1. GEOMETRY AND LINEAR ALGEBRA

1.1. Convex Sets

Euclidean space $\mathbb{R}^n$ consists of all ordered n-tuples, x, of real numbers. Here x is written as a column and its *ith* entry is written x_i, $i = 1$ to n. $\mathbb{R}^n$ is a real linear space (or vector space) with operations defined entrywise. If $x,y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$,

$$(x + y)_i := x_i + y_i \quad \text{and} \quad (\alpha x)_i := \alpha x_i , \quad i = 1 \text{ to } n .$$

For x,y in $\mathbb{R}^n$, the *inner product* $\langle x,y \rangle$ is the real number $\sum_{i=1}^{n} x_i y_i$ and the *length* $\|x\|$ of x is $\sqrt{\langle x,x \rangle}$, that is, the non-negative number $\sqrt{(x_1^2 + x_2^2 + \ldots + x_n^2)}$; x is called a *unit* vector if $\|x\| = 1$. In $\mathbb{R}^n$ the *canonical unit vectors* e_i, $i = 1$ to n, are defined, for $j = 1$ to n, by

$$(e_i)_j := \begin{cases} 1, & j = i , \\ 0, & j \neq i . \end{cases}$$

If $x,y \in \mathbb{R}^n$ then x and y are *orthogonal* if $\langle x,y \rangle = 0$. A subset S of $\mathbb{R}^n$ is called *orthonormal* if it consists solely of unit vectors which are pairwise orthogonal. If $x \in \mathbb{R}^n$, then $\langle x,e_i \rangle = x_i$, $i = 1$ to n, and x has the *canonical orthonormal decomposition* as a linear combination of e_1 to e_n expressed by

$$x = \sum_{i=1}^{n} x_i e_i = \sum_{i=1}^{n} \langle x,e_i \rangle e_i .$$

Example 1.1. If $x = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$, $y = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$ in $\mathbb{R}^2$, then $x + y = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$;

$(-\frac{1}{2})x = \begin{pmatrix}-1\\ \frac{1}{2}\end{pmatrix}$; $\langle x,y \rangle = -2 - 3 = -5$; $\|x\| = 5$; $e_1 = \begin{pmatrix}1\\0\end{pmatrix}$ and $e_2 = \begin{pmatrix}0\\1\end{pmatrix}$; $x_1 := \langle x,e_1 \rangle = 2$ and $x_2 := \langle x,e_2 \rangle = -1$. The canonical orthonormal decomposition of x is $x = 2e_1 - e_2$. See Figure 1.1. //

A subset C of $\mathbb{R}^n$ is said to be *convex* if $\lambda x + (1 - \lambda)y \in C$ whenever $x,y \in C$ and $0 \leqslant \lambda \leqslant 1$. This simply asserts that for every pair of points in C the straight line segment linking these points lies wholly within C. Evidently, the empty set, each singleton $\{x\}$ and the whole space $\mathbb{R}^n$ are all convex in $\mathbb{R}^n$.

Example 1.2. (a) In $\mathbb{R}^2$ straight lines, half-planes, circular discs, and triangle interiors are all convex. See Figure 1.2.
(b) In $\mathbb{R}^2$ doubletons, complements of circular discs, triangle interiors less one point, and triangle interiors together with two vertices are all examples of sets which are not convex. //

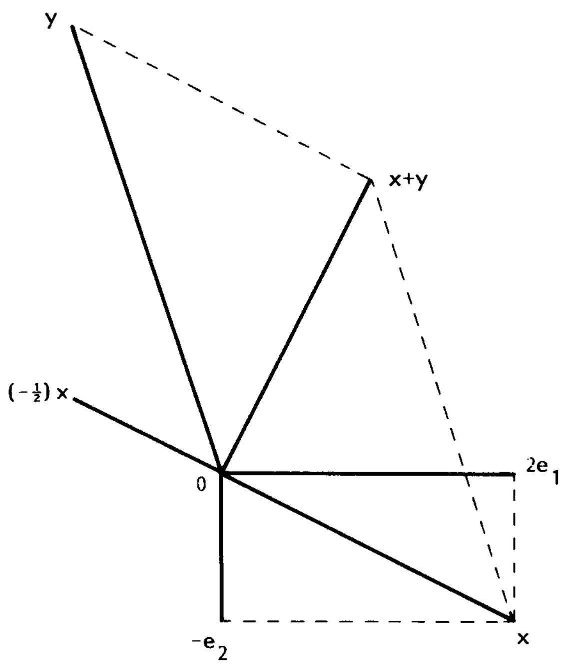

Figure 1.1. See Example 1.1.

