| | |
|---|---|
| Title | 有限次元ベクトル空間の辞書式分離とその応用 |
| Author(s) | 池田，潔 |
| Citation | |
| Issue Date | 2003-03 |
| Type | Thesis or Dissertation |
| Text version | author |
| URL | http://hdl.handle.net/10119/936 |
| Rights | |
| Description | Supervisor:小野 寛晰，情報科学研究科，博士 |

# Lexicographical Separation in Finite-Dimensinal Vector Spaces and its Applications

by

Kiyoshi IKEDA

**submitted to**
**Japan Advanced Institute of Science and Technology**
**in partial fulfillment of the requirements**
**for the degree of**
**Doctor of Philosophy**

*Supervisor:* Professor Hiroakira Ono

*School of Information Science*
*Japan Advanced Institute of Science and Technology*

Feb 14, 2003

# Abstract

The aim of the thesis is to prove a lexicographical separation theorem and to give its applications to linear inequality systems, lexicographic expected utility, and extensive measurement.

The main theorem of the thesis is a lexicographical separation theorem stating that a convex cone and its convex complement in $\mathbb{F}^n$ can be separated by linear functions and a lexicographic order, where $\mathbb{F}$ stands for an ordered field such that $\mathbb{Q} \subseteq \mathbb{F} \subseteq \mathbb{R}$. We also give several other versions of lexicographical separation theorems, all of which are obtained from the main theorem. We provide a proof of the main theorem from our original standpoint, which makes use of the fact that a lexicographic order can be described by polynomials whose variable is an infinitesimal.

As one of the applications of the lexicographical separation theorem, we give a necessary and sufficient condition for the existence of solutions to infinite systems of linear inequalities, where the solutions are allowed to be polynomials whose variable is an infinitesimal. The result is a generalization of the well-known "theorem of the alternatives" for linear inequality systems. We also give a Farkas type theorem for lexicographical inequality systems.

As other application of the lexicographical separation theorem, we presented two kinds of lexicographic utility representations: one is about lexicographic expected utility, and the other is about lexicographic extensive utility. The lexicographic expected utility representation given in this thesis is a modification of Hausner's lexicographic expected utility theory, by omitting the existence of irrational-valued probabilities: we restrict our attention to rational-valued probabilities, and show that lexicographic expected utility theory can be founded on the domain of rational-valued lotteries. On the other hand, the lexicographical extensive utility representation given in this thesis is a modification of classical Hahn's embedding theorem: we establish a scheme of conditions which is necessary and sufficient for the existence of extensive utilities on indivisible items.

# Acknowledgements

First of all, I greatly appreciate the help of my supervisor Professor Hiroakira Ono, who gave me advice and encouraged me every time. I should like to thank Professor Mamoru Kaneko for his critical reading of the manuscript. I am also indebted to my vice-supervisor Professor Milan Vlach, who invited me to work at this subject. I am grateful to Professor Juan Enrique Martínez-Legaz for his helpful comments and suggestions. Lastly, I thank my colleague Yousuke Katou for our fruitful discussions, which inspired the idea of the thesis.

# Contents

# Chapter 1

# Introduction

In the past century, the importance of lexicographic orders was recognized by several authors, including Hahn [14], Hausner and Wendel [18], Chipman [4], Martínez-Legaz and Singer [30]. These authors showed that a variety of relations can be represented by lexicographic orders, and proved lexicographical representation theorems without continuity or an Archimedean condition. The present thesis is written along the line of this literature. In this thesis we prove another lexicographical representation theorem, and provide its applications to linear inequality systems, lexicographic expected utility, and extensive measurement.

The main result of this thesis strengthens the lexicographical representation theorems of the above mentioned literature. The result is concerned with separation of two convex sets in a finite-dimensional vector space over $\mathbb{F}$, where $\mathbb{F}$ denotes an arbitrary ordered subfield of real numbers including the field of rational numbers. We present applications of the main result to linear inequality systems and to lexicographic utility theory. Some of the results in this thesis are presented in terms of a polynomial ring whose variable is infinitely small, which gives us a clear picture of lexicographic orders.

