circle and of a square

23.1.1 Sperner's Lemma

For $n \in \mathbb{N}$ denote the zero vector in $\mathbb{R}^n$ by $\vec{0}$. Recall that a set $X \subseteq \mathbb{R}^n$ is called *convex* if $\lambda x + (1-\lambda)y \in X$ for all $x, y \in X$ and all $\lambda \in [0,1]$.

Definition 23.1 *Let $x^0, x^1, \ldots, x^k$ be vectors in $\mathbb{R}^n$. The* convex hull *of $x^0, x^1, \ldots, x^k$, denoted* $\operatorname{conv}(x^0, x^1, \ldots, x^k)$*, is the smallest convex set (with respect to set inclusion) that contains $x^0, x^1, \ldots, x^k$.*

When $x = \sum_{l=0}^{k} \alpha^l x^l$ for nonnegative numbers $(\alpha^j)_{j=1}^k$ whose sum is 1, we say that x is a *convex combination* of the points $x^0, x^1, \ldots, x^k$. It follows (see Exercise 23.1) that

$$\operatorname{conv}(x^0, x^1, \ldots, x^k) = \left\{x \in \mathbb{R}^n : x \text{ is a convex combination of } x^0, \ldots, x^k\right\}. \quad (23.1)$$

Definition 23.2 *Let $X \subseteq \mathbb{R}^n$ be a convex set. A vector $x \in X$ is an* extreme point *of X if for every two distinct points $y, z \in X$, and for every $\alpha \in (0,1)$,*

$$\alpha y + (1-\alpha)z \neq x. \quad (23.2)$$

In words, a point is an extreme point of a convex set if it cannot be expressed as a convex combination of two different points in that set. Equivalently, a point is an extreme point of a convex set if there is no open interval that both contains it and is contained in that set. If none of the vectors $x^0, x^1, \ldots, x^k$ is a convex combination of the other vectors, then the extreme points of the convex hull $\operatorname{conv}(x^0, x^1, \ldots, x^k)$ are $x^0, x^1, \ldots, x^k$ (Exercise 23.2). Figure 23.1 depicts the extreme points of a square (four points) and of a circle (all the boundary points of the circle) in $\mathbb{R}^2$.

Definition 23.3 *The vectors $x^0, x^1, \ldots, x^k$ in $\mathbb{R}^n$ are called* affine independent *if the only solution of the following system of equations with unknowns $(\alpha^l)_{l=0}^k$ in $\mathbb{R}$:*

$$\sum_{l=0}^{k} \alpha^l x^l = \vec{0}, \quad (23.3)$$

$$\sum_{l=0}^{k} \alpha^l = 0, \quad (23.4)$$

is given by $\alpha^0 = \alpha^1 = \cdots = \alpha^k = 0$. If this condition is not satisfied, the vectors $x^0, x^1, \ldots, x^k$ in $\mathbb{R}^n$ are called affine dependent.

We say that a vector $y \in \mathbb{R}^n$ is affine independent of the vectors $x^0, x^1, \ldots, x^k$ if the vectors $(x^0, x^1, \ldots, x^k, y)$ are affine independent.

It follows from the definition that every subset of a set of affine-independent vectors is also affine independent (why?), and every superset of a set of affine-dependent vectors

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

is also affine dependent. In addition, affine dependence implies linear dependence, which means that linear independence implies affine independence. The following example shows that the converse does not hold. The example shows that the space $\mathbb{R}^n$ may contain $n+1$ vectors that are affine independent. Since any $n+1$ vectors in $\mathbb{R}^n$ are linearly dependent, it follows that affine independence does not imply linear independence. Thus, the concepts of linear independence and affine independence are not identical.

Example 23.4 For $n > 1$ denote by $e^l = (0, \ldots, 0, 1, 0, \ldots, 0)$ the l unit vector in $\mathbb{R}^n$; by this we mean the vector all of whose coordinates are 0 except for the l-th coordinate, which is equal to 1. The set $\{e^1, e^2, \ldots, e^n, \vec{0}\}$ is a set of $n+1$ affine-independent vectors.

To see this, let $(\alpha^l)_{l=1}^{n+1}$ be a solution of the following system of equations:

$$\sum_{l=1}^{n} \alpha^l e^l + \alpha^{n+1}\vec{0} = \vec{0}, \tag{23.5}$$

$$\sum_{l=1}^{n+1} \alpha^l = 0. \tag{23.6}$$

For $l = 1, 2, \ldots, n$, the l-th coordinate on the left-hand side of Equation (23.5) equals α^l, and therefore $\alpha^l = 0$ for every $l = 1, 2, \ldots, n$. Equation (23.6) implies that $\alpha^{n+1} = 0$, and therefore the vectors $e^1, e^2, \ldots, e^n, \vec{0}$ are affine independent. ◀

Note that a set of vectors containing the vector $\vec{0}$ is necessarily linearly dependent. In contrast, as Example 23.4 shows, a set of vectors containing the vector $\vec{0}$ may be affine independent.

Example 23.5 Let $x^0, x^1, \ldots, x^k$ be affine-independent vectors. Denote their center of gravity by $y := \frac{1}{k+1}\sum_{l=0}^{k} x^l$. Then the set $\{x^1, \ldots, x^k, y\}$ is affine independent (note that x^0 is not in this set). To see this, let $(\alpha^l)_{l=1}^{k+1}$ be a solution of the following system of equations:

$$\sum_{l=1}^{k} \alpha^l x^l + \alpha^{k+1} y = \vec{0}, \tag{23.7}$$

$$\sum_{l=1}^{k+1} \alpha^l = 0. \tag{23.8}$$

Define

$$\beta^l := \begin{cases} \alpha^l + \frac{\alpha^{k+1}}{k+1} & \text{if } 1 \le l \le k, \\ \frac{\alpha^{k+1}}{k+1} & \text{if } l = 0. \end{cases} \tag{23.9}$$

Then $\sum_{l=0}^{k} \beta^l = \sum_{l=1}^{k+1} \alpha^l = 0$ (check that this true), and by (23.7),

$$\sum_{l=0}^{k} \beta^l x^l = \frac{\alpha^{k+1}}{k+1} x^0 + \sum_{l=1}^{k}\left(\alpha^l + \frac{\alpha^{k+1}}{k+1}\right) x^l = \sum_{l=1}^{k} \alpha^l x^l + \alpha^{k+1} y = \vec{0}. \tag{23.10}$$

Since the vectors $x^0, x^1, \ldots, x^k$ are affine independent, $\beta^l = 0$ for all $l = 0, 1, \ldots, k$, and therefore $\alpha^l = 0$ for all $l = 1, 2, \ldots, k+1$. In other words, the vectors $x^1, \ldots, x^k, y$ are affine independent. ◀

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

In Example 23.5, we can replace each of the vectors $x^1, \dots, x^k$ (not necessarily x^0) with the center of weight y and the resulting set of vectors will still remain affine independent. In fact, we can replace each of the vectors $\{x^0, x^1, \dots, x^k\}$ with any convex combination $z := \sum_{l=0}^{k} \gamma^l x^l$ in which all the weights $(\gamma^l)_{l=0}^{k}$ are positive (but not necessarily equal) and the resulting set of vectors will still remain affine independent (Exercise 23.6).

The last two examples show that there may be $n+1$ affine-independent vectors in $\mathbb{R}^n$. The next theorem shows that any $n+2$ vectors in $\mathbb{R}^n$ are affine dependent.

Theorem 23.6 *Every subset of $\mathbb{R}^n$ containing $n+2$ or more vectors is affine dependent.*

Proof: Since every superset of a set of affine-dependent vectors is also affine dependent, it suffices to prove that every set of $n+2$ vectors in $\mathbb{R}^n$ is affine dependent. Suppose, then, by contradiction that there exist $n+2$ affine-independent vectors $x^1, x^2, \dots, x^{n+2}$ in $\mathbb{R}^n$. Since any subset of a set of affine-independent vectors is itself a set of affine-independent vectors, the vectors $x^1, x^2, \dots, x^{n+1}$ are also affine independent. Since any $n+1$ vectors in $\mathbb{R}^n$ are linearly dependent, there exist real numbers $(\beta^l)_{l=1}^{n+1}$, not all of them 0, such that

$$\sum_{l=1}^{n+1} \beta^l x^l = \vec{0}. \tag{23.11}$$

Since the vectors $x^1, x^2, \dots, x^{n+1}$ are affine independent, $\beta := \sum_{l=1}^{n+1} \beta^l \neq 0$. Suppose without loss of generality that $\beta^1 \neq 0$. Applying the same reasoning to the $n+1$ vectors $x^2, x^3, \dots, x^{n+2}$ leads to the conclusion that there exist real numbers $(\gamma^l)_{l=2}^{n+2}$, not all of them 0, such that

$$\sum_{l=2}^{n+2} \gamma^l x^l = \vec{0}, \qquad \gamma := \sum_{l=2}^{n+2} \gamma^l \neq 0. \tag{23.12}$$

Define

$$\alpha^l := \begin{cases} \gamma\beta^1 & \text{if } l = 1, \\ \gamma\beta^l - \beta\gamma^l & \text{if } 2 \leq l \leq n+1, \\ -\beta\gamma^{n+2} & \text{if } l = n+2. \end{cases} \tag{23.13}$$

Then

$$\sum_{l=1}^{n+2} \alpha^l x^l = \gamma \sum_{l=1}^{n+1} \beta^l x^l - \beta \sum_{l=2}^{n+2} \gamma^l x^l = \vec{0}, \tag{23.14}$$

$$\sum_{l=1}^{n+2} \alpha^l = \gamma \sum_{l=1}^{n+1} \beta^l - \beta \sum_{l=2}^{n+2} \gamma^l = \gamma\beta - \beta\gamma = 0, \tag{23.15}$$

and morever $\alpha^1 = \gamma\beta^1 \neq 0$. It follows that the vectors $x^1, x^2, \dots, x^{n+2}$ are affine dependent, contradicting the assumption we started with. The conclusion is that there do not exist $n+2$ affine-independent vectors in $\mathbb{R}^n$. □

For $n = 1$ the space $\mathbb{R}^n$ is the real line, and vectors are real numbers. The theorem states that in that space every triple of real numbers is affine dependent. Any given three

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

vectors in $\mathbb{R}^n$ are affine dependent if and only if none of them is a convex combination of the other two (Exercise 23.3).

The following theorem shows how to find affine-independent vectors using linearly independent vectors. Its proof is left to the reader (Exercise 23.5).

Theorem 23.7 *Let $x^0, x^1, \ldots, x^k$ be affine-independent vectors in $\mathbb{R}^n$, and let y be a vector that is linearly independent of $\{x^0, x^1, \ldots, x^k\}$. Then the vectors $x^0, x^1, \ldots, x^k, y$ are affine independent.*

Definition 23.8 *A set $S \subseteq \mathbb{R}^n$ is called* a k-dimensional simplex *if it is the convex hull of $k+1$ affine-independent vectors.*

If $x^0, x^1, \ldots, x^k$ are affine-independent vectors whose convex hull is the k-dimensional simplex S, then we write

$$S = \langle\langle x^0, x^1, \ldots, x^k \rangle\rangle := \operatorname{conv}(x^0, x^1, \ldots, x^k). \tag{23.16}$$

A simplex is a compact and convex set. A zero-dimensional simplex is a set containing only one point. A one-dimensional simplex is a closed interval, and a two-dimensional simplex is a triangle.

The next theorem, whose proof is left to the reader (Exercise 23.7), states that every vector in a simplex can be uniquely represented as a convex combination of the extreme points of the simplex.

Theorem 23.9 *Let $x^0, x^1, \ldots, x^k$ be affine-independent vectors in $\mathbb{R}^n$. Let $y \in \langle\langle x^0, x^1, \ldots, x^k \rangle\rangle$. Then y has a unique representation as a convex combination of $x^0, x^1, \ldots, x^k$. In other words, the following system of equations with unknowns $(\alpha^l)_{l=0}^k$ in $\mathbb{R}$ has a unique solution:*

$$\sum_{l=0}^{k} \alpha^l x^l = y, \tag{23.17}$$

$$\sum_{l=0}^{k} \alpha^l = 1, \tag{23.18}$$

$$\alpha^l \geq 0 \quad \forall l = 0, 1, \ldots, k. \tag{23.19}$$

If the system of Equations (23.17)–(23.18) has a solution (omitting Equation (23.19)) we say that y is an *affine combination* of $(x^l)_{l=0}^k$. It follows that if y is a convex combination of $(x^l)_{l=0}^k$, then it is an affine combination of those vectors (but the converse is not true).

Theorem 23.9 leads to the definition of a new coordinate system for vectors in the simplex $\langle\langle x^0, x^1, \ldots, x^k \rangle\rangle$: the coordinates of a point y in the simplex are the weights $(\alpha^l)_{l=0}^k$ satisfying Equations (23.17)–(23.19). Every point $y \in \mathbb{R}^n$ that is a linear combination of $x^0, x^1, \ldots, x^k$ but is not in their convex hull can also be represented by weights $(\alpha^l)_{l=0}^k$ satisfying Equations (23.17)–(23.18) but not Equation (23.19) (Exercise 23.9). This coordinate system is called the *barycentric coordinate system* (relative to the vectors $x^0, x^1, \ldots, x^k$).

The barycentric coordinates of a point in the simplex $\langle\langle x^0, x^1, \ldots, x^k\rangle\rangle$ can be given the following physical interpretation: if $(\alpha^l)_{l=0}^k$ are the barycentric coordinates of y (relative to the vectors $x^0, x^1, \ldots, x^k$), then if we place a weight α^l at each point x^l, the center of gravity of the resulting system is y.

Since every subset of a set of affine-independent vectors is itself a set of affine-independent vectors, we deduce the following theorem.

Theorem 23.10 *Let $\langle\langle x^0, x^1, \ldots, x^k\rangle\rangle$ be a k-dimensional simplex in $\mathbb{R}^n$. Then for every set $\{x^{l_0}, x^{l_1}, \ldots, x^{l_t}\} \subseteq \{x^0, x^1, \ldots, x^k\}$, the convex hull of $x^{l_0}, x^{l_1}, \ldots, x^{l_t}$ is a t-dimensional simplex in $\mathbb{R}^n$.*

The simplex $\langle\langle x^{l_0}, x^{l_1}, \ldots, x^{l_t}\rangle\rangle$ is called a t-dimensional *face* of S. The face of a simplex S, of any dimension, is a subsimplex of S, that is, a simplex contained in S. In particular, the simplex S is itself a face of S.

Definition 23.11 *Let $S = \langle\langle x^0, x^1, \ldots, x^k\rangle\rangle$ be a simplex in $\mathbb{R}^n$. The* boundary *of S is the union of all the $(k-1)$-dimensional subsimplices of S, i.e.,*[1]

$$\bigcup_{l=0}^{k} \langle\langle x^0, x^1, \ldots, x^{l-1}, x^{l+1}, \ldots, x^k\rangle\rangle.$$

In other words, the boundary of S is the union of all the $(k-1)$ faces of S, and it contains all the points y in S whose barycentric coordinate representation contains at least one coordinate that is zero (Exercise 23.10).

Definition 23.12 *Let $S \subseteq \mathbb{R}^n$ be a k-dimensional simplex in $\mathbb{R}^n$. A* simplicial partition *of S is a collection $\mathcal{T} = \{T_1, T_2, \ldots, T_M\}$ of simplices in $\mathbb{R}^n$ satisfying.*[2]

1. *$\bigcup_{m=1}^M T_m = S$: the union of all the simplices in $\mathcal{T}$ is the simplex S.*
2. *For every j, m, $0 \le j \le m \le M$, the intersection $T_j \cap T_m$ is either the empty set or a face both of T_j and of T_m.*
3. *If T is a simplex in the collection $\mathcal{T}$, then all of its faces are also elements of this collection.*
4. *If T is an l-dimensional simplex in $\mathcal{T}$, for $l < k$, then it is contained in an $l+1$-dimensional simplex in $\mathcal{T}$.*

Example 23.13 Let x^0, x^1, x^2 be three affine-independent vectors in $\mathbb{R}^2$. Figure 23.2 depicts three partitions of $\langle\langle x^0, x^1, x^2\rangle\rangle$.

In Partition A, the collection T_1, T_2 and all their faces is not a simplicial partition, because T_1 is not a simplex (T_1 is the convex hull of 4 points in $\mathbb{R}^2$, and in $\mathbb{R}^2$ any four points are affine dependent (see Theorem 23.6 on page 919)).

In Partition B, the collection of simplices T_1, T_2, T_3 and all their faces is not a simplicial partition, because the intersection $T_1 \cap T_2$ (as well as the intersection $T_1 \cap T_3$) is not a face of T_1.

1 For $l = 0$ and $l = k$, the simplices in the union are $\langle\langle x^1, x^2, \ldots, x^k\rangle\rangle$ and $\langle\langle x^0, x^1, \ldots, x^{k-1}\rangle\rangle$, respectively.
2 The fourth property follows from the other properties, and is therefore superfluous. We will not prove this fact here.