The union of convex sets need not be convex. For example, consider convex C, D non-parallel straight lines in $\mathbb{R}^2$; then $z = \frac{1}{2}x + \frac{1}{2}y$ is not in $C \cup D$, if x is in C, y is in D and $x \neq y$.

Theorem 1.1. *If* F *is a family of convex subsets of* $\mathbb{R}^n$ *then* $D = \cap\{C \mid C \in F\}$ *is convex.*

Proof. If the intersection is empty, it is vacuously convex. Suppose it is not empty and let $x, y \in D$, $0 \leqslant \lambda \leqslant 1$. Then $x, y \in C$ for each C in F and C is convex so $z = \lambda x + (1 - \lambda)y \in C$. Thus $z \in D$, so D is convex. //

If $a \in \mathbb{R}^n$, $a \neq 0$, and $b \in \mathbb{R}$ the set $\{x \in \mathbb{R}^n \mid \langle a, x\rangle = b\}$, defined by the equation $a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$, is called a

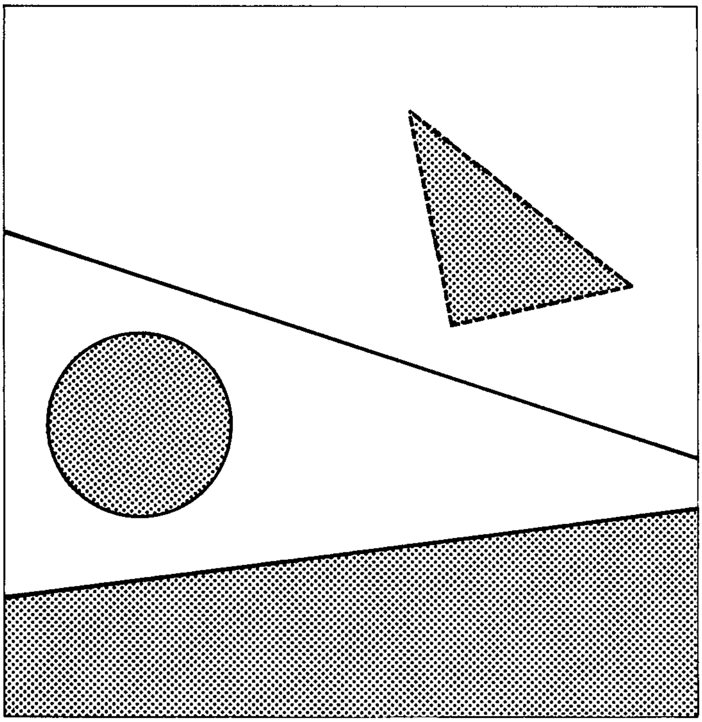

Figure 1.2. See Example 1.2.(a).

hyperplane in $\mathbb{R}^n$. (In $\mathbb{R}^2$ such sets are the straight lines and in $\mathbb{R}^3$ the planes.) The set $\{x \in \mathbb{R}^n | \langle a,x \rangle \leqslant b\}$, defined by the inequality

$$a_1x_1 + a_2x_2 + \dots + a_nx_n \leqslant b, \qquad (1.1)$$

is called a *closed half-space* in $\mathbb{R}^n$. (In $\mathbb{R}^2$ such sets are half-planes.) Note that $\{x \in \mathbb{R}^n | \langle a,x \rangle \geqslant b\}$ can be alternatively written $\{x \in \mathbb{R}^n | \langle -a,x \rangle \leqslant -b\}$ so the former are also half-spaces; all closed half-spaces can be expressed in the form (1.1). The complement of a closed half-space is called an *open* half-space; it has the form $\{x \in \mathbb{R}^n | \langle a,x \rangle > b\}$.