In Section 1.1 we first explain the backgrounds of the thesis. In Section 1.2 we give an overview of our results, and in Section 1.3 we mention the organization of the thesis.

## 1.1 Backgrounds

In this section, we survey (i) separation theorems, (ii) finite systems of linear inequalities, (iii) extensive measurement of preference, (iv) von Neumann–Morgenstern utility functions. In each subsection we also mention our contributions in this thesis.

### Separation Theorems

The notion of separation is one of the most fundamental notions in convexity theory. We say that two disjoint sets $C_1$ and $C_2$ in $\mathbb{R}^n$ are *separated* if there exists a hyperplane such that $C_1$ is contained in one of the closed half-spaces associated with the hyperplane and $C_2$ lies in the opposite closed half-space. It is known that there exist several separation theorems stating that two disjoint convex sets[1] in $\mathbb{R}^n$ can be separated by a hyperplane. Here we shall present a version:

[1]See Section 2.2 for the definition of convex sets.

**Proposition 1.1.1** (Separating Hyperplane Lemma) Let $B$ be a closed convex set in $\mathbb{R}^n$, and let $x = (x_1, \cdots, x_n)$ be a point not in $B$. Then there exist real numbers $p_1, \ldots, p_n,\ p_{n+1}$ such that

$$\sum_{i=1}^{n} p_i\, x_i \;=\; p_{n+1}\,, \tag{1.1}$$

and

$$\sum_{i=1}^{n} p_i\, y_i \;>\; p_{n+1} \quad \text{for all } y = (y_1, \cdots, y_n) \in B\,. \tag{1.2}$$

( Geometrically, this means that there exists a hyperplane through $x$ such that $B$ lies entirely "above" the hyperplane.)

**Proof.** See Appendix. □

There are many other versions of separation theorems. See Rockafellar [39]. This author provides comprehensive information about the role of separation theorems in convexity theory and their applications. We will see in the next subsection that the separating hyperplane lemma can directly be applicable to the theory of finite linear inequality systems.

The separating hyperplane lemma (Proposition 1.1.1) relys on the assumption that a given convex set should be closed. However, it is sometimes difficult to verify whether a given set in a topological space is closed or not; so that from a theoretical point of view it is important to examine the consequences of removing this assumption. As will be seen next, dropping this assumption modifies the separating hyperplane lemma (Proposition 1.1.1) by allowing "lexicographical" separation. Let $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$. (In the following, the elements of $\bar{\mathbb{R}}^n$ will be regarded as column vectors, and the superscript $^T$ will mean transpose.) We consider the lexicographic ordering[2] $<_L$ on $\bar{\mathbb{R}}^n$, that is, given $x = (x_1, \ldots, x_n)^T$, $y = (y_1, \ldots, y_n)^T \in \bar{\mathbb{R}}^n$ with $x \neq y$, we have $x <_L y$ if $x_k < y_k$ for $k = \min\{\, i \mid x_i \neq y_i \,\}$. We also write $x \leqq_L y$ if $x <_L y$ or $x = y$. The following theorem is due to Martínez-Legaz [27]:

**Proposition 1.1.2** Let $B$ be a convex set in $\mathbb{R}^n$, and let $x = (x_1, \cdots, x_n)^T$ be a point not in $B$. Then there exist an orthogonal matrix $A$ of order $n$ and a vector $t \in \bar{\mathbb{R}}^n$ such that

$$Ax \;\leq_L\; t \tag{1.3}$$

and

$$Ay \;>_L\; t \quad \text{for all } y = (y_1, \cdots, y_n)^T \in B\,. \tag{1.4}$$

□

That is to say, a convex set and a point not in the convex set can be separated by a linear operator and the lexicographic ordering on $\bar{\mathbb{R}}^n$. For a proof see Martínez-Legaz [27], who provided a simple induction proof of Proposition 1.1.2.