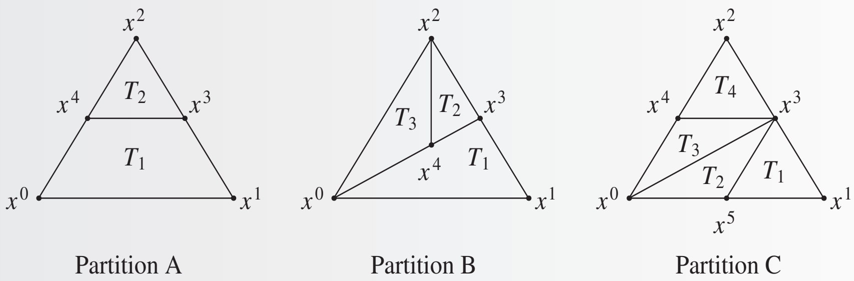


Figure 23.2 Examples of partitions of a two-dimensional simplex

In Partition C, the collection of simplices T_1, T_2, T_3, T_4 and all their faces is a simplicial partition. In contrast, the collection T_1, T_2, T_3, T_4, $T_1 \cup T_2$ and all their faces is not a simplicial partition, because the intersection of T_1 and $T_1 \cup T_2$ is T_1, which is not a face of $T_1 \cup T_2$. ◀

For every simplicial partition $\mathcal{T} = \{T_1, T_2, \dots, T_M\}$, denote by $Y(\mathcal{T})$ the set of all the extreme points of the simplices $(T_m)_{m=1}^M$. In other words, every T_m is the convex hull of some of the points in $Y(\mathcal{T})$. For example, for Partition C in Figure 23.2,

$$Y(\mathcal{T}) = \{x^0, x^1, x^2, x^3, x^4, x^5\}. \tag{23.20}$$

The following theorem follows from Properties (1) and (4) in Definition 23.12 (Exercise 23.14).

Theorem 23.14 *Let S be a k-dimensional simplex in $\mathbb{R}^n$, and $\mathcal{T}$ be a simplicial partition of S. Then S equals the union of all the k-dimensional simplices in $\mathcal{T}$.*

For every simplex $S = \langle\!\langle x^0, x^1, \dots, x^k \rangle\!\rangle$ in $\mathbb{R}^n$ denote by H_S the affine space spanned by the vectors in S,

$$H_S := \left\{ \sum_{l=0}^{k} \alpha^l x^l : \sum_{l=0}^{k} \alpha^l = 1 \right\} \subseteq \mathbb{R}^n. \tag{23.21}$$

The affine space H_S is a k-dimensional space, just like S.

Theorem 23.15 *Let $S = \langle\!\langle x^0, x^1, \dots, x^k \rangle\!\rangle$ be a k-dimensional simplex in $\mathbb{R}^n$, let $\mathcal{T}$ be a simplicial partition of S, and let $T \in \mathcal{T}$ be a $(k-1)$-dimensional simplex. If T is in the boundary of S, then T is contained in a unique k-dimensional simplex in $\mathcal{T}$. If T is not in the boundary of S, then T is contained in two k-dimensional simplices in $\mathcal{T}$.*

Consider partition C in Figure 23.2 (Example 23.13). This is the partition containing T_1, T_2, T_3, T_4 and all their faces. The one-dimensional simplex $\langle\!\langle x^1, x^3 \rangle\!\rangle$ is on the boundary of S and contained in a single two-dimensional simplex, T_1. In contrast, the one-dimensional simplex $\langle\!\langle x^3, x^5 \rangle\!\rangle$ is not on the boundary of S and is contained in two two-dimensional simplices, T_1 and T_2.

Proof: Let $T = \langle\langle y^0, y^1, \dots, y^{k-1} \rangle\rangle$ be a $(k-1)$-dimensional simplex in $\mathcal{T}$. By Property (4) in Definition 23.12, every $(k-1)$-dimensional simplex in $\mathcal{T}$ is contained in at least one k-dimensional simplex in $\mathcal{T}$, and by Property (2) it is a face of every such simplex.

Step 1: Preparations.
Let $\widehat{T} = \langle\langle y^0, y^1, \dots, y^{k-1}, y \rangle\rangle$ and $\widehat{T}' = \langle\langle y^0, y^1, \dots, y^{k-1}, y' \rangle\rangle$ be two k-dimensional simplices in $\mathcal{T}$ such that T is a face of both of them (the possibility that $\widehat{T} = \widehat{T}'$ is not ruled out). In particular, y and y' are not in T. Since $\widehat{T}$ and $\widehat{T}'$ are contained in S, it follows that $H_{\widehat{T}}$ and $H_{\widehat{T}'}$ are contained in H_S. These three spaces are all k-dimensional affine spaces, and therefore they coincide (Exercise 23.11),

$$H_{\widehat{T}} = H_{\widehat{T}'} = H_S. \tag{23.22}$$

It follows that the vector y' can be written as an affine combination of the vectors $y^0, y^1, \dots, y^{k-1}, y$,

$$y' = \sum_{l=0}^{k-1} \alpha^l y^l + \beta y, \quad \sum_{l=0}^{k-1} \alpha^l + \beta = 1. \tag{23.23}$$

Since $y' \notin T$ it follows that $\beta \neq 0$. Indeed, if $\beta = 0$, then by Equation (23.23) we would deduce that y' is an affine combination of $(y^l)_{l=0}^k$, and then $H_{\widehat{T}'}$ would be a $(k-1)$-dimensional affine space, contradicting the fact that it is a k-dimensional affine space.

Step 2: If $\beta > 0$, then $\widehat{T} = \widehat{T}'$.
For every $\varepsilon \in (0, 1]$ the vector $z^\varepsilon := (1-\varepsilon) \sum_{l=0}^{k-1} \frac{1}{k} y^l + \varepsilon y'$ is in $\widehat{T}'$ but not in T (since $y' \notin T$). For $\varepsilon > 0$ sufficiently small, the vector z^ε is also in $\widehat{T}$, since

$$z^\varepsilon = (1-\varepsilon) \sum_{l=0}^{k-1} \frac{1}{k} y^l + \varepsilon y' = \sum_{l=0}^{k-1} \left((1-\varepsilon)\frac{1}{k} + \varepsilon \alpha^l \right) y^l + \varepsilon \beta y, \tag{23.24}$$

and all the coefficients in the right-hand expression are positive for $\varepsilon > 0$ sufficiently small. It follows that the intersection $\widehat{T} \cap \widehat{T}'$ contains both T and z^ε for $\varepsilon > 0$ sufficiently small. Therefore $\widehat{T} \cap \widehat{T}'$ strictly contains T. On the other hand, by Property (2) this intersection is a face both of $\widehat{T}$ and of $\widehat{T}'$. But the only face of the simplex $\widehat{T}$ properly containing T is $\widehat{T}$ itself, and therefore $\widehat{T} \cap \widehat{T}' = \widehat{T}$. Similarly, $\widehat{T} \cap \widehat{T}' = \widehat{T}'$. We deduce from this that $\widehat{T} = \widehat{T}'$.

Step 3: If T is in the boundary of S, then it is contained in a unique k-dimensional simplex in $\mathcal{T}$.
We will show that if T is in the boundary of S then $\beta > 0$, and by Step 2 it then follows that T is contained in a unique k-dimensional simplex in $\mathcal{T}$. Recall that $x^0, x^1, \dots, x^k$ are the extreme points of S. By assumption, the simplex T is in the boundary of S, which is equal to $\bigcup_{i=0}^k \langle\langle x^0, \dots, x^{i-1}, x^{i+1}, \dots, x^k \rangle\rangle$. Since every simplex is a convex set, T is in one of the $(k-1)$-dimensional faces of S. Suppose, without loss of generality, that $T \subseteq \langle\langle x^0, x^1, \dots, x^{k-1} \rangle\rangle$. It follows that each of the vectors $(y^l)_{l=0}^{k-1}$ can be represented as a convex combination of $x^0, x^1, \dots, x^{k-1}$,

$$y^l = \sum_{j=0}^{k-1} \alpha^{jl} x^j, \quad \sum_{j=0}^{k-1} \alpha^{jl} = 1, \qquad l = 0, 1, \dots, k-1. \tag{23.25}$$

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

Since the vectors y and y' are in S, they can be represented as convex combinations of the extreme points of S,

$$y = \sum_{j=0}^{k} \gamma^j x^j, \qquad y' = \sum_{j=0}^{k} \gamma'^j x^j, \qquad \sum_{j=0}^{k} \gamma^j = \sum_{j=0}^{k} \gamma'^j = 1, \tag{23.26}$$

where $\gamma^j, \gamma'^j \geq 0$ for every $j \in \{0, 1, \ldots, k\}$.

Since y and y' are not in T, the coefficients γ^k are γ'^k are in fact positive. Plugging in these representations of the vectors into Equation (23.23), we get

$$\sum_{j=0}^{k} \gamma'^j x^j = y' = \sum_{l=0}^{k-1} \alpha^l y^l + \beta y \tag{23.27}$$

$$= \sum_{l=0}^{k-1} \left(\alpha^l \sum_{j=0}^{k-1} \alpha^{jl} x^j \right) + \beta \sum_{j=0}^{k} \gamma^j x^j \tag{23.28}$$

$$= \sum_{j=0}^{k-1} \left(\sum_{l=0}^{k-1} \alpha^l \alpha^{jl} + \beta \gamma^j \right) x^j + \beta \gamma^k x^k. \tag{23.29}$$

Since every vector in a simplex can be represented in a unique way as a convex combination of extreme points of the simplex, we deduce that the coefficient of x^k in both representations must be identical: $\gamma'^k = \beta \gamma^k$. Since both γ'^k and γ^k are positive, the conclusion is that $\beta > 0$, which is what we wanted to show.

Step 4: If T is not in the boundary of S, then it is contained in at least two k-dimensional simplices in $\mathcal{T}$.

By Property (4) of simplicial partitions, T is contained in at least one k-dimensional simplex $\widehat{T}$ in $\mathcal{T}$. Suppose that such a simplex $\widehat{T}$ is given by $\widehat{T} = \langle\langle y^0, y^1, \ldots, y^{k-1}, y \rangle\rangle$. We will show that there exists an additional k-dimensional simplex $\widehat{T}' \neq \widehat{T}$ that contains T.

The simplex $\widehat{T}$ is the collection of all points that can be represented as a convex combination of its extreme points. For every $n \in \mathbb{N}$, the vector $z^n := \sum_{l=0}^{k-1} \frac{1-\frac{1}{n}}{k} y^l - \frac{1}{n} y$ is not in $\widehat{T}$, since the coefficient of y is a negative number (Exercise 23.8). The sequence $(z^n)_{n \in \mathbb{N}}$ converges to $z^* = \sum_{l=0}^{k-1} \frac{1}{k} y^l \in T$, which lies in the relative interior[3] of T.

Since T is not in the boundary of S, the vector z^* is in the relative interior of S, and therefore there is $n_0 \in \mathbb{N}$ such that $z^n \in S$ for every $n \geq n_0$. Since S is the union of all the k-dimensional simplices in $\mathcal{T}$ (Theorem 23.14), and since there are a finite number of such simplices, there exists a subsequence $(n_j)_{j \in \mathbb{N}}$ such that $(z^{n_j})_{j \in \mathbb{N}}$ are all contained in the same k-dimensional simplex $\widehat{T}' \in \mathcal{T}$, which differs from $\widehat{T}$. Since the simplex is a closed set, $z^* \in \widehat{T}'$, and therefore $z^* \in \widehat{T} \cap \widehat{T}'$. Since the smallest simplex (with respect to set inclusion) in $\mathcal{T}$ that contains z^* is T (why?), the intersection $\widehat{T} \cap \widehat{T}'$, which is a face both of $\widehat{T}$ and of $\widehat{T}'$, equals T. It follows that T is contained also in $\widehat{T}'$, which is what we wanted to show.

3 A point x in a simplex T is in the *relative interior* of T if all its barycentric coordinates relative to T are positive.

Step 5: If T is not in the boundary of S, then it is contained in exactly two k-dimensional simplices in $\mathcal{T}$.

Suppose by contradiction that there exists a third k-dimensional simplex $\widehat{T}'' = \langle\langle y^0, y^1, \ldots, y^{k-1}, y''\rangle\rangle \in \mathcal{T}$ containing T, in addition to $\widehat{T}$ and $\widehat{T}'$. As we saw in the first part of the proof,

$$H_{\widehat{T}} = H_{\widehat{T}'} = H_{\widehat{T}''} = H_S. \tag{23.30}$$

Therefore in particular $y'' \in H_{\widehat{T}'}$ and $y' \in H_{\widehat{T}}$. Write $y'' = \sum_{l=0}^{k-1} \alpha'^l y^l + \beta' y'$, where $\sum_{l=0}^{k-1} \alpha'^l + \beta' = 1$. Recall that $y' = \sum_{l=0}^{k-1} \alpha^l y^l + \beta y$ (Equation (23.23)), and that $\beta < 0$ since $\widehat{T}' \neq \widehat{T}$ (step 2 of the proof). Similarly, since $\widehat{T}' \neq \widehat{T}''$ we deduce that $\beta' < 0$. Hence

$$y'' = \sum_{l=0}^{k-1} \alpha'^l y^l + \beta' y' \tag{23.31}$$

$$= \sum_{l=0}^{k-1} \alpha'^l y^l + \beta' \left(\sum_{l=0}^{k-1} \alpha^l y^l + \beta y \right) \tag{23.32}$$

$$= \sum_{l=0}^{k-1} (\alpha'^l + \beta'^l) y^l + \beta' \beta y. \tag{23.33}$$

Since $\beta < 0$ and $\beta' < 0$, it follows that $\beta'\beta > 0$, and therefore by the results of Step 2, $\widehat{T}'' = \widehat{T}$, contradicting the assumption that $\widehat{T}'' \neq \widehat{T}$. The contradiction proves that there are exactly two k-dimensional simplices in $\mathcal{T}$ containing T. $\square$

Let $S = \langle\langle x^0, x^1, \ldots, x^k\rangle\rangle$ be a simplex in $\mathbb{R}^n$. By Theorem 23.9 (page 920), every vector $y \in S$ has a unique representation as a convex combination of $x^0, x^1, \ldots, x^k$,

$$\sum_{l=0}^{k} \alpha^l x^l = y, \tag{23.34}$$

where $(\alpha^l)_{l=0}^k$ are nonnegative numbers whose sum is 1 (the barycentric coordinates of y). Denote

$$\mathrm{supp}_S(y) = \{l : 0 \leq l \leq k, \quad \alpha^l > 0\}. \tag{23.35}$$

This set is called the *support* of y relative to S. This is the set of positive barycentric coordinates of y. Recall that if $\mathcal{T}$ is a simplicial partition of S, then $Y(\mathcal{T})$ is the set of all extreme points of the simplices in $\mathcal{T}$.

Definition 23.16 *Let $S = \langle\langle x^0, x^1, \ldots, x^k\rangle\rangle$ be a k-dimensional simplex in $\mathbb{R}^n$, for $n \geq k$, and let $\mathcal{T}$ be a simplicial partition of S. A* coloring *of $\mathcal{T}$ is a function $c : Y(\mathcal{T}) \to \{0, 1, \ldots, k\}$ associating every vertex y in $Y(\mathcal{T})$ with an index $c(y)$ in $\{0, 1, \ldots, k\}$ that is called the* color *of the vertex. A coloring c is called* proper *if for every $y \in Y(\mathcal{T})$ the color of y is one of the indices in the support of y (using its barycentric representation relative to S).*

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

Example 23.17 Three colorings of a simplicial partition of a two-dimensional simplex $S = \langle\langle x^0, x^1, x^2\rangle\rangle$ in $\mathbb{R}^2$ are depicted in Figure 23.3. The color of each vertex is noted next to it. Colorings A and B in Figure 23.3 are proper colorings. Coloring C in Figure 23.3 is not proper, because one of the vertices in the simplicial partition, which is contained in the simplex $\langle\langle x^0, x^2\rangle\rangle$ and is identified by an arrow, is colored by the color 1.

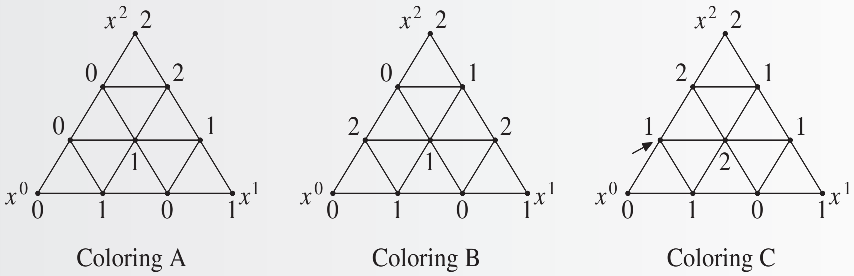


Figure 23.3 Examples of colorings

◀

Definition 23.18 *Let $\mathcal{T}$ be a simplicial partition of a k-dimensional simplex S, and let c be a proper coloring of $\mathcal{T}$. The k-dimensional simplex $T \in \mathcal{T}$ is* perfectly colored *if its vertices are colored with $k+1$ different colors* $\{0, 1, \ldots, k\}$.

Example 23.17 (*Continued*) Figure 23.4 depicts the colorings in Figure 23.3, with the two-dimensional perfectly colored simplices shaded in grey. In colorings A and B, which are proper colorings, the number of such simplices is indeed odd (1 in coloring A, and 5 in coloring B). When the coloring is not proper, the number of two-dimensional perfectly colored simplices may be even, as coloring C of the figure shows.

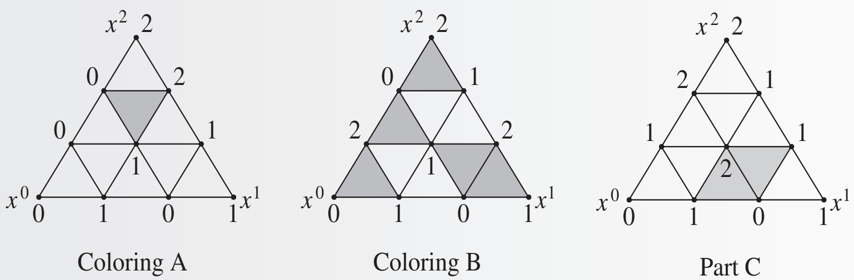


Figure 23.4 The perfectly colored simplices in Example 23.17

◀

Theorem 23.19 (Sperner's Lemma) *Let S be a k-dimensional simplex in $\mathbb{R}^n$, for $n \geq k$, and let $\mathcal{T}$ be a simplicial partition of S. Let c be a proper coloring of $\mathcal{T}$. Then the number of perfectly colored k-dimensional simplices $T \in \mathcal{T}$ is odd.*

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

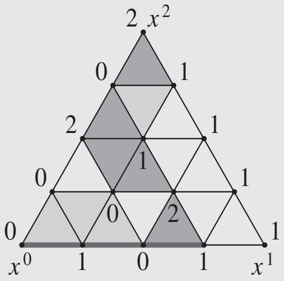


Figure 23.5 Collections of simplices $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$

In particular, there is at least one perfectly colored k-dimensional simplex in $\mathcal{T}$.