Theorem 1.2. *Every half-space in* $\mathbb{R}^n$ *is convex.*

Proof. Suppose, without loss of generality, that the defining inequality is $\langle a,x \rangle \leqslant b$. Let x,y belong to the half-space and $0 \leqslant \lambda \leqslant 1$. Then, from the definition of inner product,

$$\langle a, \lambda x + (1-\lambda)y \rangle = \lambda\langle a,x \rangle + (1-\lambda)\langle a,y \rangle.$$

But $\lambda \geqslant 0$, $1-\lambda \geqslant 0$, $\langle a,x \rangle \leqslant b$ and $\langle a,y \rangle \leqslant b$ so

$$\langle a, \lambda x + (1-\lambda)y \rangle \leqslant \lambda b + (1-\lambda)b = b. \;/\!/$$

Corollary 1. *Every hyperplane in* $\mathbb{R}^n$ *is convex.*

Proof. This can be proved directly or by noting that a hyperplane with defining equation $\langle a,x \rangle = b$ is the intersection of two convex sets, namely half-spaces with defining inequalities $\langle a,x \rangle \geqslant b$ and $\langle a,x \rangle \leqslant b$. //

Corollary 2. *The first orthant* $\{x \in \mathbb{R}^n | x_i \geqslant 0,\ i = 1 \text{ to } n\}$ *of* $\mathbb{R}^n$ *is convex.*

Proof. This can be proved directly or by noting that $x_i \geqslant 0$ if and only if $\langle e_i,x \rangle \geqslant 0$ so that the first orthant is convex, as the intersection of n half-spaces. //

It is conventional to write $x \geqslant 0$ iff $x_i \geqslant 0$, $i = 1$ to n.

Let S be a subset of $\mathbb{R}^n$. Consider F, the family of *all* convex sets in $\mathbb{R}^n$ containing S. Certainly F is not empty since, for example, $\mathbb{R}^n$ is convex and contains S. Then the set $\langle S \rangle$ defined as $\cap \{C | C \in F\}$ is convex and contains S; it is in fact the smallest convex set containing S and is called the *convex hull* of S. Clearly $\langle S \rangle = S$ if and only if S is convex. (Other notations, such as $co(S)$, are often used in the literature instead of $\langle S \rangle$.)

Theorem 1.3. *Let* C *be a convex subset of* $\mathbb{R}^n$. *If* $x_i \in C$, $\lambda_i \geq 0$, $i = 1$ *to* k, *and* $\lambda_1 + \lambda_2 + \ldots + \lambda_k = 1$ *then the convex combination* $\sum_{i=1}^{k} \lambda_i x_i$ *of* x_1 *to* x_k, *belongs to* C.

Proof. We prove this by induction on k.
If $k = 1$ the assertion is simply $x_1 \in C \Rightarrow x_1 \in C$, evidently true.
Suppose the result is true for k. Then (for $\lambda_{k+1} \neq 1$)

$$\sum_{i=1}^{k+1} \lambda_i x_i = \sum_{i=1}^{k} \lambda_i x_i + \lambda_{k+1} x_{k+1} = (1 - \lambda_{k+1}) \sum_{i=1}^{k} \mu_i x_i + \lambda_{k+1} x_{k+1}$$

where $\mu_i = \lambda_i/(1 - \lambda_{k+1})$, $i = 1$ to k.

But then $\mu_i \geq 0$, $i = 1$ to k and $\sum_{i=1}^{k} \mu_i = \frac{\sum_{i=1}^{k} \lambda_i}{1-\lambda_{k+1}} = \frac{1 - \lambda_{k+1}}{1 - \lambda_{k+1}} = 1$,

so by the result for k, $y = \sum_{i=1}^{k} \mu_i x_i$ is in C.

Immediately, by convexity of C, $\sum_{i=1}^{k+1} \lambda_i x_i = (1 - \lambda_{k+1})y + \lambda_{k+1} x_{k+1}$ is in C. //

Theorem 1.4. *Let* S *be a non-empty subset of* $\mathbb{R}^n$. *Then* $x \in \langle S \rangle$ *if and only if there exist* x_i *in* S, $\lambda_i \geq 0$, $i = 1$ *to* k, *for some positive integer* k, *where* $\lambda_1 + \lambda_2 + \ldots + \lambda_k = 1$, *such that* $x = \sum_{i=1}^{k} \lambda_i x_i$.