Other theorems on lexicographical separation in $\mathbb{R}^n$ can be found in Hausner and Wendel [18], Klee [23], Martínez-Legaz and Singer [30]; all these theorems are essentially equivalent, in the sense that they are all concerned with lexicographic separation of two

[2]See Definition 2.1.2 for the precise definition of the lexicographic ordering.

convex sets in $\mathbb{R}^n$. In Chapter 3 we will prove a modification of the above lexicographical separation theorems, by considering not only the real numbers $\mathbb{R}$ but also other ordered fields $\mathbb{F}$: we will show that a convex cone and its convex complement in $\mathbb{F}^n$ can be separated lexicographically (where $\mathbb{F}$ stands for an ordered field such that $\mathbb{Q} \subseteq \mathbb{F} \subseteq \mathbb{R}$).

### Finite Systems of Linear Inequalities

In this subsection, we present an application of the separating hyperplane lemma (Proposition 1.1.1) to finite system of linear inequalities. Let $A$ be a finite set of points in $\mathbb{R}^n$. Consider the following system of (strict) linear inequalities:

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n \; > \; 0 \quad \text{for all} \;\; (a_1, a_2, \ldots, a_n) \in A\,. \tag{1.5}$$

If $A$ includes $(0, 0, \ldots, 0)$, for example, then obviously the system (1.5) has no solutions $x_1, x_2, \ldots, x_n$. This indicates that the solvability of the system (1.5) depends on the structure of $A$.

It is known that the following proposition, called "the theorem of the alternatives," gives a necessary and sufficient condition for the existence of solutions $x_1, \ldots, x_n$ to (1.5). Let $\langle x, y\rangle$ denote the inner product of $x$ and $y$, i.e. $\langle x, y\rangle = x_1y_1 + \cdots + x_ny_n$ where $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$.

**Proposition 1.1.3** (Theorem of the Alternatives) Let $a_i \in \mathbb{R}^n$ for $i = 1, \ldots, m$. Then exactly one of the following alternatives holds:

(i) There exist nonnegative real numbers $p_1, \ldots, p_m$ such that $\sum_{i=1}^m p_i = 1$ and

$$\sum_{i=1}^{m} p_i\, a_i \; = \; 0\,.$$

(ii) There exists a vector $x \in \mathbb{R}^n$ such that

$$\langle a_i,\, x\rangle \; > \; 0 \qquad \text{for } i = 1, \ldots, m\,.$$

**Proof.** This can be proved as a consequence of the separating hyperplane lemma (Proposition 1.1.1). See Appendix. □

There are several versions of linear existence theorems which give solutions to finite system of linear inequalities; they are known by various names, including the theorem of the alternatives, Farkas' lemma, Motzkin's lemma, Gordan's lemma, and so on. See Gale [12], Rockafellar [39], Slaka [42] for more comprehensive information on finite systems of linear inequalities. In Chapter 4, we will treat infinite systems of linear inequalities, and give a generalization of the theorem of the alternatives by allowing an "infinitely small" solution.

We shall introduce another linear existence theorem known as Farkas' lemma. An inequality $\langle a_0,\, x\rangle \le 0$ is said to be a *consequence* of the system

$$\langle a_i,\, x\rangle \; \le \; 0 \qquad \text{for } i = 1, \ldots, m \tag{1.6}$$

if it is satisfied by every $x$ which satisfies the system (1.6).

**Proposition 1.1.4** (Farkas' Lemma) Let $a_i \in \mathbb{R}^n$ for $i = 0, 1, \ldots, m$. Then, $\langle a_0, x \rangle \leq 0$ is a consequence of the system

$$\langle a_i, x \rangle \leq 0 \qquad \text{for } i = 1, \ldots, m$$

if and only if there exist nonnegative real numbers $p_1, \ldots, p_m$ such that

$$\sum_{i=1}^{m} p_i \, a_i = a_0 .$$

□

For a proof see e.g. Rockafellar [39]. The role of Farkas' lemma in linear programming theory was discussed in Kuhn and Tucker [24]. In Chapter 4, we will give a generalization of Farkas' lemma for lexicographical inequality systems.