Proof: The proof is conducted by induction on the dimension of the simplex, k.

Step 1: The case $k = 0$.
In this case, the simplex contains only one point, $S = \langle\langle x^0 \rangle\rangle$, and the only simplicial partition is the one containing the entire simplex, $\mathcal{T} = \{S\}$. In this case, $Y(\mathcal{T}) = \{x^0\}$. The only proper coloring associates the vertex x^0 with the color 0. It follows that the number of perfectly colored zero-dimensional simplices is 1, which is an odd number.

Step 2: Defining collections of simplices when $k > 0$.
Let $k > 0$. Suppose that the statement of the theorem is true for every $(k-1)$-dimensional simplex, and let $S = \langle\langle x^0, x^1, \ldots, x^k \rangle\rangle$ be a k-dimensional simplex. Let $\mathcal{T}$ be a simplicial partition of S and c a proper coloring of $\mathcal{T}$. We will make use of the following notation:

- $\mathcal{A}$ is the collection of all $(k-1)$-dimensional simplices in $\mathcal{T}$ contained in the boundary of S and colored by $\{0, 1, \ldots, k-1\}$.
- $\mathcal{B}$ is the collection of all the k-dimensional simplices in $\mathcal{T}$ (not necessarily in the boundary of S) whose vertices are colored by $\{0, 1, \ldots, k-1\}$. In other words, these are the simplices colored by $\{0, 1, \ldots, k-1\}$ satisfying the property that two of their vertices are colored by the same color.
- $\mathcal{C}$ is the collection of all the k-dimensional simplices in $\mathcal{T}$ whose vertices are colored by $\{0, 1, \ldots, k\}$.

In the example in Figure 23.5, $\mathcal{A}$ has three one-dimensional simplices, denoted by thick lines. $\mathcal{B}$ has four two-dimensional simplices, denoted by light shading. $\mathcal{C}$ has five two-dimensional simplices, denoted by darker shading.

Step 3: The number of simplices in $\mathcal{A}$ is odd.
First, we show that every simplex in $\mathcal{A}$ is contained in the face $\langle\langle x^0, x^1, \ldots, x^{k-1} \rangle\rangle$ of S. Recall that the boundary of the simplex $\langle\langle x^0, x^1, \ldots, x^k \rangle\rangle$ is the union of the $(k-1)$-dimensional faces,

$$\bigcup_{l=0}^{k} \langle\langle x^0, x^1, \ldots, x^{l-1}, x^{l+1}, \ldots, x^k \rangle\rangle.$$

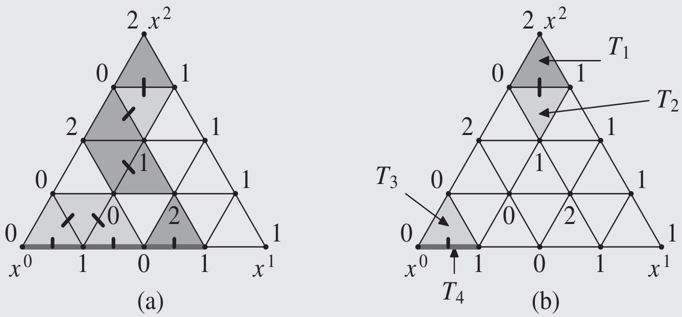


Figure 23.6 The collections $\mathcal{A}$ (dark lines), $\mathcal{B}$ (light triangles), and $\mathcal{C}$ (dark triangles) in Figure 23.5 and their associated graphs

It follows that if a vertex is colored by a color $l \in \{0, 1, \dots, k-1\}$, then it cannot be on the face $\langle\langle x^0, x^1, \dots, x^{l-1}, x^{l+1}, \dots, x^k \rangle\rangle$, because the coloring is proper, and the index l is not in the support of any vector in this face of S. Since every simplex in $\mathcal{A}$ is colored by all the colors $0, 1, \dots, k-1$, such a simplex cannot be in $\bigcup_{l=0}^{k-1} \langle\langle x^0, x^1, \dots, x^{l-1}, x^{l+1}, \dots, x^k \rangle\rangle$, and it must therefore be in the face $\langle\langle x^0, x^1, \dots, x^{k-1} \rangle\rangle$ of S.

We therefore deduce that the simplices in $\mathcal{A}$ are exactly the $(k-1)$-dimensional simplices contained in the simplex $\langle\langle x^0, x^1, \dots, x^{k-1} \rangle\rangle$ whose vertices are colored by $\{0, 1, \dots, k-1\}$. By the induction hypothesis, the number of such simplices is odd.

Step 4: Completing the proof.
Define an undirected graph as follows:

- The set of vertices is $\mathcal{A} \cup \mathcal{B} \cup \mathcal{C}$; i.e., every simplex in the union $\mathcal{A} \cup \mathcal{B} \cup \mathcal{C}$ is a vertex of the graph.
- Let T_1 and T_2 be two different simplices in $\mathcal{A} \cup \mathcal{B} \cup \mathcal{C}$. Then there exists an edge connecting T_1 and T_2 if and only if the intersection $T_1 \cap T_2$ is a $(k-1)$-dimensional simplex whose vertices are colored by $\{0, 1, \dots, k-1\}$.

The graph corresponding to the coloring in Figure 23.5 appears in Figure 23.6(a). In this figure, the one-dimensional simplices outlined with dark lines and the two-dimensional colored simplices are vertices of the graph. The lines connecting pairs of simplices denote the edges. Figure 23.6(b) illustrates the rule according to which the edges of the graph are determined. For example, the simplices T_1 and T_2 denoted in Figure 23.6(b) are two vertices of the graph: T_1 is in $\mathcal{C}$ and T_2 is in $\mathcal{B}$. Their intersection is a one-dimensional simplex that is colored by the colors 0 and 1, and therefore there is an edge connecting T_1 and T_2. The simplices T_3 and T_4, depicted in Figure 23.6(b), are also vertices in the graph: T_3 is in B and T_4 is in $\mathcal{A}$. Their intersection is the simplex T_4, which is a one-dimensional simplex colored by the colors 0 and 1, and hence there is an edge connecting T_3 and T_4. This explains all the edges in Figure 23.6(a).

Denote by R the number of edges in the graph. Using Theorem 23.15, we can count how many edges emanate from each vertex in the graph.

- There is only one edge emanating from each vertex in $\mathcal{A}$ to a vertex in $\mathcal{B}$ or $\mathcal{C}$. Indeed, by Theorem 23.15, every simplex $T = \langle\!\langle y^0, y^1, \ldots, y^{k-1}\rangle\!\rangle \in \mathcal{A}$ is contained in a single k-dimensional simplex $\widehat{T} = \langle\!\langle y^0, y^1, \ldots, y^{k-1}, y^k\rangle\!\rangle$ in $\mathcal{T}$: if the color of vertex x^k is k, the simplex $\widehat{T}$ is contained in $\mathcal{C}$. If not, $\widehat{T}$ is in $\mathcal{B}$. In any case, $\widehat{T} \cap T = T$, and since T is a $(k-1)$-dimensional simplex colored with $\{0, 1, \ldots, k-1\}$ there is an edge connecting T and $\widehat{T}$.
- Two edges emanate from every vertex in $\mathcal{B}$. To see this, suppose that $T = \langle\!\langle x^0, x^1, \ldots, x^{k-1}, x^k\rangle\!\rangle \in \mathcal{B}$ and assume without loss of generality that the vertices x^{k-1} and x^k are colored in the same color. The only two $(k-1)$-dimensional subsimplices of T that are colored in the colors $\{0, 1, \ldots, k-1\}$ are $T_1 := \langle\!\langle x^0, x^1, \ldots, x^{k-2}, x^{k-1}\rangle\!\rangle$ and $T_2 := \langle\!\langle x^0, x^1, \ldots, x^{k-2}, x^k\rangle\!\rangle$. For $i = 1, 2$, consider the simplex T_i. Using Theorem 23.15 we deduce that there are two possibilities.
  - In this case, T_i is in $\mathcal{A}$ and there is an edge connecting T_i and T (because $T \cap T_i = T_i$).
  - T_i is not in the boundary of S. It is then contained in two k-dimensional simplices, one of which is T; denote the other by $\widehat{T_i}$. If the simplex $\widehat{T_i}$ is colored in the colors $\{0, 1, \ldots, k\}$, then $\widehat{T_i} \in \mathcal{C}$. If not, $\widehat{T_i} \in \mathcal{B}$. In both cases there is an edge connecting T and $\widehat{T_i}$ (because $\widehat{T_i} \cap T = T$).

  We conclude that there are two edges emanating from T.
- There is only one edge emanating from every vertex in $\mathcal{C}$. To see this, suppose that $T = \langle\!\langle x^0, x^1, \ldots, x^{k-1}, x^k\rangle\!\rangle \in \mathcal{C}$. Since all the vertices are colored in different colors $\{0, 1, \ldots, k\}$, we will suppose without loss of generality that the subsimplex $T_1 = \langle\!\langle x^0, x^1, \ldots, x^{k-1}\rangle\!\rangle$ is colored with the colors $\{0, 1, \ldots, k-1\}$ (and this is the only $(k-1)$-dimensional subsimplex in T_1 colored in these colors). If T_1 is in the boundary of S, then it is in A and there is an edge connecting T and T_1 (because $T_1 \cap T = T_1$). If T_1 is not in the boundary of S, by Theorem 23.15 the simplex T_1 is contained in two k-dimensional simplices. One of those simplices is T. As in the previous case, the second k-dimensional simplex T_2 containing T_1 is either in $\mathcal{B}$ or in $\mathcal{C}$, and there is an edge connecting T_2 and T (because $T_2 \cap T = T_1$).

The sum total of edges emanating from all the vertices in the graph is twice the number of edges R, because every edge is counted twice. It follows that

$$2R = |\mathcal{A}| + 2|\mathcal{B}| + |\mathcal{C}|.$$

This implies that $|\mathcal{A}| + |\mathcal{C}|$ is an even number. Since the number of elements in A is odd (by Step 3), the number of elements in C must be odd. In other words, the number of perfectly colored k-dimensional simplices in $\mathcal{T}$ is odd. This establishes the statement for k-dimensional simplices, concluding the induction step and the proof of Sperner's Lemma. □

Definition 23.20 *Let $S = \langle\!\langle x^0, x^1, \ldots, x^k\rangle\!\rangle$ be a simplex in $\mathbb{R}^n$. The* diameter *of S, denoted by $\rho(S)$, is defined as*[4]

$$\rho(S) = \max_{0 \le i < j \le k} \|x^i - x^j\|. \tag{23.36}$$

4 Recall that $\|x^i - x^j\|$ is the Euclidean distance in $\mathbb{R}^n$ between the vectors x^i and x^j.

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

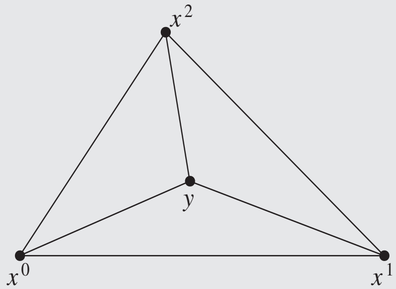


Figure 23.7 The collection $\mathcal{T}$ described in Theorem 23.23, for a two-dimensional simplex

The diameter of a simplex is also equal to the greatest (Euclidean) distance between two points in the simplex (not necessarily vertices; see Exercise 23.22).

Definition 23.21 *Let $\mathcal{T} = \{T_1, T_2, \dots, T_M\}$ be a simplicial partition of a simplex S. The* diameter *of a simplicial partition $\mathcal{T}$ is denoted by $\rho(\mathcal{T})$ and defined by*

$$\rho(\mathcal{T}) = \max_{m=1,..,M} \rho(T_m). \tag{23.37}$$

The diameter of $\mathcal{T}$ is the greatest (Euclidean) distance between two points located in the same simplex in $\mathcal{T}$. One simplicial partition of a simplex S is the partition that contains all faces of S. The next theorem enables us to deduce that every simplicial partition can be refined to a simplicial partition with an arbitrarily small diameter.

Theorem 23.22 *Let $k \ge 1$ and let $\mathcal{T}$ be a simplicial partition of a k-dimensional simplex $S = \langle\langle x^0, x^1, \dots, x^k \rangle\rangle$ in $\mathbb{R}^n$. Then there exists a simplicial partition $\mathcal{T}'$ of S satisfying the following properties:*

(i) For every $T' \in \mathcal{T}'$ there exists $T \in \mathcal{T}$ such that $T' \subseteq T$.
(ii) $\rho(\mathcal{T}') \le \frac{k}{k+1}\rho(\mathcal{T})$.

For proving Theorem 23.22 we make use the following two auxiliary theorems.

Theorem 23.23 *Let $S = \langle\langle x^0, x^1, \dots, x^k \rangle\rangle$ be a k-dimensional simplex in $\mathbb{R}^n$, and denote $y = \frac{1}{k+1}\sum_{l=0}^{k} x^l$. Then the collection $\mathcal{T}$ containing the $k+1$ simplices $\langle\langle x^0, x^1, \dots, x^{l-1}, x^{l+1}, \dots, x^k, y \rangle\rangle$, $l = 0, 1, \dots, k$, and all of their faces, is a simplicial partition of S.*

For a two-dimensional simplex $S = \langle\langle x^0, x^1, x^2 \rangle\rangle$, the collection $\mathcal{T}$ defined in Theorem 23.23 is depicted in Figure 23.7, and it is indeed a simplicial partition of S.

Proof: By Theorem 23.7 on page 920, for every $l = 0, 1, \dots, k$, the vectors $x^0, x^1, \dots, x^{l-1}, x^{l+1}, \dots, x^k, y$ are affine-independent vectors (see Example 23.5), and

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

$\mathcal{T}$ is therefore a collection of simplices. We now show that

$$S = \bigcup_{i=0}^{k} \langle\langle x^0, x^1, \ldots, x^{l-1}, x^{l+1}, \ldots, x^k, y\rangle\rangle. \tag{23.38}$$

Let $z \in S$. Since S is a simplex, there exist nonnegative numbers $\alpha^0, \alpha^1, \ldots, \alpha^k$ whose sum is 1 such that $z = \sum_{l=0}^k \alpha^l x^l$. Suppose that $\alpha^j = \min_{l=0,1,\ldots,k} \alpha^l$. We will show that the vector z is contained in the simplex $\langle\langle x^0, x^1, \ldots, x^{j-1}, x^{j+1}, \ldots, x^k, y\rangle\rangle$. Since α^j is minimal, we have $\alpha^j \le \frac{1}{k+1}$, or equivalently that $(k+1)\alpha^j \le 1$. Therefore,

$$z = \sum_{l=0}^{k} \alpha^l x^l = \alpha^j \sum_{l=0}^{k} x^l + \sum_{l\neq j} (\alpha^l - \alpha^j) x^l \tag{23.39}$$

$$= (k+1)\alpha^j \frac{1}{k+1} \sum_{l=0}^{k} x^l + \sum_{l\neq j} (\alpha^l - \alpha^j) x^l \tag{23.40}$$

$$= (k+1)\alpha^j y + \sum_{l\neq j} (\alpha^l - \alpha^j) x^l. \tag{23.41}$$

Since $\alpha^l - \alpha^j \ge 0$ for all $l \neq j$, and

$$(k+1)\alpha^j + \sum_{l\neq j} (\alpha^l - \alpha^j) = \sum_{l=0}^{k} \alpha^l = 1, \tag{23.42}$$

it follows that $z \in \langle\langle x^0, x^1, \ldots, x^{j-1}, x^{j+1}, \ldots, x^k, y\rangle\rangle$, which is what we wanted to show. We leave it to the reader to ascertain that any pair of simplices in $\mathcal{T}$ are either disjoint, or their intersection is also a simplex in $\mathcal{T}$ (Exercise 23.23). □

Remark 23.24 *For the simplicial partition described in Theorem 23.23, we compute here an upper bound to the value of $\|x^j - y\|$, which we will need later. Let $j \in \{0, 1, \ldots, k\}$. By the triangle inequality,*

$$\|x^j - y\| = \left\| x^j - \frac{1}{k+1} \sum_{l=0}^{k} x^l \right\| = \left\| \frac{1}{k+1} \sum_{l\neq j} (x^j - x^l) \right\| \tag{23.43}$$

$$\le \frac{1}{k+1} \sum_{l\neq j} \|x^j - x^l\| \le \frac{k}{k+1} \max_{l\neq j} \|x^j - x^l\| = \frac{k}{k+1} \rho(S). \tag{23.44}$$

♦

For every k-dimensional simplex $S = \langle\langle x^0, x^1, \ldots, x^k\rangle\rangle$ in $\mathbb{R}^n$ and every vector $y \in \mathbb{R}^n$ that is affine independent of the vectors $x^0, x^1, \ldots, x^k$, denote by $\langle\langle S, y\rangle\rangle$ the $(k+1)$-dimensional simplex whose extreme points are y and the extreme points of S,

$$\langle\langle S, y\rangle\rangle := \langle\langle x^0, x^1, \ldots, x^k, y\rangle\rangle. \tag{23.45}$$

Theorem 23.25 *Let $S = \langle\langle x^0, x^1, \ldots, x^k\rangle\rangle$ be a k-dimensional simplex in $\mathbb{R}^n$, and let $\mathcal{T}$ be a simplicial partition of S. Let $y \in \mathbb{R}^n$ be a vector that is affine independent of the*

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

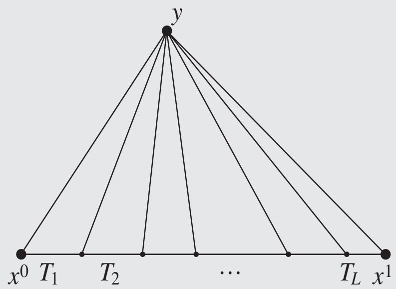


Figure 23.8 The collection $\widehat{\mathcal{T}}$ described in Theorem 23.25, for the one-dimensional simplex $S = \langle\langle x^0, x^1 \rangle\rangle$

vectors $x^0, x^1, \ldots, x^k$, and let $\widehat{S} = \langle\langle S, y \rangle\rangle$. Then the following collection of simplices $\widehat{\mathcal{T}}$ is a simplicial partition of the simplex $\widehat{S}$,

$$\widehat{\mathcal{T}} := \mathcal{T} \bigcup \{\langle\langle y \rangle\rangle\} \bigcup \{\langle\langle T, y \rangle\rangle : T \in \mathcal{T}\}. \tag{23.46}$$

Figure 23.8 corresponds to the one-dimensional case of the theorem, i.e., when $S = \langle\langle x^0, x^1 \rangle\rangle$ is a one-dimensional simplex (the base of the triangle in the figure) and the partition $\mathcal{T}$ of S is composed of the points and intervals at the base of the triangle in the figure. The collection $\widehat{\mathcal{T}}$ defined in Theorem 23.25 contains, in addition to the simplices in $\mathcal{T}$, the triangles appearing in the figure, their sides, and the upper vertex.