Proof. ($\Leftarrow$) The 'if' part follows directly from Theorem 1.3 where $C := \langle S \rangle$.

($\Rightarrow$) It is easily shown that the set C of all convex combinations from within S,

$$C := \{ \sum_{i=1}^{k} \lambda_i x_i \mid x_i \in S,\ \lambda_i \geq 0,\ i = 1 \text{ to } k,\ \sum_{i=1}^{k} \lambda_i = 1,\ k \geq 1\}$$

is convex. Namely, consider $y = \sum_{i=1}^{k} \lambda_i y_i$ and $z = \sum_{j=1}^{\ell} \mu_j z_j$ where $y_i \in S$, $\lambda_i \geq 0$, $i = 1$ to k, $\sum_{i=1}^{k} \lambda_i = 1$, and $z_j \in S$, $\mu_j \geq 0$, $j = 1$ to ℓ, $\sum_{j=1}^{\ell} \mu_j = 1$, and let $0 \leq \lambda \leq 1$. Then $\lambda y + (1 - \lambda)z = \sum_{i=1}^{k} \lambda\lambda_i y_i + \sum_{j=1}^{\ell} (1 - \lambda)\mu_j z_j$ where $\lambda\lambda_i \geq 0$, $i = 1$ to k, $(1 - \lambda)\mu_j \geq 0$, $j = 1$ to ℓ, and $\sum_{i=1}^{k} \lambda\lambda_i + \sum_{j=1}^{\ell} (1 - \lambda)\mu_j$ $= \lambda \sum_{i=1}^{k} \lambda_i + (1 - \lambda) \sum_{j=1}^{\ell} \mu_j = \lambda + (1 - \lambda) = 1$. Also this set of convex combinations contains S (each x in S can be written as $x = 1x$). By the definition of $\langle S \rangle$ as the intersection of *all* convex supersets of S we deduce that $\langle S \rangle$ is contained in C. //

Thus the convex hull of S is the set of all (finite) convex combinations from within S. (Compare this with $[S]$, the *subspace* spanned by a non-empty subset S of $\mathbb{R}^n$ consisting of all (finite) *linear* combinations from within S.) Also see exercise 1.8.5.

Example 1.3. The convex hull of $\{e_1, e_2\}$ in $\mathbb{R}^2$ is the line segment joining e_1 and e_2. The convex hull of $\{0, e_1, e_2\}$ in $\mathbb{R}^2$ is the (closed) triangular region with vertices $0, e_1, e_2$. See Figure 1.3. If C and D are non-parallel straight lines in $\mathbb{R}^2$ the convex hull of $C \cup D$ is the whole plane $\mathbb{R}^2$. //

A particularly simple kind of convex set is a *polytope*, that is a convex set C such that $C = \langle S \rangle$ for some *finite* subset S. (What is here described as a polytope is called by some authors a polyhedron. A polytope, in our terminology, is bounded, while to us the first orthant is an example of a convex polyhedron. See section 1.8.) Examples in $\mathbb{R}^3$ are, for example, compact convex polygonal regions, and compact convex polyhedral volumes. See Figure 1.4. (For clarity, only the skeleton of the polyhedral volume is shown; its convex hull is the polytope.) Counterexamples of convex sets which are not polytopes include straight lines, circular discs, and spherical balls.

If C is a non-empty convex subset of $\mathbb{R}^n$ and $x \in C$ then x is called a *vertex* (or *extreme point*) of C if it is not internal to any line segment in C, that is if $y, z \in C$, $0 < \lambda < 1$ and $x = \lambda y + (1 - \lambda)z$ then necessarily $x = y = z$.

Convex polytopes have the obvious corners as vertices while open circular discs have no vertices. The vertices of a closed circular disc in $\mathbb{R}^2$ are its infinitely many boundary points. It is a theorem, unproved here, that each convex polytope C can be expressed as the convex hull of its set, V, of vertices, $C = \langle V \rangle$, and furthermore V is a *minimal* subset S of C such that $C = \langle S \rangle$. For example, the polytope $\langle 0, \frac{1}{2}e_1 + \frac{1}{4}e_2, e_1, e_2 \rangle$ has vertices 0, e_1 and e_2 and can be minimally expressed as $\langle 0, e_1, e_2 \rangle$; see Figure 1.3. If a convex polytope in $\mathbb{R}^n$ has $n + 1$ vertices then it is known as an n-dimensional *simplex*; for example in $\mathbb{R}^3$ the tetrahedral volume $\langle 0, e_1, e_2, e_3 \rangle$ is a simplex.