## Extensive Measurement of Preference

In this subsection, we introduce the theory of fundamental measurement of preference. For more detailed discussion, see Roberts [38]. Measurement of preference is considered as an analog of measurement of temperature or mass: in the case of temperature, measurement is the assignment of numbers that preserve the observed relation "warmer than;" in the case of mass, the relation preserved is the relation "heavier than." Thus, measurement of preference is assignment of numbers preserving the observed relation "preferred to." If $S$ is a set of alternatives and $a \prec b$ holds if and only if you prefer $a$ to $b$, then we would like to assign a real number $u(a)$ to each $a \in S$ such that for all $a, b \in S$,

$$a \prec b \quad \text{iff} \quad u(a) < u(b) . \tag{1.7}$$

The function $u$ is called a *utility function* or an *ordinal utility function* or an *order-preserving utility function*, and the value $u(a)$ is called the *utility* of $a$.

To give a concrete example, let { coffee, tea, juice } be a set of alternatives, and suppose we have a preference order "coffee $\succ$ tea $\succ$ juice" (that is, coffee $\succ$ tea, tea $\succ$ juice, and coffee $\succ$ juice). Then, there exists a utility function $u$ preserving the order $\succ$, such as $u_1(\text{coffee}) = 3$, $u_1(\text{tea}) = 2$, and $u_1(\text{juice}) = 1$, or $u_2(\text{coffee}) = 100$, $u_2(\text{tea}) = 10$, and $u_2(\text{juice}) = 0$.

As observed in the above example, there are several different utility functions preserving the same preference order. Hence, it comes into question which function is to be adopted. The answer depends on how demanding we want to be in our measurement.

In the case of mass, we actually demand more of our measure. We want it to be "additive" in the sense that the mass of the combination of two objects is the sum of their masses. Formally, we speak of a binary operation $\circ$ on the set $S$ of objects. We want a real-valued function $f$ on $S$ that not only satisfies (1.7) but also preserves the binary relation $\circ$, in the sense that

$$f(a \circ b) = f(a) + f(b) . \tag{1.8}$$

There might be a comparable operation in the case of preference; we might want to allow compound alternatives, such as coffee and suger $(a \circ b)$, and we might want to require utility to be additive, that is, to satisfy

$$u(a \circ b) = u(a) + u(b) . \tag{1.9}$$

A utility function that is also additive is often called a *cardinal utility function.*

Additive properties, such as mass, have traditionally been called *extensive* in the literature of measurement, and so the problem of finding (necessary and) sufficient conditions for the existence of a cardinal utility function is called the problem of *extensive measurement*. See Roberts [38], Narens [36] for the developments on extensive measurement.

It is well-known that classical Hölder's Theorem gives sufficient conditions for extensive measurement. A system $\langle A, R, \circ, e\rangle$ is called an *Archimedean ordered group* if it satisfies the following conditions:

O1 $\langle A, \circ, e\rangle$ is a group,

O2 $R$ is a total order on $A$,

O3 (Monotonicity) for all $x, y, z \in A$,

$$xRy \quad \text{iff} \quad (x \circ z)\, R\, (y \circ z) \quad \text{iff} \quad (z \circ x)\, R\, (z \circ y)\,,$$

O4 (Archimedean) for all $x, y \in S$, if $eRy$ then there is a positive integer $k$ such that $x\, R\, ky$.

It can easily be verified that $\langle \mathbb{R}, <, +, 0\rangle$ is an Archimedean ordered group. A system $\langle A, R, \circ, e\rangle$ is said to be *homomorphic* to $\langle \mathbb{R}, <, +, 0\rangle$ if there exists a real-valued function $u$ on $A$ such that (1.7) and (1.9) hold for all $a, b \in A$.

**Proposition 1.1.5** (Hölder's Theorem) Every Archimedean ordered group is homomorphic to $\langle \mathbb{R}, <, +, 0\rangle$. □

For a proof see e.g. Birkhoff [1]. See also Roberts [38] for other modifications of Hölder's Theorem giving sufficient conditions for extensive measurement.

Hahn's Theorem [14] gave a non-Archimedean extension of Hölder's Theorem. Hahn [14] showed that it is possible to omit the Archimedean condition O4 by introducing lexicographically ordered vectors in place of real-values. In Chapter 5 we will give a modification of Hahn's Theorem, giving a necessary and sufficient condition for "lexicographical" extensive measurement.