Proof: We prove that the union of the simplices in the collection $\widehat{\mathcal{T}}$ equals $\widehat{S}$. Ascertaining that the intersection of any two simplices in the collection $\widehat{\mathcal{T}}$ is either the empty set, or is a simplex in $\widehat{\mathcal{T}}$, is left to the reader (Exercise 23.24). Let z be a vector in $\widehat{S}$. Since $\widehat{S}$ is a simplex, there exist nonnegative numbers $(\alpha^i)_{i=0}^{k+1}$ whose sum is 1 such that $z = \sum_{i=0}^k \alpha^i x^i + \alpha^{k+1} y$. If $\alpha^{k+1} = 1$ then $z = y$, and therefore $z \in \langle\langle y \rangle\rangle$. Since $\langle\langle y \rangle\rangle$ is a simplex in $\widehat{\mathcal{T}}$, the vector z is in the union of the simplices in $\widehat{\mathcal{T}}$. If $\alpha^{k+1} < 1$, define $\widehat{z} := \sum_{i=0}^k \frac{\alpha^i}{1-\alpha^{k+1}} x^k$. Since the numbers $\{\frac{\alpha^i}{1-\alpha^{k+1}}\}_{i=0}^k$ are nonnegative numbers whose sum is 1, $\widehat{z} \in S$, and therefore it is contained in one of the simplices T in the simplicial partition $\mathcal{T}$. Since $z = (1-\alpha^{k+1})\widehat{z} + \alpha^{k+1} y$, it follows that $z \in \langle\langle T, y \rangle\rangle$, and therefore in this case the vector z is in the union of the simplices in $\widehat{\mathcal{T}}$ as well. □

The triangle inequality implies that the diameter of the partition $\widehat{\mathcal{T}}$ constructed in Theorem 23.25 is

$$\rho(\widehat{\mathcal{T}}) = \max\{\rho(\mathcal{T}), \|x^0 - y\|, \|x^1 - y\|, \ldots, \|x^k - y\|\}; \tag{23.47}$$

see Exercise 23.25.

Proof of Theorem 23.22: The proof is by induction on k.

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

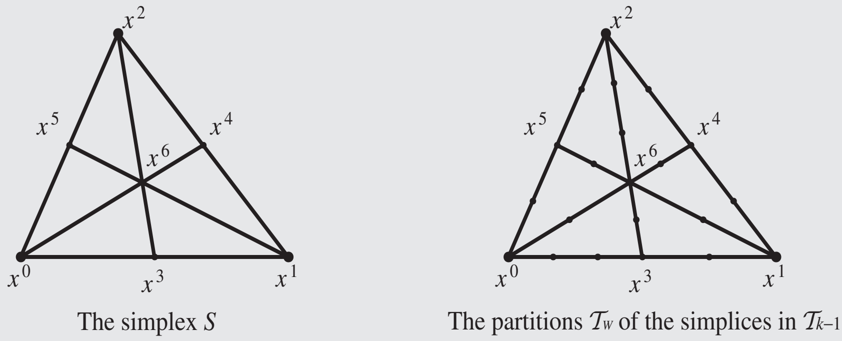


Figure 23.9 The case $k = 2$ in the proof of Theorem 23.22: the simplex S and the partition $(\mathcal{T}_W)_{W \in \mathcal{T}_{k-1}}$

Step 1: The case $k = 1$.
When $k = 1$ the simplex S is the interval $S = \langle\langle x^0, x^1 \rangle\rangle$ and the simplicial partition is given by a finite number of points in this interval; i.e., there exist points $z^1, z^2, \ldots, z^L$ such that

$$\mathcal{T} = \{\langle\langle x^0, z^1\rangle\rangle, \langle\langle z^1, z^2\rangle\rangle, \ldots, \langle\langle z^{L-1}, z^L\rangle\rangle, \langle\langle z^L, x^1\rangle\rangle, \langle\langle x^0\rangle\rangle, \langle\langle z^1\rangle\rangle, \ldots, \langle\langle z^L\rangle\rangle, \langle\langle x^1\rangle\rangle\}. \tag{23.48}$$

Define a new simplicial partition $\mathcal{T}'$ by partitioning each one of the one-dimensional simplices in $\mathcal{T}$ into two equal parts: for each $l \in \{1, \ldots, L+1\}$ let $y^l = \frac{1}{2}z^{l-1} + \frac{1}{2}z^l$ where $z^0 := x^0$ and $z^{L+1} := x^1$. The simplicial partition $\mathcal{T}'$ is therefore

$$\begin{aligned}\mathcal{T}' = \{&\langle\langle x^0, y^1\rangle\rangle, \langle\langle y^1, z^1\rangle\rangle, \langle\langle z^1, y^2\rangle\rangle, \langle\langle y^2, z^2\rangle\rangle, \ldots, \langle\langle y^L, x^1\rangle\rangle, \langle\langle x^0\rangle\rangle,\\ &\langle\langle y^1\rangle\rangle, \langle\langle z^1\rangle\rangle, \ldots, \langle\langle y^L\rangle\rangle, \langle\langle x^1\rangle\rangle\}.\end{aligned} \tag{23.49}$$

Then $\mathcal{T}'$ is a refinement of $\mathcal{T}$, and its diameter is half the diameter of $\mathcal{T}$.

Step 2: The case $k > 1$.
Suppose that the claim of the theorem holds for every $(k-1)$-dimensional simplex. We will prove that it then holds for every k-dimensional simplex. Let $\mathcal{T}$ be a simplicial partition of a k-dimensional simplex S. Let $\mathcal{T}_{k-1}$ be the collection of all the $(k-1)$-dimensional simplices in $\mathcal{T}$ and let $\mathcal{T}_k$ be the collection of all the k-dimensional simplices in $\mathcal{T}$. The construction of the desired simplicial partition $\mathcal{T}'$ of S is conducted in several steps:

- Using the induction hypothesis, we will prove that for every $(k-1)$-dimensional simplex $W \in \mathcal{T}_{k-1}$ there exists a simplicial partition $\mathcal{T}_W$ of W whose diameter is less than or equal to $\frac{k-1}{k}\rho(\mathcal{T})$. Figure 23.9 illustrates the simplex S in the case that $k = 2$ and the simplicial partition is $(\mathcal{T}_W)_{W \in \mathcal{T}_{k-1}}$.

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

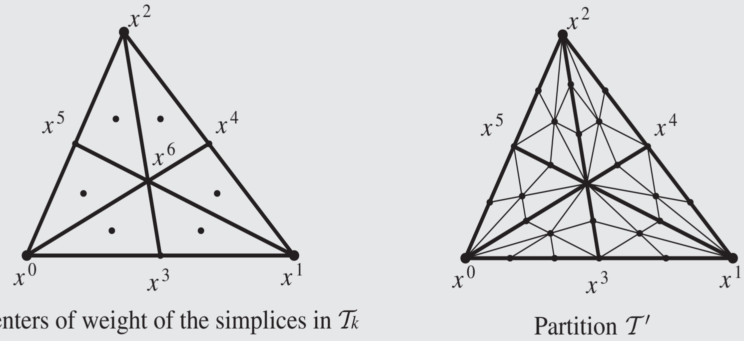


Figure 23.10 The case $k = 2$ in the proof of Theorem 23.22: centers of weight and the subdivision $\mathcal{T}'$

- For every k-dimensional simplex $T \in \mathcal{T}$ denote by y_T its center of weight. These centers of weight are illustrated in Figure 23.10.
- Every simplex $T \in \mathcal{T}$ can be subdivided into the simplices $\langle\langle R, y_T\rangle\rangle$, for all the simplices $R \in \mathcal{T}_W$, for each face W of T. This partition is illustrated in Figure 23.10.
- Since $\cup_{T\in\mathcal{T}_k} T = S$, we obtain a refinement of the simplicial partition $\mathcal{T}$ of S.
- Finally, we prove that the diameter of this simplicial partition is less than or equal to $\frac{k}{k+1}\rho(\mathcal{T})$.

We next turn to the formal construction of the simplicial division that refines $\mathcal{T}$. For each $(k-1)$-dimensional simplex $W \in \mathcal{T}_{k-1}$, let $\widehat{\mathcal{T}}_W$ be the trivial partition containing all the faces of W. Since W is a simplex in the simplicial partition T one has $\rho(\widehat{\mathcal{T}}_W) \leq \rho(\mathcal{T})$. By the induction hypothesis there exists a simplicial partition $\mathcal{T}_W$ of W satisfying

$$\rho(\mathcal{T}_W) \leq \frac{k-1}{k}\rho(\widehat{\mathcal{T}}_W) \leq \frac{k-1}{k}\rho(\mathcal{T}). \tag{23.50}$$

For each k-dimensional simplex $T = \langle\langle y^0, y^1, \ldots, y^k\rangle\rangle \in \mathcal{T}_k$, let y_T be its center of weight:

$$y_T := \sum_{l=0}^{k} \frac{1}{k+1} y^l. \tag{23.51}$$

Let $\mathcal{T}'$ be the collection of all the simplices $\langle\langle R, y_T\rangle\rangle$ and their faces, where $T \in \mathcal{T}_k$, $W \in \mathcal{T}_{k-1}$ is a face of T and $R \in \mathcal{T}_W$. To conclude the proof, we need to show that the collection $\mathcal{T}'$ is a simplicial partition and that $\rho(\mathcal{T}') \leq \frac{k}{k+1}\rho(\mathcal{T})$. By construction, Properties (3) and (4) in the definition of simplicial partitions hold for the collection $\mathcal{T}'$. By Theorems 23.14 and 23.23 it follows that the union of the simplices in $\mathcal{T}'$ is S. We next show that the intersection of two simplices in $\mathcal{T}'$ is either empty or a face of both of them. Let R and $\widehat{R}$ be two simplices in $\mathcal{T}'$ whose intersection is nonempty.

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

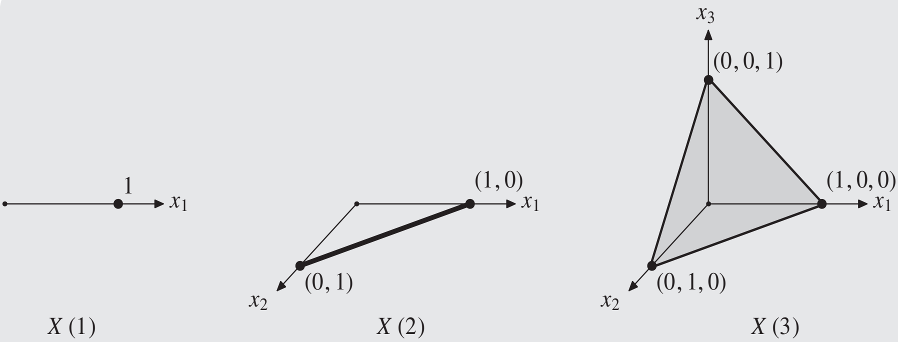


Figure 23.11 The standard simplices $X(1)$, $X(2)$, and $X(3)$

- If R and $\widehat{R}$ are contained in the same k-dimensional simplex T in $\mathcal{T}$, Theorems 23.23 and 23.25 imply that R and $\widehat{R}$ are elements of a simplicial partition of T, and therefore in particular $R \cap \widehat{R}$ is a face of both R and $\widehat{R}$.
- Assume that R and $\widehat{R}$ are contained in two different k-dimensional simplices in $\mathcal{T}$, which we denote as T and $\widehat{T}$ respectively. Since $\mathcal{T}$ is a simplicial partition, the intersection $T \cap \widehat{T}$ is a face both of T and of $\widehat{T}$, and it is a simplex of a dimension smaller than k. Let W be a $(k-1)$-dimensional simplex in $\mathcal{T}$ containing $T \cap \widehat{T}$. Since $\mathcal{T}_W$ is a simplicial partition, it follows by construction that $R \cap \widehat{R}$ is an element in $\mathcal{T}_W$, and that this element is a face of both R and $\widehat{R}$.

By construction, and with the use of Equations (23.43)–(23.44), we deduce that the diameter of $\mathcal{T}'$ is given by

$$\rho(\mathcal{T}') = \max\{\max\{\rho(\mathcal{T}_W), W \in \mathcal{T}_{k-1}\}, \max\{\|x^l - y\|, 0 \leq l \leq k\}\} \tag{23.52}$$

$$\leq \max\left\{\frac{k-1}{k}\rho(\mathcal{T}), \frac{k}{k+1}\rho(\mathcal{T})\right\} = \frac{k}{k+1}\rho(\mathcal{T}), \tag{23.53}$$

which is what we needed to show. This completes the proof of Theorem 23.22. □

By repeated use of Theorem 23.22 we obtain the following corollary.

Corollary 23.26 *For every simplex S, and every $\varepsilon > 0$, there exists a simplicial partition $\mathcal{T}_\varepsilon$ of S satisfying $\rho(\mathcal{T}_\varepsilon) \leq \varepsilon$.*

23.1.2 Brouwer's Fixed Point Theorem

For every $i = 1, 2, \ldots, n$, the $(n-1)$-*dimensional standard simplex* is the simplex in $\mathbb{R}^n$ whose vertices are the unit vectors $e^1, e^2, \ldots, e^n$. This is the set $X(n) \subseteq \mathbb{R}^n$ defined as

$$X(n) := \left\{x \in \mathbb{R}^n : \sum_{i=1}^{n} x_i = 1, \quad x_i \geq 0 \quad \forall i\right\}. \tag{23.54}$$

The standard simplices $X(1)$, $X(2)$, and $X(3)$ are depicted in Figure 23.11.

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

When vectors in a simplex $S = \langle\langle x^0, x^1, \ldots, x^k \rangle\rangle$ are presented in barycentric coordinates, the simplex $X(k+1)$ is obtained. To see this, note that for each l, $0 \leq l \leq k$, the barycentric coordinates of x^l are the unit vector e^l and every point in the simplex S can be uniquely presented as a convex combination of extreme points in the simplex (Theorem 23.9).

Brouwer's Fixed Point Theorem states that every continuous function from a convex and compact set X in $\mathbb{R}^n$ to itself has a fixed point. We first prove the theorem for the case in which the convex and compact set is $X(n)$. The statement of the theorem refers to a function defined on a standard simplex, but since every k-dimensional simplex S is equivalent to the standard simplex $X(k+1)$ when its points are represented in barycentric coordinates, the theorem holds for functions defined on any simplex.

Theorem 23.27 (Special case of Brouwer's Fixed Point Theorem) *Let $f : X(n) \to X(n)$ be a continuous function. Then there exists $x^* \in X(n)$ such that $f(x^*) = x^*$.*

In the proof of the theorem we will use the sup-norm in $\mathbb{R}^n$: for each vector $x \in \mathbb{R}^n$ denote

$$\|x\|_\infty := \max_{i=1,\ldots,n} |x_i|. \tag{23.55}$$

Proof: Since $y = \sum_{i=1}^n y_i e^i$ for every $y \in X(n)$ where $(e^i)_{i=1}^n$ are the unit vectors (see Example 23.4 on page 918), the support of y relative to the vertices of the simplex $X(n)$ is the collection of the indices i for which $y_i > 0$.

Step 1: For every $y \in X(n)$ there exists an index $i \in \operatorname{supp}(y)$ satisfying $f_i(y) \leq y_i$.
Let $y \in X(n)$, and suppose by contradiction that the claim does not hold. Then:

- $f_i(y) > y_i$ for every index $i \in \operatorname{supp}(y)$.
- $f_i(y) \geq 0 = y_i$ for every index $i \notin \operatorname{supp}(y)$.

Sum together these equations for $i = 1, 2, \ldots, n$. Since $\operatorname{supp}(y)$ contains at least one index, it follows that

$$1 = \sum_{i=1}^n f_i(y) > \sum_{i=1}^n y_i = 1. \tag{23.56}$$

This contradiction shows that the initial assumption was wrong, and therefore there exists an index $i \in \operatorname{supp}(y)$ satisfying $f_i(y) \leq y_i$.