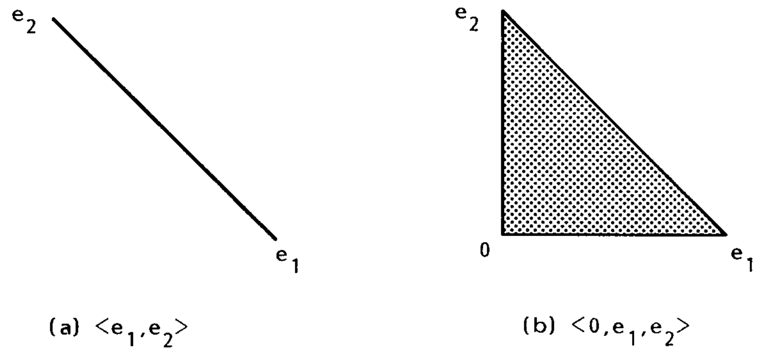

Figure 1.3. See Example 1.3.

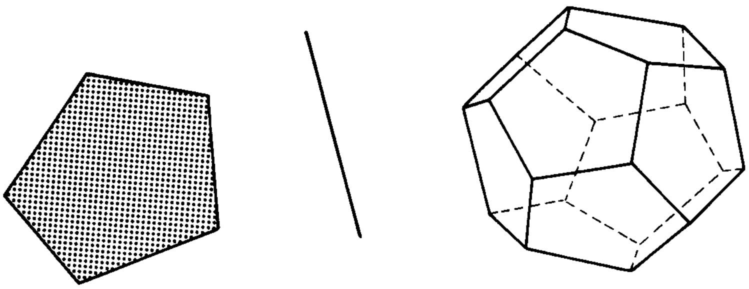

Figure 1.4. Examples of convex polytopes in $\mathbb{R}^3$.

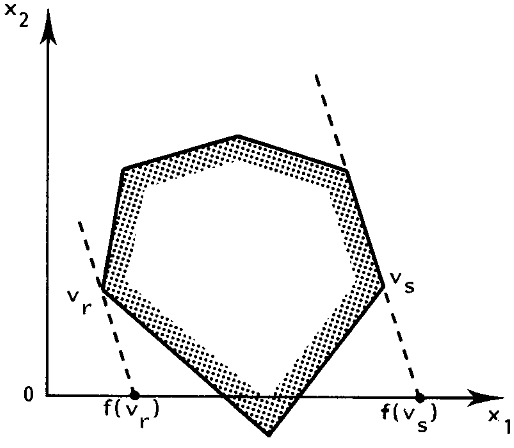

Figure 1.5. See Theorem 1.5. (Here $n = 2$, $k = 6$.)

If $c \in \mathbb{R}^n$, the function $f: \mathbb{R}^n \to \mathbb{R}$ defined by

$$f(x) := \langle c,x \rangle = c_1x_1 + c_2x_2 + \ldots + c_nx_n$$

is known as a *linear form* (or linear *functional*) on $\mathbb{R}^n$.

Theorem 1.5. *If* C *is a convex polytope in* $\mathbb{R}^n$ *and* f *is a linear form on* $\mathbb{R}^n$ *then* $\min_{x \in C} f(x)$ *and* $\max_{x \in C} f(x)$ *both exist and are achieved at vertices of* C.

Proof. Writing $C = \langle V \rangle$ where $V = \{v_1, v_2, \ldots, v_k\}$ is the finite set of vertices of C, if $x \in C$ then

$$x = \sum_{i=1}^{k} \lambda_i v_i \quad \text{for some } \lambda_i \geq 0,\ i = 1 \text{ to } k, \text{ with } \sum_{i=1}^{k} \lambda_i = 1 .$$

Then $f(x) := \langle c,x \rangle = \sum_{i=1}^{k} \lambda_i \langle c,v_i \rangle = \sum_{i=1}^{k} \lambda_i f(v_i)$.