## von Neumann–Morgenstern Utility Functions

In this subsection, we consider measurement of preference in the context of decisions under risk. In brief, we think of a risky situation by assigning "probabilities" to the alternatives. For example, if there are two alternatives $a$, $b$ to be chosen under risk, we assign a probability $\lambda$ to $a$, and $(1-\lambda)$ to $b$, and therefore the situation is expressed by the combination $\lambda a \circ (1-\lambda)b$. If there exists a preference $\prec$ on such a situation, we would like to find a utility function $u$ which preserves the combination, in the sense that

$$u(\, \lambda a \circ (1-\lambda)b\,) \;=\; \lambda u(a) \,+\, (1-\lambda)u(b)\,. \tag{1.10}$$

This means that utility of the combination of two alternatives is equivalent to the "expected value" of the utilities of the alternatives. Such expected value of utilities is called *expected utility*.

Von Neumann and Morgenstern [45] were the first to develop the theory of expected utility, establishing necessary and sufficient conditions for the existence of a utility function satisfying (1.7) and (1.10). Their expected utility theory has been modified by several authors, including Herstein and Milnor [19], Jensen [21], Fishburn [6]. Let us summarize their results as follows:

Let $S$ be a set of items. A *probability distribution on $S$ with finite support* or, more briefly, a *lottery on $S$* is a mapping $\lambda : S \to [0,1]$ with the following properties:

$$\text{(i)} \sum_{s\in S} \lambda(s) = 1, \quad \text{(ii) there is a finite } S' \subseteq S \text{ such that } \lambda(s) = 0 \text{ for all } s \in S \setminus S'.$$

Let $\Delta(S)$ denote the set of all lotteries. Given $\lambda, \mu \in \Delta(S)$ and $\theta \in [0,1]$, we define the *convex combination* $\theta\lambda + (1-\theta)\mu$ by

$$[\,\theta\lambda + (1-\theta)\mu\,](s) \;=\; \theta\lambda(s) + (1-\theta)\mu(s) \quad \text{for all } s \in S\,.$$

One can easily verifies that the convex combination $\theta\lambda + (1-\theta)\mu$ also belongs to $\Delta(S)$. Let $\prec$ be a preference relation on $\Delta(S)$, where $x \prec y$ means "$y$ is preferred to $x$." Consider the following axioms: For all $x, y, z$ in $\Delta(S)$,

L1 $\prec$ is a weak order[3] on $\Delta(S)$,

L2 if $x \prec y$ then $\theta x + (1-\theta)z \;\prec\; \theta y + (1-\theta)z$ for all $\theta \in (0,1)$,

L3 if $x \prec y$ and $y \prec z$, then there are $\theta, \xi \in (0,1)$ such that

$$\theta x + (1-\theta)z \;\prec\; y \quad \text{and} \quad y \;\prec\; \xi x + (1-\xi)z\,.$$

These axioms L1, L2 and L3 are refered to as the ordering, independence, and Archimedean axioms, respectively.

**Proposition 1.1.6** (Jensen [21]) A preference relation $\prec$ on $\Delta(S)$ satisfies L1–L3 if and only if there is a utility function $u : S \to \mathbb{R}$ such that for all $\lambda, \mu \in \Delta(S)$,

$$\lambda \prec \mu \quad \text{iff} \quad \sum_{s\in S} \lambda(s)u(s) \;<\; \sum_{s\in S} \mu(s)u(s)\,. \tag{1.11}$$

Moreover, $u$ is unique up to a positive affine transformation, i.e. if $v : S \to \mathbb{R}$ is another function with the same property, then there exist real constants $a > 0$ and $b$ such that

$$v(s) = a\,u(s) + b \quad \text{for all } s \in S\,.$$

**Proof.** See e.g. Jensen [21], Fishburn [6], Hammond [16]. □

That is to say, if we assume the axioms L1–L3, we can obtain a utility function $u : S \to \mathbb{R}$ satisfying (1.11) for all $\lambda, \mu \in \Delta(S)$, which preserves both the relation $\prec$ and the "comvex combination." See Fishburn [6] [9] [10], Hammond [16] for more information about the foundations of expected utility and its role in decision theory.

[3] See Section 2.1 for the definition of weak orders.