Step 2: Defining a coloring.
Recall that every continuous function defined on a compact space is uniformly continuous.[5] Let $\varepsilon > 0$. Since f is uniformly continuous, there exists $\delta = \delta(\varepsilon) > 0$ such that if $\|x - y\|_\infty \leq \delta$, then $\|f(x) - f(y)\|_\infty \leq \varepsilon$. Note that δ may be chosen to be sufficiently small so as to be smaller than ε.

5 Let X be a subset of $\mathbb{R}^n$. A function $f : X \to \mathbb{R}^n$ is *uniformly continuous* if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\|f(x) - f(y)\|_\infty \leq \varepsilon$ for every $x, y \in X$ satisfying $\|x - y\|_\infty \leq \delta$.

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

Color the simplicial partition $\mathcal{T}_\varepsilon$ with the colors $\{1, 2, \ldots, n\}$ corresponding to the vertices of $X(n)$, which are $e^1, e^2, \ldots, e^n$, as follows. For every $y \in Y(\mathcal{T}_\varepsilon)$, let the color $c(y)$ be an index i satisfying $i \in \operatorname{supp}(y)$ and $f_i(y) \leq y_i$ (recall that in Step 1 we showed that there exists at least one such index).

Since the color of every $y \in Y(\mathcal{T}_\varepsilon)$ is an index in $\operatorname{supp}(y)$, this coloring satisfies the condition of Sperner's Lemma.

Step 3: Existence of a fixed point.
By Sperner's Lemma (Theorem 23.19 on page 926) there exists a simplex $T_\varepsilon = \langle\langle x^1, x^2, \ldots, x^n \rangle\rangle \in \mathcal{T}_\varepsilon$, all of whose vertices are colored by $\{1, 2, \ldots, n\}$. Suppose without loss of generality that $c(x^i) = i$ for each $i = 1, 2, \ldots, n$, i.e., $f_i(x^i) \leq x_i^i$ (x_i^i is the i-th coordinate of x^i). Let x^ε be a vector in T_ε. Since the diameter of $\mathcal{T}_\varepsilon$ is at most δ, it follows that $\|x^\varepsilon - x^i\|_\infty \leq \delta$ for each i, and since f is uniformly continuous, it follows that

$$f_i(x^\varepsilon) \leq f_i(x^i) + \varepsilon \leq x_i^i + \varepsilon \leq x_i^\varepsilon + \varepsilon + \delta \leq x_i^\varepsilon + 2\varepsilon, \quad \forall i = 1, 2, \ldots, n. \tag{23.57}$$

The last inequality holds because $\delta \leq \varepsilon$. This is true for every $\varepsilon > 0$, and therefore for every $\varepsilon > 0$ there exists a point $x^\varepsilon \in X(n)$ satisfying Equation (23.57), that is,

$$f_i(x^\varepsilon) \leq x_i^\varepsilon + 2\varepsilon, \quad \forall i = 1, 2, \ldots, n. \tag{23.58}$$

Since $X(n)$ is a compact set, there exists a sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ converging to 0 such that the sequence $(x^{\varepsilon_k})_{k=1}^\infty$ converges to a limit in $X(n)$, denoted by x^*. Since f is continuous, by taking the limit in Equation (23.58) one has

$$f_i(x^*) \leq x_i^*, \quad \forall i = 1, 2, \ldots, n, \tag{23.59}$$

and therefore

$$1 = \sum_{i=1}^n f_i(x^*) \leq \sum_{i=1}^n x_i^* = 1. \tag{23.60}$$

If there existed i, $1 \leq i \leq n$, for which the inequality in Equation (23.59) were a strict inequality, the inequality in Equation (23.60) would also be a strict inequality, which is impossible, and therefore $f_i(x^*) = x_i^*$ for all $i = 1, 2, \ldots, n$; i.e., x^* is a fixed point of f. □

Brouwer's Fixed Point Theorem (Theorem 23.27 on page 936) will now be generalized to a convex and compact set $X \subset \mathbb{R}^n$. We first show that for every point x that is not in a closed and convex set X there exists a unique point in X that is closest to it.

Recall that the distance of a point in $y \in \mathbb{R}^n$ from a set $X \subseteq \mathbb{R}^n$ is defined by

$$d(y, X) := \inf_{x\in X} d(y, x), \tag{23.61}$$

where $d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$ is the Euclidean distance between x and y.

Theorem 23.28 *Let $X \subset \mathbb{R}^n$ be a closed and convex set, and let $x \notin X$. Then there exists a unique point $y \in X$ such that $d(x, X) = d(x, y)$.*

Proof: Since the set X is closed and convex, there is at least one point $y \in X$ satisfying $d(x, X) = d(x, y)$. Suppose by contradiction that there are two such points y and $\widehat{y}$. Since

X is a convex set, the point $\frac{y+\widehat{y}}{2}$ is contained in X. We will now show that $d(x, \frac{y+\widehat{y}}{2}) < d(x, y) = d(x, \widehat{y})$, in contradiction to the assumption that y and $\widehat{y}$ are points in X whose distance from x is minimal.

Define a function $D : \mathbb{R} \to \mathbb{R}$ by

$$D(\alpha) = d(x, \alpha y + (1-\alpha)\widehat{y}) \tag{23.62}$$

$$= \sum_{i=1}^{n} (x_i - \alpha y_i - (1-\alpha)\widehat{y}_i)^2 = \sum_{i=1}^{n} (x_i - \widehat{y}_i - \alpha(y_i - \widehat{y}_i))^2. \tag{23.63}$$

This is a nonnegative and nonconstant quadratic function of α. Since $d(x, y) = d(x, \widehat{y})$, one has $D(0) = D(1)$. The coefficient of α^2 in this function is positive, and therefore it has a unique minimum attained at the point $\alpha = \frac{1}{2}$. In other words, $d(x, \frac{y+\widehat{y}}{2}) < d(x, y)$, as we wanted to show. □

Let $X \subset \mathbb{R}^n$ be a closed and convex set. The following theorem, which is proved using the triangle inequality, states that the function associating each point in $\mathbb{R}^n$ with the point in X that is closest to it is a Lipschitz[6] function, and therefore, in particular, it is continuous (Exercise 23.27).

Theorem 23.29 *Let $X \subset \mathbb{R}^n$ be a closed and convex set. Define a function $g : \mathbb{R}^n \to X$ as follows: $g(x)$ is the closest point in X to x. Then $d(g(x), g(\widehat{x})) \le d(x, \widehat{x})$ for every $x, \widehat{x} \in \mathbb{R}^n$.*

This last result is used to prove the following theorem.

Theorem 23.30 (Brouwer's Fixed Point Theorem) *Let $X \subseteq \mathbb{R}^n$ be a convex, compact, and nonempty set. Then every continuous function $f : X \to X$ has a fixed point.*

Proof: Since X is compact, there exists a simplex S sufficiently large to contain X. Define a function $h : S \to S$ by

$$h(x) := f(g(x)), \tag{23.64}$$

where g is the function associating every point x with the point in X that is closest to it. Note that the function g is well defined by Theorem 23.28 and that it satisfies $g(x) = x$ for every $x \in X$. By Theorem 23.29, the function g is continuous; hence h, as the composition of two continuous functions, is also continuous. By Theorem 23.27 the function h has a fixed point x^*; that is, $x^* = h(x^*)$. The range of h is the set X, and therefore $x^* \in X$. Since $g(x) = x$ for each $x \in X$, we have $x^* = h(x^*) = f(g(x^*)) = f(x^*)$, i.e., x^* is also a fixed point of f. □

23.1.3 Kakutani's Fixed Point Theorem

Brouwer's Fixed Point Theorem is generalized by Kakutani's Fixed Point Theorem, proved by Kakutani [1941], which we present in this section. This theorem does not deal with a function f, but rather with a correspondence, i.e., a set-valued function.

6 Let $X \subseteq \mathbb{R}^n$. A function $f : X \to \mathbb{R}$ is a *Lipschitz function* if there exists a nonnegative real number K satisfying $|f(x) - f(y)| \le K d(x, y)$ for every $x, y \in X$.

Definition 23.31 *Let $X \subset \mathbb{R}^n$. A* correspondence *(or a* set-valued function*) from X to X is a function F associating every $x \in X$ with a subset $F(x)$ of X. The* graph *of a correspondence F is the set $\{(x, y) \in X \times X : y \in F(x)\}$, which is a subset of $X \times X$. A correspondence F whose graph is closed is called an* upper semi-continuous correspondence.

Equivalently, a correspondence F is a function $F : X \to 2^X$, where 2^X is the power set of X, i.e., the set of all subsets of X. Note that if the graph of F is a closed set in $X \times X$, then in particular for each $x \in X$ the set $F(x)$ is a closed subset of X.

Theorem 23.32 (Kakutani [1941]) *Let $X \subset \mathbb{R}^n$ be a compact and convex set, and let F be an upper semi-continuous correspondence from X to X satisfying that for every $x \in X$ the set $F(x)$ is nonempty and convex. Then there exists a point $x^* \in X$ satisfying $x^* \in F(x^*)$.*

A point x^* satisfying $x^* \in F(x^*)$ is called a *fixed point* of the correspondence F.

Remark 23.33 *The Brouwer Fixed Point Theorem (Theorem 23.27) is a special case of Kakutani's Fixed Point Theorem: if $f : X \to X$ is a continuous function, then the correspondence F from X to X defined by $F(x) = \{f(x)\}$ for every $x \in X$ is an upper semi-continuous correspondence with nonempty and convex values. By Kakutani's Fixed Point Theorem it follows that this correspondence has a fixed point x^*. Every such fixed point is a fixed point of f (Exercise 23.28).* ♦

For every subset A of $\mathbb{R}^n$, and every $\varepsilon > 0$, denote the ε*-neighborhood* of A by $B(A, \varepsilon)$,

$$B(A, \varepsilon) := \{y \in \mathbb{R}^n : d(y, A) \leq \varepsilon\}. \tag{23.65}$$

If A is a convex set, then $B(A, \varepsilon)$ is also a convex set (Exercise 23.26). The following lemma states that for an upper semi-continuous correspondence F, if x is "close" to x^0 then $F(x)$ is also "close" to $F(X^0)$.

Lemma 23.34 *Let $F : X \to X$ be an upper semi-continuous correspondence. For each $x^0 \in X$ and each $\varepsilon > 0$ there exists $\delta > 0$ such that $F(x) \subseteq B(F(x^0), \varepsilon)$ for every $x \in X$ satisfying $d(x, x^0) \leq \delta$.*

Proof: Suppose by contradiction that the claim does not hold. Then there exists $x^0 \in X$, and $\varepsilon > 0$ such that for every $\delta > 0$ there is $x^\delta \in X$ satisfying $d(x^\delta, x^0) \leq \delta$, and there is $y^\delta \in F(x^\delta)$ such that $d(y^\delta, F(x^0)) > \varepsilon$. Since X is a compact set, by taking subsequences of (x^δ) and (y^δ), there exist sequences $(x^k)_{k\in\mathbb{N}}$ and $(y^k)_{k\in\mathbb{N}}$ satisfying (a) $\lim_{k\to\infty} x^k = x^0$; (b) $y^k \in F(x^k)$ for every $k \in \mathbb{N}$, and the limit $\widehat{y} := \lim_{k\to\infty} y^k$ exists; and (c) $d(y^k, F(x^0)) > \varepsilon$ for every $k \in \mathbb{N}$. By (a) and (b), and using the fact that the graph of F is a closed set, we deduce that $\widehat{y} \in F(x^0)$. But by (c), it follows that $d(\widehat{y}, F(x^0)) \geq \varepsilon$. These two conclusions contradict each other, with the contradiction showing that the original supposition does not hold; hence the statement of the theorem holds. □

Proof of Kakutani's Fixed Point Theorem (Theorem 23.32): The idea behind the proof is to use the correspondence F to construct a sequence of continuous functions $(f^m)_{m\in\mathbb{N}}$ from X to X. By Brouwer's Fixed Point Theorem, for each $m \in \mathbb{N}$ the function f^m so defined has a fixed point x^m. The functions $(f^m)_{m\in\mathbb{N}}$ will be defined in such a

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

way that each accumulation point x^* of the sequence $(x^m)_{m\in\mathbb{N}}$ is a fixed point of the correspondence F.

Step 1: Defining sequences of continuous functions.
We first define, for every $m \in \mathbb{N}$, a continuous function $f^m : X \to X$. Let $m \in \mathbb{N}$. Since X is a compact set, it can be covered by a finite number of open balls K^m, each of which is of radius $\frac{1}{m}$. Denote the centers of these balls by $(x_k^m)_{k=1}^{K^m}$. Then for every $x \in X$ there exists $k \in \{1, 2, \ldots, K^m\}$ such that $d(x, x_k^m) < \frac{1}{m}$. For every $k \in \{1, 2, \ldots, K^m\}$ choose $y_k^m \in F(x_k^m)$. Denote the set of points whose distance from x_k^m is at least $\frac{1}{m}$ by C_k^m,

$$C_k^m := \left\{x \in X : d(x, x_k^m) \geq \frac{1}{m}\right\}. \tag{23.66}$$

It follows that $d(x, x_k^m) < \frac{1}{m}$ if and only if $x \notin C_k^m$, and since C_k^m is closed, this can happen if and only if $d(x, C_k^m) > 0$. Define, for every $x \in X$ and for every $k \in \{1, 2, \ldots, K^m\}$,

$$\lambda_k^m(x) := \frac{d\left(x, C_k^m\right)}{\sum_{l=1}^{K^m} d\left(x, C_l^m\right)}. \tag{23.67}$$

The denominator in the definition of $\lambda_k^m(x)$ is a sum of nonnegative numbers, and hence is nonnegative. Since for every x there exists $k \in \{1, 2, \ldots, K^m\}$ such that $d(x, x_k^m) < \frac{1}{m}$, the denominator is positive. Since the distance function $x \mapsto d(x, C_k^m)$ is continuous, and the sum of a finite number of continuous functions is a continuous function, both the numerator and the denominator in the definition of λ_k^m are continuous functions. Since the ratio of two continuous functions where the denominator is positive is a continuous function, one deduces that λ_k^m is a continuous function. Note that

$$\sum_{k=1}^{K^m} \lambda_k^m(x) = 1, \quad \forall x \in X. \tag{23.68}$$

Define a function $f^m : X \to X$ as

$$f^m(x) := \sum_{k=1}^{K^m} \lambda_k^m(x) y_k^m. \tag{23.69}$$

In other words, $f^m(x)$ is a convex combination of the points $(y_k^m)_{k=1}^{K^m}$, with weights $(\lambda_k^m(x))_{k=1}^{K^m}$. Since the set X is convex, and since the points $(y_k^m)_{k=1}^{K^m}$ are in X, the range of f^m is contained in X.

Step 2: Using Brouwer's Fixed Point Theorem.
For every $m \in \mathbb{N}$, f^m is a continuous function, because it is a sum of a finite number of continuous functions. By Brouwer's Fixed Point Theorem (Theorem 23.30), this function has a fixed point: there exists $x^{*,m} \in X$ satisfying $x^{*,m} = f(x^{*,m})$.

The sequence of fixed points $(x^{*,m})_{m\in\mathbb{N}}$ is contained in the compact set X, and it therefore contains a convergent subsequence. Denote the subsequence by $(x^{*,m_l})_{l\in\mathbb{N}}$, and denote its limit by x^*.

Step 3: x^* is a fixed point of F.
The idea behind this is as follows. Since x^{*,m_l} is a fixed point of f^{m_l}, it follows that

$$x^{*,m_l} = f^{m_l}(x^{*,m_l}) = \sum_{k=1}^{K^{m_l}} \lambda_k^{m_l}(x^{*,m_l}) y_k^{m_l}. \tag{23.70}$$

The coefficient $\lambda_k^{m_l}(x^{*,m_l})$ is greater than 0 only if $x_k^{m_l}$ is close to x^{*,m_l}, which is close to x^*. Lemma 23.34 then implies that for every k such that $\lambda_k^{m_l}(x^{*,m_l}) > 0$, the point $y_k^{m_l}$ is close to $F(x^*)$. Since $F(x^*)$ is convex, we also deduce that x^{*,m_l}, as a convex combination of points close to $F(x^*)$, is also close to $F(x^*)$. By letting l go to infinity we conclude that $x^* \in F(x^*)$, since the graph of F is a closed set.