Certainly there exist v_r, v_s in V such that, for all $i = 1$ to k,

$$f(v_r) \leq f(v_i) \leq f(v_s) .$$

Then

$$f(v_r) = f(v_r) \sum_{i=1}^{k} \lambda_i \leq \sum_{i=1}^{k} \lambda_i f(v_i) \leq f(v_s) \sum_{i=1}^{k} \lambda_i = f(v_s) .$$

That is, $f(v_r) \leq f(x) \leq f(v_s)$. //

For example, for the convex polytope C in $\mathbb{R}^2$ with vertices $\binom{1}{1}, \binom{2}{0}, \binom{2}{2}, \binom{4}{1}$ and $f(x) := 2x_1 - 3x_2$, see exercise 1.8.10.

1.2. Independence, Bases and Dimension

A non-empty finite set $S = \{y_1, y_2, \ldots, y_k\}$ in $\mathbb{R}^n$ is said to be *linearly dependent* if it contains 0 or at least one member which is a linear combination of the others. Otherwise the set is *linearly independent*. Independence of S means equivalently

$$\left(\alpha_i \in \mathbb{R},\ i = 1 \text{ to } k, \text{ and } \sum_{i=1}^{k} \alpha_i y_i = 0\right) \Rightarrow (\alpha_i = 0,\ i = 1 \text{ to } k).$$

For example, $S_1 := \left\{\begin{pmatrix}1\\-1\end{pmatrix}, \begin{pmatrix}2\\3\end{pmatrix}\right\}$ is independent in $\mathbb{R}^2$. However $S_2 := \left\{\begin{pmatrix}1\\-1\end{pmatrix}, \begin{pmatrix}2\\3\end{pmatrix}, \begin{pmatrix}-1\\0\end{pmatrix}\right\}$ is dependent since, for example, $\begin{pmatrix}-1\\0\end{pmatrix} = -\frac{3}{5}\begin{pmatrix}1\\-1\end{pmatrix} - \frac{1}{5}\begin{pmatrix}2\\3\end{pmatrix}$. See Figure 1.6.

If S is linearly independent it is said to be a *basis* for $[S]$, the subspace of $\mathbb{R}^n$ spanned by S. In this situation each member x of $[S]$ has a *unique* representation as a linear combination of y_1 to y_k. For example, S_1 as defined above is a basis for $[S_1] = \mathbb{R}^2$ and the representation of $x = \begin{pmatrix}x_1\\x_2\end{pmatrix}$ in $\mathbb{R}^2$ as $\alpha_1\begin{pmatrix}1\\-1\end{pmatrix} + \alpha_2\begin{pmatrix}2\\3\end{pmatrix}$ is unique for given x, namely $\alpha_1 = \frac{3}{5}x_1 - \frac{2}{5}x_2$, $\alpha_2 = \frac{1}{5}x_1 + \frac{1}{5}x_2$. Although

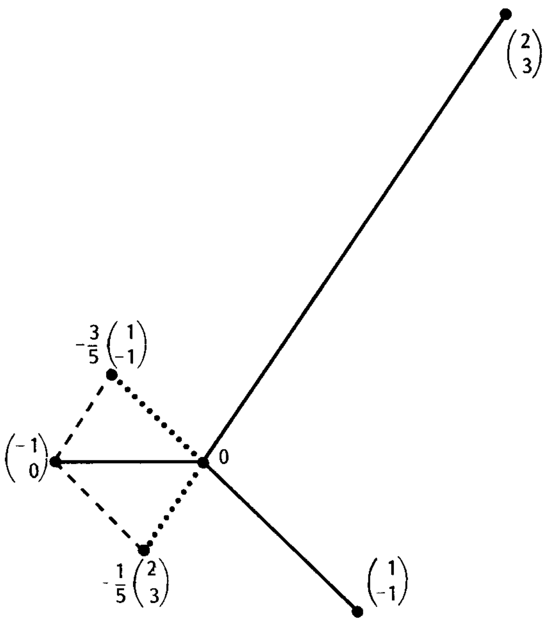

Figure 1.6. $S_2 := \left\{\begin{pmatrix}1\\-1\end{pmatrix}, \begin{pmatrix}2\\3\end{pmatrix}, \begin{pmatrix}-1\\0\end{pmatrix}\right\}$ is dependent.