It has sometimes been pointed out in the literature that the role of the Archimedean axiom L3 in single person decision theory is merely technical, namely, to ensure the existence of a real-valued utility function. Hausner [17], Chipman [4], Fishburn [7] [10], Nakamura [33] [34] developed lexicographic extensions of the classical expected utility theory without the Archimedean axiom L3, introducing lexicographically ordered vectors in place of real-values. Readers should consult the excellent survey paper of Martínez-Legaz [29] for the recent developments on lexicographic utility.

As can be seen in the definition, the expected utility theory is founded on the domain of real-valued lotteries. In Chapter 5, we will give a modification of Hausner's lexicographic expected utility theory by omitting the existence of irrational-valued lotteries.

## 1.2 Overview

In this section, we give an overview of our results in this thesis. Let $<_L$ (or $>_L$) denote the lexicographic order on $\mathbb{R}^n$, that is, given $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ with $x \neq y$, we have $x <_L y$ if $x_k < y_k$ for $k = \min\{ i \mid x_i \neq y_i \}$.

The main result of the thesis is the following lexicographical separation theorem (see Theorem 3.1.1 in Chapter 3):

**Main Theorem** *Let $\mathbb{F}^n$ be the $n$-dimensional vector space over $\mathbb{F}$, where $\mathbb{F}$ stands for an ordered field such that $\mathbb{Q} \subseteq \mathbb{F} \subseteq \mathbb{R}$. Let $P$ be a positive cone*[4] *in $\mathbb{F}^n$ with its complement $\mathbb{F}^n \setminus P$ being a convex cone in $\mathbb{F}^n$. Then, there exist real-valued linear functions $g_1, \ldots, g_n$ on $\mathbb{F}^n$ such that for all $x \in \mathbb{F}^n$,*

$$x \in P \quad \textit{iff} \quad (\, g_1(x), \ldots, g_n(x)\,) >_L (\, 0, \ldots, 0\,).$$

□

This means that, in any finite-dimensional vector space over $\mathbb{F}$, a convex cone $P$ and its convex complement $\mathbb{F}^n \setminus P$ can be separated by a set of linear functions and a lexicographic order. We also show a kind of uniqueness of the linear functions.

In case $\mathbb{F} = \mathbb{R}$, equivalent versions of this theorem was proved by Hausner and Wendel [18], Klee [23], Martínez-Legaz and Singer [30]; so that the above theorem is a generalization of their theorems, considering an ordered field $\mathbb{F}$ other than $\mathbb{R}$.

We give a proof of the lexicographical separation theorem from our original standpoint, using an *infinitely small number*, or an *infinitesimal* $\varepsilon$, that is, $0 < \varepsilon$ and $\varepsilon < 1/k$ for all positive integer $k$. Let us introduce $\mathbb{R}[\varepsilon]$ the smallest ring containing both $\mathbb{R}$ and an infinitesimal $\varepsilon$, i.e.

$$\mathbb{R}[\varepsilon] \;=\; \{\, r_0 + r_1\varepsilon + \cdots + r_n\varepsilon^n \mid n \in \mathbb{N},\; r_0, r_1, \ldots, r_n \in \mathbb{R}\,\}.$$

In this thesis, we use the following fact that the lexicographic order on $\mathbb{R}^n$ can be described by the polynomials in $\mathbb{R}[\varepsilon]$ (see Definition 2.1.5 in Chapter 2):

**Proposition** *Let $\varepsilon$ be an infinitesimal. For all $a_0, a_1, \ldots, a_n$, $b_0, b_1, \ldots, b_n \in \mathbb{R}$,*

$$(a_0, a_1, \ldots, a_n) <_L (b_0, b_1, \ldots, b_n) \quad \textit{iff} \quad a_0 + a_1\varepsilon + \cdots + a_n\varepsilon^n < b_0 + b_1\varepsilon + \cdots + b_n\varepsilon^n. \tag{1.12}$$

[4] In this thesis, a *positive cone* is defined as a convex cone in which the origin is not contained. See Section 2.2.