The formal proof is as follows. Let $\varepsilon > 0$, and let $\delta > 0$ be the number obtained by applying Lemma 23.34 for $x^0 = x^*$. Let L be sufficiently large such that for all $l \geq L$, (a) $\frac{1}{m_l} < \frac{\delta}{2}$, and (b) $d(x^*, x^{*,m_l}) < \frac{\delta}{2}$. Let $l \geq L$. For every k such that $\lambda_k^{m_l}(x^{*,m_l}) > 0$, one has $d(x_k^{m_l}, x^{*,m_l}) < \frac{1}{m_l} < \frac{\delta}{2}$. Therefore, by the triangle inequality, for each such k,

$$d\left(x^*, x_k^{m_l}\right) \leq d(x^*, x^{*,m_l}) + d\left(x^{*,m_l}, x_k^{m_l}\right) < \frac{\delta}{2} + \frac{\delta}{2} = \delta. \tag{23.71}$$

By Lemma 23.34, $y_k^{m_l} \in F(x_k^{m_l}) \in B(F(x^*), \varepsilon)$ for every such k; i.e., for every $k = 1, 2, \ldots, K^{m_l}$, either $\lambda_k^{m_l}(x^{*,m_l}) = 0$, or $y_k^{m_l} \in B(F(x^*), \varepsilon)$. By Equation (23.70), we then deduce that x^{*,m_l} is a convex combination of points located in the convex set $B(F(x^*), \varepsilon)$, and therefore $x^{*,m_l} \in B(F(x^*), \varepsilon)$. This holds for all $\varepsilon > 0$, and because $F(x^*)$ is a closed set, x^*, as the limit of the sequence $(x^{*,m_l})_{l\in\mathbb{N}}$, is also in $F(x^*)$. □

Remark 23.35 *The properties of the set X that we used in the proof are: (a) convexity, (b) compactness, and (c) every continuous function $f : X \to X$ has a fixed point. Since Brouwer's Fixed Point Theorem holds also for compact and convex sets in infinite-dimensional spaces (see Schauder [1930] or Dunford and Schwartz [1988, Section V.10]), Kakutani's Fixed Point Theorem also holds in such spaces. Proofs of Kakutani's Fixed Point Theorem in infinite-dimensional spaces are given in Bohnenblust and Karlin [1950] and Glicksberg [1952].* ♦

23.1.4 The KKM Theorem

The next theorem we present here, known as the KKM Theorem, after the three researchers who first proved it, Knaster, Kuratowski, and Mazurkiewicz, is a central theorem in topology. It is equivalent to Brouwer's Fixed Point Theorem. A guided proof of the KKM Theorem, using Brouwer's Fixed Point Theorem, is given in Exercise 23.31, and a guided proof of Brouwer's Fixed Point Theorem, using the KKM Theorem, is given in Exercise 23.32. We present here a direct proof of the KKM Theorem, using Sperner's Lemma.

Theorem 23.36 (KKM) *Let $X^1, X^2, \ldots, X^n$ be compact subsets of $X(n)$ satisfying*

$$X^i \supseteq \{x \in X(n) \colon x_i = 0\}, \quad i = 1, \ldots, n, \tag{23.72}$$

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

whose union is $X(n)$,

$$\bigcup_{i=1}^{n} X^i = X(n). \tag{23.73}$$

Then their intersection is nonempty:

$$\bigcap_{i=1}^{n} X^i \neq \emptyset. \tag{23.74}$$

Proof: We first prove the theorem in the special case where the sets $(X^i)_{i=1}^n$ are relatively open sets[7] in $X(n)$. Suppose by contradiction that $\bigcap_{i=1}^n X^i = \emptyset$. In particular, for every $x \in X(n)$ there exists an index i satisfying $x \notin X^i$.

For every $k \in \mathbb{N}$, let $\mathcal{T}_k$ be a simplicial partition of $X(n)$ with diameter less than $\frac{1}{k}$. Define a coloring c of $\mathcal{T}_k$ as follows: for every vertex $y \in Y(\mathcal{T}_k)$, the color of y is one of the indices i such that $y \notin X^i$. Note that if $y_i = 0$, then $y \in X^i$ by Equation (23.72), and in particular the color of y is not i. It follows that if the color of y is i, then necessarily $y_i > 0$, i.e., the color of y is one of the indices in the support of y. In particular, the coloring c is proper. By Sperner's Lemma (Theorem 23.19), there exists a perfectly colored $(n-1)$-dimensional simplex $T_k \in \mathcal{T}_k$. It follows that for every $i = 1, 2, \ldots, n$ there exists a vertex $x^{k,i}$ of T_k whose color is i, and it is therefore not in X^i.

Let $i \in \{1, 2, \ldots, n\}$, and consider the sequence $(x^{k,i})_{k\in\mathbb{N}}$. Since the set $X(n)$ is compact, the sequence has a subsequence converging to $x^{*,i}$. We next show that the following two claims regarding the limits $x^{*,1}, x^{*,2}, \ldots, x^{*,n}$ hold:

1. $x^{*,i} = x^{*,j}$ for all $i, j \in \{1, 2, \ldots, n\}$.
2. $x^{*,i} \notin X^i$ for all $i \in \{1, 2, \ldots, n\}$.

Claim (1) implies that there exists $x^{**} \in X(n)$ satisfying $x^{*,i} = x^{**}$ for all $i \in \{1, 2, \ldots, n\}$, and Claim (2) implies that $x^{**} \notin \cup_{i=1}^n X^i$. This will contradict the assumption that $\cup_{i=1}^n X^i = X(n)$ and lead to the conclusion that $\cap_{i=1}^n X^i \neq \emptyset$, thus completing the proof in the case there the sets $(X^i)_{i=1}^n$ are relatively open.

We start by proving Claim (1). Since $x^{k,i}$ is in T_k for every $i = 1, 2, \ldots, n$, and since the diameter of $\mathcal{T}_k$ is smaller than $\frac{1}{k}$, the distance between $x^{k,i}$ and $x^{k,j}$ is less than $\frac{1}{k}$ for every i, j. Taking the limit $k \to \infty$ yields $x^{*,i} = x^{*,j}$ for every i, j.

We next prove Claim (2). Since $x^{k,i} \notin X^i$ for every $k \in \mathbb{N}$, it follows that $x^{k,i} \in X(n) \setminus X^i$. Since the set X^i is relatively open in $X(n)$, its complement $X(n) \setminus X^i$ is relatively closed in $X(n)$, and therefore the limit $x^{*,i}$ is contained in it, i.e., $x^{*,i} \notin X^i$, which is what we needed to show.

Finally, we show that the statement of the theorem holds when the sets $(X^i)_{i=1}^n$ are closed sets. For every $\delta > 0$ let $X_0^{i,\delta}$ be the open δ-neighborhood of X^i,

$$X_0^{i,\delta} := \{y \in X(n) \colon d(X^i, y) < \delta\}. \tag{23.75}$$

7 A set $A \subseteq \mathbb{R}^n$ is called *relatively open* in a set $C \subseteq \mathbb{R}^n$ if it is the intersection of an open set U in $\mathbb{R}^n$ with C. A set $A \subseteq \mathbb{R}^n$ is called *relatively closed* in a set $C \subseteq \mathbb{R}^n$ if its complement in C, the set $C \setminus A$, is relatively open in C. If A is a relatively closed set in C, then for every sequence of points in A converging to y in C one has $y \in A$.

For every $i \in \{1, 2, \ldots, n\}$, the set $X^{i,\delta}$ is relatively open in $X(n)$, and contains X^i. In particular, $\cup_{i=1}^n X^{i,\delta} \supseteq \cup_{i=1}^n X^i = X(n)$. By the first part of the proof, the intersection $\cap_{i=1}^n X^{i,\delta}$ is nonempty: there exists a point x^δ satisfying $x^\delta \in X^{i,\delta}$ for every $i = \{1, 2, \ldots, n\}$. In particular, $d(X^i, x^\delta) < \delta$. Since $X(n)$ is a compact set, there is a subsequence of the sequence $(x^\delta)_{\delta>0}$ converging to a limit denoted $\widehat{x}$. By passing to the limit, one has $d(X^i, \widehat{x}) = 0$, i.e., $\widehat{x} \in X^i$, for every $i \in N$. Therefore, $\cap_{i=1}^n X^i \neq \emptyset$, which is what we wanted to prove. □

23.2 The Separating Hyperplane Theorem

Recall that the inner product of two vectors in $\mathbb{R}^n$ is defined by

$$\langle x, y\rangle := \sum_{i=1}^n x_i y_i. \tag{23.76}$$

The relationship between the inner product of vectors and the distance between them is given by

$$d(x, y) = \sqrt{\langle x - y, x - y\rangle}. \tag{23.77}$$

Definition 23.37 *A* hyperplane $H(\alpha, \beta)$ *in* $\mathbb{R}^n$ *is defined by*

$$H(\alpha, \beta) := \{x \in \mathbb{R}^n : \langle \alpha, x\rangle = \beta\}, \tag{23.78}$$

where $\alpha \in \mathbb{R}^n$ *and* $\beta \in \mathbb{R}$.

Denote

$$H^+(\alpha, \beta) = \{x \in \mathbb{R}^n : \langle \alpha, x\rangle \geq \beta\} \tag{23.79}$$

and

$$H^-(\alpha, \beta) = \{x \in \mathbb{R}^n : \langle \alpha, x\rangle \leq \beta\}. \tag{23.80}$$

$H^+(\alpha, \beta)$ and $H^-(\alpha, \beta)$ are called the half-spaces defined by the hyperplane $H(\alpha, \beta)$ (see Figure 23.12).

The definition implies (see Exercise 23.33) that

$$H^+(\alpha, \beta) = H^-(-\alpha, -\beta) \text{ and } H^+(\alpha, \beta) \cap H^-(\alpha, \beta) = H(\alpha, \beta). \tag{23.81}$$

A hyperplane separates a set from a point if the set is contained in one of the half-spaces defined by the hyperplane, and the point is contained in the other half-space.

Definition 23.38 *Let* $S \subseteq \mathbb{R}^n$ *be a set, and* $x \in \mathbb{R}^n$ *be a vector. The hyperplane* $H(\alpha, \beta)$ separates *the vector* x *from the set* S *if:*

(i) $x \in H^+(\alpha, \beta)$ *and* $S \subseteq H^-(\alpha, \beta)$*, or*
(ii) $x \in H^-(\alpha, \beta)$ *and* $S \subseteq H^+(\alpha, \beta)$.

If the set S and the point x do not touch the separating hyperplane $H(\alpha, \beta)$, i.e., if:

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

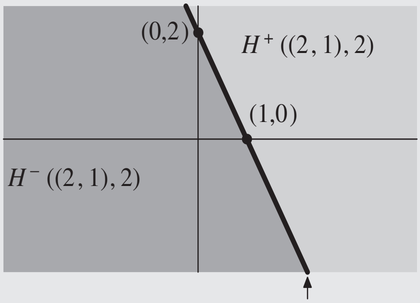


Figure 23.12 The hyperplane $H((2, 1), 2)$ in $\mathbb{R}^2$, which is the line $2x_1 + x_2 = 2$

- $x \in H^+(\alpha, \beta) \setminus H(\alpha, \beta)$ and $S \subseteq H^-(\alpha, \beta) \setminus H(\alpha, \beta)$, or
- $x \in H^-(\alpha, \beta) \setminus H(\alpha, \beta)$ and $S \subseteq H^+(\alpha, \beta) \setminus H(\alpha, \beta)$,

then we say that the separation of x and S is a *strict separation*.

Recall that if S is a closed set, then for every point $x \notin S$ there is at least one point $y \in S$ that is closest to x from among the points in S. If S is convex, the closest point is unique (Theorem 23.28 on page 937).

The following geometric claim, called the *Separating Hyperplane Theorem*, states that for every closed and convex set, and every point that is not in that set, there exists a hyperplane separating the point from the set. The theorem also shows how to construct such a separating hyperplane.

Theorem 23.39 (The Separating Hyperplane Theorem) *Let S be a closed and convex set, and let $x \notin S$. Let y be the closest point in S to x. Then the hyperplane $H(x - y, \langle x - y, y\rangle)$ separates x from S,*

$$S \subseteq H^-(x - y, \langle x - y, y\rangle), \tag{23.82}$$

$$x \in H^+(x - y, \langle x - y, y\rangle). \tag{23.83}$$

In addition, if the condition of the theorem is met, then there exists a hyperplane that strictly separates x from S (Exercise 23.43) (as in Figure 23.13).

Proof: Step 1: $x \in H^+(x - y, \langle x - y, y\rangle)$.
This statement is equivalent to the condition

$$\langle x - y, x\rangle \geq \langle x - y, y\rangle. \tag{23.84}$$

Since the inner product is bilinear, this condition is equivalent to the condition

$$\langle x - y, x - y\rangle \geq 0. \tag{23.85}$$

The left-hand side of this equation is equal to $\sum_{i=1}^n (x_i - y_i)^2$, a nonnegative number, and therefore Equation (23.85) indeed holds. Note that since $x \notin S$, it follows that $x \neq y$ and therefore $\langle x - y, x - y\rangle > 0$, that is, $x \in H^+(x - y, \langle x - y, y\rangle) \setminus H(x - y, \langle x - y, y\rangle)$.

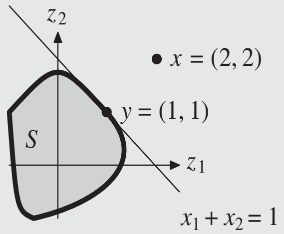


Figure 23.13 The line $x_1 + x_2 = 2$ is a hyperplane separating x from S

Step 2: $S \subseteq H^-(x - y, \langle x - y, y\rangle)$.
We will show that $z \in H^-(x - y, \langle x - y, y\rangle)$ for every $z \in S$. This statement is equivalent to the condition

$$\langle x - y, z\rangle \leq \langle x - y, y\rangle, \quad \forall z \in S. \tag{23.86}$$

Since S is convex, $(1 - \lambda)y + \lambda z \in S$ for every $\lambda \in [0, 1]$. Since y is the closest point in S to x, it follows that

$$d(x, y) \leq d(x, (1 - \lambda)y + \lambda z), \quad \forall \lambda \in [0, 1]. \tag{23.87}$$

Since $d(x, y) = \sqrt{\langle x - y, x - y\rangle}$ for every pair of vectors $x, y \in \mathbb{R}^n$, it follows from Equation (23.87) that

$$\langle x - y, x - y\rangle \leq \langle x - (1 - \lambda)y - \lambda z, x - (1 - \lambda)y - \lambda z\rangle \tag{23.88}$$
$$= \langle (x - y) - \lambda(z - y), (x - y) - \lambda(z - y)\rangle \tag{23.89}$$
$$= \langle x - y, x - y\rangle - 2\lambda\langle x - y, z - y\rangle + \lambda^2\langle z - y, z - y\rangle. \tag{23.90}$$

Subtracting the term $\langle x - y, x - y\rangle$ yields, for all $\lambda \in [0, 1]$, the following

$$0 \leq -2\lambda\langle x - y, z - y\rangle + \lambda^2\langle z - y, z - y\rangle. \tag{23.91}$$

This inequality holds for all $\lambda \in [0, 1]$, and in particular for $\lambda \in (0, 1]$, in which case we can divide by λ to obtain

$$0 \leq -2\langle x - y, z - y\rangle + \lambda\langle z - y, z - y\rangle. \tag{23.92}$$

Letting λ go to 0 yields the conclusion that $\langle x - y, z - y\rangle \leq 0$, which further implies that

$$\langle x - y, z\rangle \leq \langle x - y, y\rangle, \quad \forall z \in S, \tag{23.93}$$

which is what we wanted to show. □

23.3 Linear programming

All the vectors appearing in this section are to be interpreted as row vectors. $y^\top$ is then the column vector corresponding to the row vector y. For $c, y \in \mathbb{R}^m$, the inner product $\langle c, y\rangle$

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

can be written as

$$\langle c, y\rangle = \sum_{i=1}^{m} c_i y_i = cy^\top. \tag{23.94}$$

The notation $y \geq c$ indicates inequality in every coordinate,

$$y_i \geq c_i, \quad i = 1, 2, \dots, m, \tag{23.95}$$

and $y = c$ indicates equality in every coordinate, $y_i = c_i$ for all $i = 1, 2, \dots, m$. Recall that for all $m \in \mathbb{N}$ we denote the zero vector in $\mathbb{R}^m$ by $\vec{0}$.

Definition 23.40 *Let $c, y \in \mathbb{R}^m$, $b \in \mathbb{R}^n$, and let A be an $n \times m$ matrix. The* (standard) linear program *in the unknowns $y = (y_i)_{i=1}^m$ defined by A, b, c is the following linear maximization program under linear constraints.*

$$\begin{array}{ll} \textit{Compute:} & Z_P := \max cy^\top, \\ \textit{subject to:} & Ay^\top \leq b^\top, \\ & y \geq \vec{0}. \end{array}$$

The linear function $cy^\top$ is called the *objective function* of the program. The conditions $Ay^\top \leq b^\top$ and $y \geq \vec{0}$ are called *constraints*. Every vector y satisfying the constraints is called a *feasible vector*. The set of feasible vectors is called the *feasible region* of the linear program,

$$R_P := \{y \in \mathbb{R}^n : Ay^\top \leq b^\top, y \geq \vec{0}\}. \tag{23.96}$$

A feasible vector maximizing $cy^\top$ among all feasible vectors is called an *optimal solution*. This maximum, Z_P, is called the *value* of the linear program; it is the maximum of the objective function over the feasible region. When the feasible region is empty, define $Z_P := -\infty$. If the objective function $cy^\top$ is not bounded over R_P (which happens only if R_P is unbounded) define $Z_P := +\infty$. Since the objective function is linear, and the constraints are weak inequalities, if $Z_P < +\infty$ then the maximum is attained at one of the extreme points of R_P, whether R_P is bounded or not.

Definition 23.41 *Given the following linear program with unknowns $y = (y_i)_{i=1}^m$*

$$\begin{array}{ll} \textit{Compute:} & Z_P := \max cy^\top, \\ \textit{subject to:} & Ay^\top \leq b^\top, \\ & y \geq \vec{0}; \end{array} \tag{23.97}$$

its dual program *is the following program with unknowns $x = (x_j)_{j=1}^n$.*

$$\begin{array}{ll} \textit{Compute:} & Z_D := \min xb^\top, \\ \textit{subject to:} & xA \geq c, \\ & x \geq \vec{0}. \end{array} \tag{23.98}$$

The original program is called the primal program. *The number Z_D is called the* value *of the dual program.*

The conditions $xA \geq c$ and $x \geq \vec{0}$ are called the *constraints* of the dual program. Note that the primal program has m unknowns $y = (y_i)_{i=1}^m$, one for each column of A, i.e., one unknown for each constraint $xA \geq c$ of the dual program. The dual program has n unknowns $(x_j)_{j=1}^n$, one for each row of A, i.e., an unknown for each of the constraints $Ay^\top \leq b^\top$ of the primal program. If the feasible region of the dual program is empty, define $Z_D := +\infty$, and if the objective function $xb^\top$ is unbounded from below in R_D (this can only happen if R_D is unbounded) define $Z_D := -\infty$.

Example 23.42 Consider the primal program given by

$$\begin{array}{ll} \text{Compute:} & Z_P := \max\{6y_1 + 4y_2 + 6y_3\}, \\ \text{subject to:} & 6y_1 + y_2 + y_3 \leq 3, \\ & y_1 + y_2 + 3y_3 \leq 4, \\ & y_1, y_2, y_3 \geq 0. \end{array} \tag{23.99}$$

In this linear program

$$c = (6, 4, 6), \qquad b = (3, 4), \qquad A = \begin{pmatrix} 6 & 1 & 1 \\ 1 & 1 & 3 \end{pmatrix}. \tag{23.100}$$

The dual program is

$$\begin{array}{ll} \text{Compute:} & Z_D := \min\{3x_1 + 4x_2\}, \\ \text{subject to:} & 6x_1 + x_2 \geq 6, \\ & x_1 + x_2 \geq 4, \\ & x_1 + 3x_2 \geq 6, \\ & x_1, x_2 \geq 0. \end{array} \tag{23.101}$$

◄

The constraints in a linear program may be given by equalities (and not necessarily only by inequalities). Every equality can be represented by two inequalities: a vector equality $Ay^\top = b^\top$ can be represented by the inequalities $Ay^\top \leq b^\top$ and $-Ay^\top \leq -b^\top$. It follows that a linear program with equalities can be rewritten as a system involving inequalities only.

Example 23.43 Consider the following linear program:

$$\begin{array}{ll} \text{Compute:} & Z_P := \max cy^\top, \\ \text{subject to:} & Ay^\top = b^\top, \\ & y \geq \vec{0}. \end{array} \tag{23.102}$$

This program can be rewritten as a system involving inequalities only:

$$\begin{array}{ll} \text{Compute:} & Z_P := \max cy^\top, \\ \text{subject to:} & Ay^\top \leq b^\top, \\ & -Ay^\top \leq -b^\top, \\ & y \geq \vec{0}. \end{array} \tag{23.103}$$

In this representation, the primal program has $2n$ constraints. The dual program must therefore have $2n$ unknowns, denoted (w, z), where $w = (w_j)_{j=1}^n$ and $z = (z_j)_{j=1}^n$. The dual program

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

corresponding to this primal program is

$$\begin{aligned} &\text{Compute:} && Z_D := \min(w-z)b^\top, \\ &\text{subject to:} && (w-z)A \geq c, \\ &&& w \geq \vec{0}, \\ &&& z \geq \vec{0}. \end{aligned} \tag{23.104}$$

Set $x = w - z$. Since w and z are nonnegative, x is not constrained, and can take on any value, positive or negative. The dual program is then

$$\begin{aligned} &\text{Compute:} && Z_D := \min x b^\top, \\ &\text{subject to:} && xA \geq c. \end{aligned} \tag{23.105}$$

In other words, for every feasible solution x of Problem (23.105) there exists a feasible solution (w, z) of Problem (23.104) satisfying $x = w - z$ (Exercise 23.48). ◀

As a corollary we deduce the following theorem.

Theorem 23.44 *The dual program to the following primal program,*

$$\begin{aligned} &\textit{Compute:} && Z_P := \max c y^\top, \\ &\textit{subject to:} && Ay^\top = b^\top, \\ &&& y \geq 0, \end{aligned} \tag{23.106}$$

is

$$\begin{aligned} &\textit{Compute:} && Z_D := \min x b^\top, \\ &\textit{subject to:} && xA \geq c. \end{aligned} \tag{23.107}$$

The effect of equalities in the constraints of the primal problem can thus be summarized as follows: each variable in the dual problem that corresponds to a constraint with equality in the primal problem is unconstrained.

Example 23.42 (*Continued*) We solve both the primal and the dual program, starting with the primal program. Since the numbers in the matrix A are nonnegative, and since the numbers in the vector c are positive, the maximum in the definition of Z_P is attained when the following equalities are satisfied:

$$6y_1 + y_2 + y_3 = 3, \tag{23.108}$$

$$y_1 + y_2 + 3y_3 = 4. \tag{23.109}$$

The solution to this system of equations is

$$y_1 = \tfrac{1}{5}(2y_3 - 1), \tag{23.110}$$

$$y_2 = \tfrac{1}{5}(21 - 17y_3). \tag{23.111}$$

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

The condition that $y_1, y_2, y_3 \geq 0$ implies that $\frac{1}{2} \leq y_3 \leq \frac{21}{17}$. The primal program is thus equivalent to

$$Z_P = \max\left\{\tfrac{6}{5}(2y_3 - 1) + \tfrac{4}{5}(21 - 17y_3) + 6y_3 : \tfrac{1}{2} \leq y_3 \leq \tfrac{21}{17}\right\} \tag{23.112}$$

$$= \max\left\{\tfrac{78}{5} - \tfrac{26}{5}y_3 : \tfrac{1}{2} \leq y_3 \leq \tfrac{21}{17}\right\} = \tfrac{78}{5} - \tfrac{13}{5} = 13. \tag{23.113}$$

The computation of the last maximum uses the fact that a linear function in y_3 with a negative slope attains its maximal value at the minimal value of y_3, which in this case is $y_3 = \frac{1}{2}$.

Next, we solve the dual program. The feasible region of the dual program is the shaded area in Figure 23.14. This is the intersection of the half-spaces corresponding to the inequalities defining the constraints of the dual program. As Figure 23.14 illustrates, the minimum of the function $3x_1 + 4x_2$ in the feasible region is attained at the point $(3, 1)$, and therefore $Z_D = 13$.

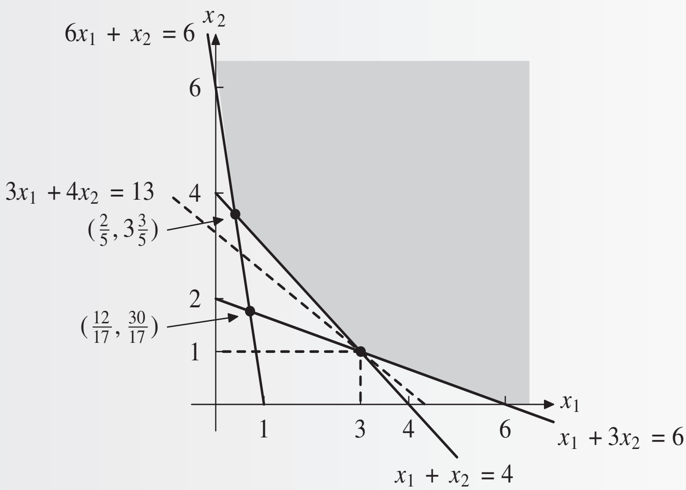


Figure 23.14 The feasible region and the solution of the dual program in Example 23.42 ◀

In Example 23.42 we computed $Z_P = Z_D$. This is not a coincidence, as the next two theorems show.

Theorem 23.45 (The Weak Duality Theorem) *Let Z_P and Z_D be, respectively, the values of the primal and the dual problems given in Definition 23.40. Then $Z_P \leq Z_D$.*

Proof: If R_P is empty, $Z_P = -\infty$, and if R_D is empty, $Z_D = +\infty$. In both cases, $Z_P \leq Z_D$ is satisfied. Suppose therefore that R_P and R_D are nonempty. Since every feasible vector x of the primal program and every feasible vector y of the dual program is nonnegative, one has

$$cy^\top \leq (xA)y^\top = x(Ay^\top) \leq xb^\top. \tag{23.114}$$

By taking the maximum on the left-hand side, and the minimum on the right-hand side, we conclude that $Z_P \leq Z_D$, which is what we wanted to show. □

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

The following theorem, for which we will not provide a proof, states that the inequality $Z_P \leq Z_D$ is actually an equality when Z_P is finite. A proof of the theorem can be found in many books on operations research, such as Vanderbei [2001].

Theorem 23.46 (The Strong Duality Theorem) *If Z_P is finite, then $Z_D = Z_P$.*

23.4 Remarks

The authors thank Nimrod Megiddo, David Schmeidler, and Rakesh Vohra for answering questions that arose during the composition of this chapter. The proof of Sperner's Lemma is from Kuhn [1968]. The proof of Brouwer's Fixed Point Theorem using Sperner's Lemma is a classical proof. The proof presented here is from Kuhn [1960].

23.5 Exercises

23.1 Prove that for every set $\{x^0, x^1, \ldots, x^k\}$ of vectors in $\mathbb{R}^n$

$$\begin{aligned}&\operatorname{conv}(x^0, x^1, \ldots, x^k)\\&\quad = \{x \in \mathbb{R}^n : x \text{ is a convex combination of } x^0, x^1, \ldots, x^k\}.\end{aligned} \tag{23.115}$$

The convex hull $\operatorname{conv}(x^0, x^1, \ldots, x^k)$ is defined in Definition 23.1 (page 917).

23.2 Let $x^0, x^1, \ldots, x^k$ be vectors in $\mathbb{R}^n$, such that none of them is a convex combination of the other vectors. Prove that the extreme points of $\operatorname{conv}(x^0, x^1, \ldots, x^k)$ are the vectors $x^0, x^1, \ldots, x^k$.

23.3 Let x^0, x^1, x^2 be three vectors in $\mathbb{R}^n$. Prove that these vectors are affine independent if and only if none of them is a convex combination of the other two.

23.4 For each of the following sets of vectors in $\mathbb{R}^4$, determine whether or not the vectors in the set are affine independent. Justify your answers.

(a) $x^0 = (1, 1, 0, 0)$, $x^1 = (0, 1, 1, 0)$, $x^2 = (0, 0, 1, 1)$, $x^3 = (1, 0, 0, -1)$, $x^4 = (1, 1, 1, 0)$.
(b) $x^0 = (1, 1, 0, 0)$, $x^1 = (0, 1, 1, 0)$, $x^2 = (0, 0, 1, 1)$, $x^3 = (1, 0, 0, -1)$, $x^4 = (1, 2, 1, 0)$.
(c) $x^0 = (1, 1, 0, 0)$, $x^1 = (0, 1, 1, 0)$, $x^2 = (0, 0, 1, 1)$, $x^3 = (1, 0, 0, -1)$, $x^4 = (1, 0, 0, 0)$.

23.5 Let $x^0, x^1, \ldots, x^k$ be affine-independent vectors in $\mathbb{R}^n$, and let y be a vector that is linearly independent of $\{x^0, x^1, \ldots, x^k\}$. Prove that $x^0, x^1, \ldots, x^k, y$ are affine-independent vectors.

23.6 Let $x^0, x^1, \ldots, x^k$ be affine-independent vectors in $\mathbb{R}^n$ and let $(\beta^l)_{l=0}^k$ be positive numbers summing to 1. Denote $y = \sum_{l=0}^k \beta^l x^l$. Prove that $x^1, \ldots, x^k, y$ are affine-independent vectors.

23.7 Prove that vectors $x^0, x^1, \ldots, x^k$ are affine independent in $\mathbb{R}^n$ if and only if every vector $y \in \operatorname{conv}\{x^0, x^1, \ldots, x^k\}$ can be represented in a unique way as a convex combination of $x^0, x^1, \ldots, x^k$.

23.8 Let $S = \langle\langle x^0, x^1, \ldots, x^k \rangle\rangle$ be a simplex in $\mathbb{R}^n$ and let $y \in \mathbb{R}^n$. Prove that $y \in S$ if and only if all the solutions of the following system of equations in the variables $(\alpha^l)_{l=0}^k$ satisfy $\alpha^l \geq 0$ for all $l \in \{1, 2, \ldots, k\}$:

$$\sum_{l=0}^{k} \alpha^l x^l = y, \tag{23.116}$$

$$\sum_{l=0}^{k} \alpha^l = 1. \tag{23.117}$$

23.9 Let $x^0, x^1, \ldots, x^n$ be affine-independent vectors in $\mathbb{R}^n$. Prove that for every vector $y \in \mathbb{R}^n$ there exists a unique solution to the following system of equations in unknowns $(\alpha^l)_{l=0}^n$:

$$\sum_{l=0}^{n} \alpha^l x^l = y, \tag{23.118}$$

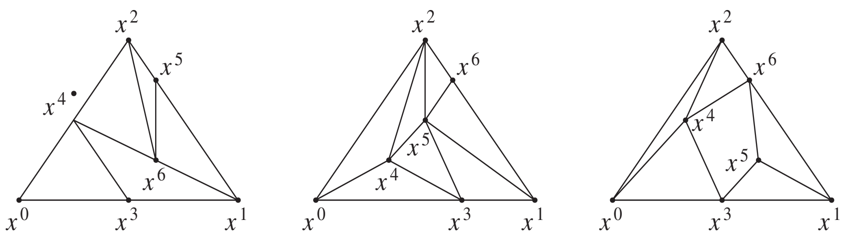
$$\sum_{l=0}^{n} \alpha^l = 1. \tag{23.119}$$

23.10 Prove that the boundary of a simplex S is the set of all the points y in S whose barycentric coordinate representation has at least one zero coordinate.

23.11 Prove that if H^1 and H^2 are two affine spaces of the same dimension k in $\mathbb{R}^n$ (that is, each one of them is spanned by a k-dimensional simplex; see page 922), and if $H^1 \subseteq H^2$, then $H^1 = H^2$.

23.12 For each of the following vectors $y^1 = (1, \frac{1}{2})$, $y^2 = (3, 0)$, $y^3 = (2, 2)$, and $y^4 = (\frac{3}{2}, \frac{1}{4})$, compute its barycentric coordinates relative to the three affine-independent vectors $x^0 = (0, 2)$, $x^1 = (1, 0)$, and $x^2 = (2, 0)$.

23.13 For each of the following partitions of a two-dimensional simplex, determine whether or not it is a simplicial partition. Justify your answer. In each case, the partition elements are the two-dimensional polytopes in it, their faces, and their vertices.



Partition A Partition B Partition C

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

23.14 Prove Theorem 23.14 (page 922): let S be a k-dimensional simplex in $\mathbb{R}^n$, and let $\mathcal{T}$ be a simplicial partition of it. Then S equals the union of all the k-dimensional simplices in $\mathcal{T}$.

23.15 Let S be a simplex, and let $\mathcal{T}$ be the collection of all the faces of S. Prove that $\mathcal{T}$ is a simplicial partition of S.

23.16 Let $T_1, T_2, \ldots, T_L$ be simplices whose union is a simplex S. Is the collection containing all the faces of $T_1, T_2, \ldots, T_L$ a simplicial partition of S? Either prove this claim, or find a counterexample.

23.17 Let $\mathcal{T}$ be a simplicial partition of a k-dimensional simplex $S = \langle\langle x^0, x^1, \ldots, x^k \rangle\rangle$. Define $\widehat{S} := \langle\langle x^0, x^1, \ldots, x^{k-1} \rangle\rangle$ and $\widehat{\mathcal{T}} := \{T \in \mathcal{T} : T \subseteq \widehat{S}\}$. Prove that $\widehat{\mathcal{T}}$ is a simplicial partition of $\widehat{S}$.

23.18 How many proper colorings are there of the simplicial partition in Example 23.17 on page 926? Justify your answer. (For the definition of a proper coloring, see Definition 23.16 on page 925.)

23.19 Find a simplex, a simplicial partition of this simplex, and a nonproper coloring of the simplicial partition, for which the number of perfectly colored n-dimensional simplices is odd. In other words, show that in Sperner's Lemma (Theorem 23.19 on page 926) the properness of the coloring is a sufficient but not necessary condition for this number to be odd.

23.20 Let $\mathcal{T}$ be a simplicial partition of a k-dimensional simplex $S = \langle\langle x^0, x^1, \ldots, x^k \rangle\rangle$, and let $c : Y(\mathcal{T}) \to \{0, 1, \ldots, k\}$ be a proper coloring (see Definition 23.16 on page 925). Denote $\widehat{S} := \langle\langle x^0, x^1, \ldots, x^{k-1} \rangle\rangle$ and $\widehat{\mathcal{T}} := \{T \in \mathcal{T} : T \subseteq \widehat{S}\}$. Prove that c restricted to $Y(\widehat{\mathcal{T}})$ is a proper coloring of $\widehat{\mathcal{T}}$.

23.21 Let $\mathcal{T}$ be the simplicial partition constructed in the proof of Theorem 23.22 (page 930). What is $Y(\mathcal{T})$?

23.22 Prove that the diameter $\rho(S)$ of a simplex S (see Definition 23.20 on page 929) equals the greatest (Euclidean) distance between two vectors in the simplex.

23.23 Complete the proof of Theorem 23.23 (page 930): show that the intersection of any two simplices in the simplicial partition $\mathcal{T}$ that is defined in the proof of the theorem is either empty or is contained in $\mathcal{T}$.

23.24 Complete the proof of Theorem 23.25 (page 931): show that the intersection of any two simplices in the simplicial partition $\widehat{\mathcal{T}}$ that is defined in the proof of the theorem is either empty or is contained in $\widehat{\mathcal{T}}$.

23.25 Prove that the diameter of the partition $\widehat{\mathcal{T}}$ constructed in the proof of Theorem 23.25 (page 931) is

$$\rho(\widehat{\mathcal{T}}) = \max\{\rho(\mathcal{T}), \|x^0 - y\|, \|x^1 - y\|, \ldots, \|x^k - y\|\}. \qquad (23.120)$$

23.26 Prove that if $A \subseteq \mathbb{R}^n$ is a convex set, then the set $B(A, \varepsilon)$ (the ε-neighborhood of A; see Equation (23.65) on page 939) is also a convex set.