□

We shall call $\mathbb{R}[\varepsilon]$ *the set of lexicographically ordered polynomials*. The correspondence (1.12) between the lexicographically ordered vectors and the lexicographically ordered polynomials will play an important role in this thesis: the lexicographical separation theorem will be proved in terms of the lexicographically ordered polynomials, rather than the lexicographically ordered vectors. The use of the polynomials makes the proof easier, since they allow both addition and multiplication (different from the lexicographically ordered vectors, which allow only addition).

The correspondence (1.12) also enables us to give a new role to an infinitesimal: in this thesis, we adopt an infinitesimal $\varepsilon$ as a solution to infinite systems of linear inequalities (as will be seen below).

As one of the applications of the lexicographical separation theorem, we obtain a generalization of the well-known "theorem of the alternatives," giving a necessary and sufficient condition for the existence of solutions to strict linear inequality systems (see Theorem 4.1.2 in Chapter 4):

**Theorem** *Let $P$ be an nonempty subset of $\mathbb{R}^n$. Then, the origin $\mathbf{0}$ is not contained in the convex hull of $P$ if and only if the inequality system*

$$\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n \;>\; 0 \quad \textit{for all} \;\; (\lambda_1, \lambda_2, \ldots, \lambda_n) \in P$$

*has solutions $x_1, x_2, \ldots, x_n$ in $\mathbb{R}[\varepsilon]$.*

□

In Chapter 4 we will provide several examples of strict linear inequality systems, from which we will see that it is not unreasonable to obtain such an infinitesimal $\varepsilon$ in our solutions. As another application, we also obtain a generalization of Farkas' lemma for lexicographical inequality systems. Further, we applied these results to game theory, giving a generalization of von Neumann's minimax theorem for semi-infinite games.

As an application of the lexicographical separation theorem to expected utility theory, we show that a similar form of the lexicographic expected utility representation considered by Hausner [17] is valid even if we omit the existence of irrational-valued probabilities: in this thesis we restrict our attention to rational-valued probabilities, in order not to consider situations in which an event occurs with an irrational-valued probability. Let $S_n = \{\alpha_1, \ldots, \alpha_n\}$ be a finite set of items, and let $\Delta(S_n)$ denote the set of all rational-valued probability distributions on $S_n$, i.e.

$$\Delta(S_n) = \{\, p_1\,\alpha_1 \,+\, \cdots \,+\, p_n\,\alpha_n \mid p_i \in \mathbb{Q},\; p_i \geq 0 \;\text{ for } i = 1, \ldots, n \;\text{ with } \sum_{i=1}^{n} p_i = 1 \,\}.$$

Suppose there exists a weak order[5] $\prec$ on $\Delta(S_n)$ satisfying the following independence conditions: for all $x, y, z \in \Delta(S_n)$ and all $0 < \theta < 1$ in $\mathbb{Q}$,

I1 if $x \prec y$ then $\theta x + (1-\theta)z \;\prec\; \theta y + (1-\theta)z$,

I2 if $x \sim y$ then $\theta x + (1-\theta)z \;\sim\; \theta y + (1-\theta)z$.

[5]See Section 2.1 for the definitions of a weak order $\prec$ and its associated indifference relation $\sim$.

As will be seen next, these independence conditions for a weak order $\prec$ on $\Delta(S_n)$ are necessary and sufficient for the existence of a finite-dimensional linear utility function on $\Delta(S_n)$ whose lexicographic order preserves the ordering $\prec$ (see Theorem 5.1.2 in Chapter 5):

**Theorem** *A weak order $\prec$ on $\Delta(S_n)$ satisfies the independence conditions (I1 and I2) if and only if there exists an $(n-1)$-dimensional utility function $U = (u_1, \ldots, u_{n-1})$ on $\Delta(S_n)$ such that for all $\sum_{i=1}^{n} p_i\alpha_i$ , $\sum_{i=1}^{n} q_i\alpha_i$ in $\Delta(S_n)$ ,*

$$\sum_{i=1}^{n} p_i\alpha_i \prec \sum_{i=1}^{n} q_i\alpha_i \quad \textit{iff} \quad \sum_{i=1}^{n} p_i\, U(\alpha_i) <_L \sum_{i=1}^{n} q_i\, U(\alpha_i)\,. \qquad \square$$

We also show that the first component $u_1$ is unique up to a positive affine transformation. Thus, when all the events are guaranteed to occur with rational-valued probabilities, we do not need to assume the existence of "irrational-valued probabilities." The above lexicographic expected utility representation can be derived as a consequence of the lexicographical separation theorem, considering lexicographical separation of two convex sets in a finite-dimensional vector space over $\mathbb{Q}$.