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

23.27 Prove Theorem 23.29 (page 938): let $X \subset \mathbb{R}^n$ be a closed and convex set. Define a function $g : \mathbb{R}^n \to X$ as follows: $g(x)$ is the closest point to x in X. Then $d(g(x), g(\widehat{x})) \leq d(x, \widehat{x})$ for all $x, \widehat{x} \in \mathbb{R}^n$.

23.28 (a) Let $X \subseteq \mathbb{R}^n$. Prove that if $f : X \to X$ is a continuous function, then the correspondence $F : X \to X$ defined by $F(x) = \{f(x)\}$ for every $x \in X$ is upper semi-continuous.

(b) Use the result of the previous item to prove Brouwer's Fixed Point Theorem (Theorem 23.30, page 938) using Kakutani's Fixed Point Theorem (Theorem 23.32, page 939).

23.29 In this exercise, we prove a generalization of Brouwer's Fixed Point Theorem to compact sets that are not necessarily convex, but are homeomorphic to a convex set.

Two compact sets X and Y in $\mathbb{R}^n$ are called *homeomorphic* if there exists a continuous bijection $g : X \to Y$ satisfying the property that g^{-1} is also continuous.[8] Prove that if $X \subseteq \mathbb{R}^n$ is a compact set that is homeomorphic to a convex and compact set Y and if $f : X \to X$ is a continuous function, then f has a fixed point.

23.30 In this exercise, we prove a generalization of Nash's Theorem (Theorem 5.10 on page 151) using Kakutani's Fixed Point Theorem (Theorem 23.32 on page 939).

Let N be a nonempty, finite set of players. For each player $i \in N$, let X_i be a convex and compact subset of $\mathbb{R}^{d_i}$, where d_i is a natural number. Denote by $X = \times_{i \in N} X_i$ the Cartesian product of the sets $(X_i)_{i \in N}$. For every player $i \in N$, let $u_i : X \to \mathbb{R}$ be a function that is continuous and quasi-concave in x_i; that is, for every real number c and all $x_{-i} \in X_{-i}$, the set $\{x_i : u_i(x_i, x_{-i}) \geq c\}$ is convex.

Define a correspondence br from X to X as follows. For every $x \in X$ and for each $i \in N$,

$$\mathrm{br}_i(x) := \left\{ y_i \in X_i : u_i(y_i, x_{-i}) = \max_{z_i \in X_i} u_i(z_i, x_{-i}) \right\}, \quad \forall i \in N, \qquad (23.121)$$

and

$$\mathrm{br}(x) := \times_{i \in N} \mathrm{br}_i(x). \qquad (23.122)$$

(a) Prove that br is an upper semi-continuous correspondence with nonempty convex values. Using Kakutani's Fixed Point Theorem deduce that br has a fixed point.

(b) Prove that every fixed point of the correspondence br is a Nash equilibrium in the game $G = (N, (X_i)_{i \in N}, (u_i)_{i \in N})$. How is this a generalization of Nash's Theorem?

23.31 In this exercise, we prove the KKM Theorem (Theorem 23.36 on page 941) using Brouwer's Fixed Point Theorem (Theorem 23.30 on page 938). Suppose by contradiction that the conditions of the KKM Theorem hold yet its conclusion does not hold, i.e., $\bigcap_{i=1}^{n} X^i = \emptyset$.

8 The function g is called a *homeomorphism* between X and Y.

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

(a) For each $\varepsilon > 0$ define

$$Y^{i,\varepsilon} = \{x \in X(n) \colon d(x, X^i) \leq \varepsilon\}. \tag{23.123}$$

Prove that $\bigcup_{i=1}^n Y^{i,\varepsilon} = X(n)$ for all $\varepsilon > 0$ and that there exists $\varepsilon_0 > 0$ such that $\bigcap_{i=1}^n Y^{i,\varepsilon_0} = \emptyset$.

For $\varepsilon < \varepsilon_0$:

(b) Denote $Z^{i,\varepsilon} = X(n) \setminus Y^{i,\varepsilon}$ for every $i = 1, 2, \ldots, n$. Prove that $\sum_{i=1}^n d(x, Z^{i,\varepsilon}) \geq \varepsilon$ for every $x \in X(n)$.
(c) Prove that $\sum_{i=1}^n x_i d(x, Y^{i,\varepsilon}) = \sum_{\{i \colon x \notin Y^{i,\varepsilon}\}} x_i d(x, Y^{i,\varepsilon})$.
(d) Define a function $f^\varepsilon : X(n) \to X(n)$ as follows:

$$f_i^\varepsilon(x) = \begin{cases} x_i - x_i d(x, Y^{i,\varepsilon}) & \text{if } x \notin Y^{i,\varepsilon}, \\ x_i + \left(\sum_{j=1}^n x_j d(x, Y^{j,\varepsilon})\right) \frac{d(x, Z^{i,\varepsilon})}{\sum_{j=1}^n d(x, Z^{j,\varepsilon})} & \text{if } x \in Y^{i,\varepsilon}. \end{cases} \tag{23.124}$$

Prove that the function f^ε is continuous and that its range is contained in $X(n)$. Deduce that f^ε has a fixed x^ε.
(e) Prove that $x^\varepsilon \in \bigcap_{i=1}^n Y^{i,\varepsilon}$.
(f) Let $(\varepsilon_k)_{k \in \mathbb{N}}$ be a sequence converging to 0 such that $x^* := \lim_{k \to \infty} x^{\varepsilon_k}$ exists. Prove that $x^* \in \bigcap_{i=1}^n X^i$, contradicting the assumption that $\bigcap_{i=1}^n X^i = \emptyset$.

23.32 Prove Brouwer's Fixed Point Theorem (Theorem 23.30 on page 938) using the KKM Theorem (Theorem 23.36 on page 941).
Hint: Define $X_i = \{x \in X(n) \colon f_i(x) \geq x_i\}$.

23.33 Prove that for every $\alpha \in \mathbb{R}^n$ and every $\beta \in \mathbb{R}$,

$$H^+(\alpha, \beta) = H^-(-\alpha, -\beta). \tag{23.125}$$

23.34 Prove that $x \in H^+(x - y, \langle x - y, y \rangle)$ for every pair of vectors $x, y \in \mathbb{R}^n$.

23.35 Let $x, y \in \mathbb{R}^m$ be two different vectors. Prove the following claims:

(a) $y \in H(y - x, \langle y - x, y \rangle)$.
(b) The hyperplane $H(y - x, \langle y - x, y \rangle)$ is perpendicular to $y - x$, i.e., $\langle y - x, y - z \rangle = 0$ for all $z \in H(y - x, \langle y - x, y \rangle)$.
(c) y is the point in $H(y - x, \langle y - x, y \rangle)$ that is closest to x, i.e., $\langle z - x, z - x \rangle > \langle y - x, y - x \rangle$ for all $z \in H(y - x, \langle y - x, y \rangle)$, $z \neq y$.

23.36 Let H be a hyperplane, let $x \notin H$ and let $y \in H$ be the point in H that is closest to x. Prove that $H = H(y - x, \langle y - x, y \rangle)$.

23.37 Let $H(\alpha, \beta)$ be a hyperplane and let $x \notin H(\alpha, \beta)$. Define

$$y := x + \frac{\beta - \langle \alpha, x \rangle}{\langle \alpha, \alpha \rangle} \alpha. \tag{23.126}$$

Prove the following claims:

(a) $y \in H(\alpha, \beta)$.
(b) $H(\alpha, \beta) = H(y - x, \langle y - x, y \rangle)$.

https://doi.org/10.1017/CBO9780511794216.024 Published online by Cambridge University Press

23.38 In each of the following items, find a closed and convex set $S \subseteq \mathbb{R}^2$, and a vector $x \notin S$, satisfying the following properties:

(a) There exists a unique hyperplane separating S from x.
(b) There exist at least two hyperplanes separating S from x, and $\alpha = \widehat{\alpha}$ for every pair of such hyperplanes $H(\alpha, \beta)$ and $H(\widehat{\alpha}, \widehat{\beta})$.

23.39 Let $H(\alpha, \beta)$ be a hyperplane separating a closed and convex set S from a vector $x \notin S$, and let $H(\widehat{\alpha}, \widehat{\beta})$ be a hyperplane separating a closed and convex set $\widehat{S}$ from a vector $\widehat{x} \notin \widehat{S}$. Does the hyperplane $H(\alpha + \widehat{\alpha}, \beta + \widehat{\beta})$ separate the set $S + \widehat{S} := \{y + \widehat{y} \colon y \in S, \widehat{y} \in \widehat{S}\}$ and the vector $x + \widehat{x}$? Either prove that this is true, or find a counterexample.

23.40 Let S be a closed set (not necessarily convex) and let $x \notin S$. Must there exist a hyperplane separating x from S? Either prove that this is true, or find a counterexample.

23.41 Each of the following items presents a closed and convex set S and a vector x. For each such pair, determine whether or not there is a hyperplane separating x from S, and if your answer is affirmative, find such a hyperplane. If your answer is negative, justify your answer.[9]

(a) $S = \{z \in \mathbb{R}^2 \colon (z_1)^2 + (z_2)^2 \leq 1\}$, $x = (0, 0)$.
(b) $S = \{z \in \mathbb{R}^2 \colon (z_1)^2 + (z_2)^2 \leq 1\}$, $x = (1, 1)$.
(c) $S = \{z \in \mathbb{R}^2 \colon (z_1)^2 + (z_2)^2 \leq 1\}$, $x = (0, 1)$.
(d) $S = [(0, 0), (1, 1)] \cup [(0, 0), (-1, 1)]$, $x = (0, \frac{1}{2})$.
(e) $S = [(0, 0), (1, 1)] \cup [(0, 0), (-1, 1)]$, $x = (\frac{1}{2}, 1)$.
(f) $S = [(0, 0), (1, 1)] \cup [(0, 0), (-1, 1)]$, $x = (\frac{1}{4}, 2)$.
(g) $S = \{z \in \mathbb{R}^3 \colon (z_1)^2 + (z_2)^2 + (z_3)^2 = 1$, $x = (0, 0, 0)$.
(h) $S = \{z \in \mathbb{R}^3 \colon (z_1)^2 + (z_2)^2 + (z_3)^2 = 1$, $x = (0, 1, 1)$.

23.42 Each of the following items presents a closed and convex set S and a vector x. For each such pair, find the hyperplane separating x from S, as described in Theorem 23.39 (page 944).

(a) $S = \{z \in \mathbb{R}^3 \colon \max\{|z_1|, |z_2|, |z_3|\} \leq 1\}$, $x = (2, 2, 2)$.
(b) $S = \{z \in \mathbb{R}^3 \colon \max\{|z_1|, |z_2|, |z_3|\} \leq 1\}$, $x = (2, 3, 4)$.
(c) $S = \{z \in \mathbb{R}^3 \colon (z_1)^2 + (z_2)^2 + (z_3)^2 = 1\}$, $x = (2, 2, 2)$.
(d) $S = \{z \in \mathbb{R}^2 \colon z_1 \geq 0, z_1 + z_2 \leq 1, z_2 \leq z_1\}$, $x = (1, 2)$.
(e) $S = \{z \in \mathbb{R}^2 \colon z_1 \geq 0, z_1 + z_2 \leq 1, z_2 \leq z_1\}$, $x = (2, 3)$.

23.43 Let $S \subseteq \mathbb{R}^n$ be a closed and convex set, and let $x \notin S$. Find a hyperplane $H(\alpha, \beta)$ strictly separating x from S; i.e., x is in the interior of $H^+(\alpha, \beta)$, and S is in the interior $H^-(\alpha, \beta)$.

23.44 In this exercise, we provide a guided proof of Farkas' Lemma: let $v \in \mathbb{R}^n$ be a vector, and let T be an $n \times m$ matrix. Then the following two claims are equivalent.

9 For every $x, y \in \mathbb{R}^2$ denote the line segment connecting x and y by $[x, y]$, i.e., $[x, y] = \{z \in \mathbb{R}^2 \colon x = \lambda x + (1 - \lambda)y, 0 \leq \lambda \leq 1\}$.

(a) $\langle u, v\rangle = uv^\top \geq 0$ for every vector $u \in \mathbb{R}^n$ satisfying $uT \geq \vec{0}$.
(b) There exists a vector $w \in \mathbb{R}^m$, $w \geq \vec{0}$, satisfying $Tw^\top = v$.

Guidance: To prove that the first claim implies the second claim, define the set $A := \{Tw^\top : w \in \mathbb{R}^m, w \geq \vec{0}\} \subseteq \mathbb{R}^n$, and suppose by contradiction that $v \notin A$. Show that by the Separating Hyperplane Theorem there exists a hyperplane $H(\alpha, \beta)$ separating A from v, and that it is possible to assume without loss of generality that $\beta = 0$. Prove that $\alpha T \geq \vec{0}$, and derive a contradiction.

23.45 Prove the following theorem, which is a generalization of Theorem 23.39 (page 944).

Theorem 23.47 *Let X and Y be two closed and convex subsets of $\mathbb{R}^n$. If X and Y are disjoint sets, then there exists a hyperplane $H(\alpha, \beta)$ strictly separating X from Y, i.e.,*

$$\langle \alpha, x\rangle > \beta > \langle \alpha, y\rangle, \quad \forall x \in X, y \in Y. \tag{23.127}$$

Guidance: Denote by $d(X, Y)$ the distance between X and Y,

$$d(X, Y) := \min_{\{x \in X, y \in Y\}} d(x, y). \tag{23.128}$$

The minimum in the definition of $d(X, Y)$ is attained because X and Y are closed sets. Let $x \in X$ and $y \in Y$ be points that minimize the distance between the points in X and the points in Y,

$$d(X, Y) = d(x, y) > 0. \tag{23.129}$$

Write down the equation for the hyperplane separating the vector x from Y, as constructed in the proof of Theorem 23.39, and the equation for the hyperplane separating y from X, as constructed in the proof of Theorem 23.39. Using these two hyperplanes, construct a hyperplane strictly separating X from Y.

23.46 In this exercise we generalize Exercise 23.45, and show that any two disjoint convex sets can be separated.

(a) Let X be a convex set. For every $\varepsilon > 0$ denote by X_ε the set of all points whose distance from X^c, the complement of X, is at least ε:

$$X_\varepsilon := \{x \in X : d(x, X^c) \geq \varepsilon\}. \tag{23.130}$$

Prove that X_ε is a closed and convex set for every $\varepsilon > 0$.
Let X and Y be two disjoint convex sets.
(b) Prove that the sets X_ε and Y_ε are disjoint for every $\varepsilon > 0$.
(c) Deduce from Exercise 23.45 that for every $\varepsilon > 0$ there exists a hyperplane $H(\alpha_\varepsilon, \beta_\varepsilon)$ satisfying

$$\langle \alpha_\varepsilon, x\rangle > \beta_\varepsilon > \langle \alpha_\varepsilon, y\rangle, \quad \forall x \in X_\varepsilon, y \in Y_\varepsilon. \tag{23.131}$$

(d) Prove that if the hyperplane $H(\alpha_\varepsilon, \beta_\varepsilon)$ is the hyperplane constructed in the proof of Theorem 23.39 (page 944), then the sequence $(\alpha_\varepsilon, \beta_\varepsilon)$ has an accumulation point (α_*, β_*) when ε converges to 0.

(e) Prove that the hyperplane $H(\alpha_*, \beta_*)$ separates X from Y, i.e., $X \subseteq H^+(\alpha_*, \beta_*)$ and $Y \subseteq H^-(\alpha_*, \beta_*)$.

23.47 Find the dual program to each of the following linear programs:

(a)

$$\begin{array}{ll} \text{Compute:} & Z_P := \max\{2y_1 - 3y_2\}, \\ \text{subject to:} & y_1 + 2y_2 \leq 1, \\ & -y_1 - 4y_2 \leq 4, \\ & -6y_1 + 5y_2 \leq 7, \\ & y_1, y_2 \geq 0. \end{array} \tag{23.132}$$

(b)

$$\begin{array}{ll} \text{Compute:} & Z_P := \max\{6y_1 + 2y_2 - 3y_3\}, \\ \text{subject to:} & -2y_1 + 4y_2 \leq -3, \\ & 5y_1 - 1y_3 \leq -6, \\ & 7y_1 + 2y_2 - 4y_3 \leq 17, \\ & y_1, y_2 \geq 0. \end{array} \tag{23.133}$$

23.48 Prove that for every feasible solution x of Problem (23.105) on page 948 there exists a feasible solution (w, z) of Problem (23.104) satisfying $x = w - z$.

23.49 Express the dual program in Equation (23.98) (page 948) in the form of a primal program, as in Equation (23.97), and show that the dual program to this program is the primal program that appears in Equation (23.97).

23.50 Consider the following maximization problem:

$$\begin{array}{ll} \text{Compute:} & Z_P := \max cy^\top, \\ \text{subject to:} & Ay^\top = b^\top. \end{array} \tag{23.134}$$

(a) Represent this problem as a standard linear program using Definition 23.40. Note that in this problem there is no constraint of the form $y \geq 0$.
(b) Show that the dual program is

$$\begin{array}{ll} \text{Compute:} & Z_D := \min xb^\top, \\ \text{subject to:} & xA = c. \end{array} \tag{23.135}$$

23.51 Consider the following maximization problem:

$$\begin{array}{ll} \text{Compute:} & Z_P := \max cy^\top, \\ \text{subject to:} & Ay^\top \leq b^\top. \end{array} \tag{23.136}$$

(a) Represent this problem as a standard linear program using Definition 23.40. Note that in this problem there is no constraint of the form $y \geq 0$.
(b) Show that the dual program is

$$\begin{array}{ll} \text{Compute:} & Z_D := \min xb^\top, \\ \text{subject to:} & xA = c, \\ & x \geq \vec{0}. \end{array} \tag{23.137}$$