Lastly, we obtain the following lexicographical extensive measurement theorem, giving a necessary and sufficient condition for the existence of extensive utilities on indivisible items. Let $S_n = \{\alpha_1, \ldots, \alpha_n\}$ be a finite set of indivisible items, and let $\Omega(S_n)$ denote the set of all *consumption plans* on $S_n$ , i.e.

$$\Omega(S_n) = \{ k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n \mid k_1, k_2 \ldots, k_n \in \mathbb{N} \}\,.$$

(A consumption plan $k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n$ means that, intuitively, a consumer consumes $k_i$ pieces of $\alpha_i$ for each $i = 1, \ldots, n$.) Suppose there exists a preference relation $\prec$ on $\Omega(S_n)$ satisfying the following conditions:

- A1 $\prec$ is a weak order on $\Omega(S_n)$,
- A2 for all $x, y, z \in \Omega(S_n)$, $x \prec y$ implies $x + z \prec y + z$,
- A3 for all $x, y, z \in \Omega(S_n)$, $x \sim y$ implies $x + z \sim y + z$.

The following theorem states that the scheme of conditions A1–A3 is necessary and sufficient for extensive measurement of preferences (see Theorem 5.2.1 in Chapter 5):

**Theorem** *A preference relation $\prec$ on $\Omega(S_n)$ satisfies A1–A3 if and only if there are lexicographically ordered polynomials $q_1, \ldots, q_n \in \mathbb{R}[\varepsilon]$ such that for all $k_1, \ldots, k_n, l_1, \ldots, l_n \in \mathbb{N}$,*

$$\sum_{i=1}^{n} k_i\alpha_i \prec \sum_{i=1}^{n} l_i\alpha_i \quad \textit{iff} \quad \sum_{i=1}^{n} k_i q_i < \sum_{i=1}^{n} l_i q_i\,. \qquad \square$$

This means that the conditions A1–A3 ensure the existence of utilities $q_1, \ldots, q_n$ of the items $\alpha_1, \ldots, \alpha_n$ , respectively, which can be added one another freely. In Chapter 5, we will discuss the meaning of the conditions A1–A3 in the context of economics.

## 1.3 Organization of the Thesis

In Chapter 2, we define algebraic concepts which will be needed in the subsequent chapters. The main theme of this chapter is to show that (i) a weak order on a vector space can be described by a convex cone, (ii) a lexicographic order can be described by polynomials whose variable is an infinitesimal ( which will be called *lexicographically ordered polynomials* in this thesis ) . These results are well-known in the literature, but we rearranged them for our later convenience. The description of a lexicographic order by means of an infinitesimal will play an important role in this thesis: we use an infinitesimal not only for the description of a lexicographic order but also as a useful tool of proving a lexicographical separation theorem in Chapter 3.

In Chapter 3, we present our main results of the thesis. The main theorem is a lexicographical separation theorem stating that a convex cone and its convex complement in $\mathbb{F}^n$ can be separated by linear functions and a lexicographic order, where $\mathbb{F}$ stands for an ordered field such that $\mathbb{Q} \subseteq \mathbb{F} \subseteq \mathbb{R}$. We also give several other versions of lexicographical separation theorems, all of which are obtained from the main theorem. We provide a proof of the main theorem from our original standpoint, using the lexicographically ordered polynomials introduced in Chapter 2. The proof is a bit lengthy, in order to derive a kind of uniqueness result. Applications of the main theorem will be discussed in Chapter 4 and Chapter 5.

In Chapter 4, we apply the lexicographical separation theorem to linear inequality systems. We give a necessary and sufficient condition for the existence of solutions to infinite systems of linear inequalities, where the solutions are allowed to be polynomials whose variable is an infinitesimal. The result is a generalization of the well-known theorem of the alternatives for finite linear inequality systems. We also give a Farkas type theorem for lexicographical inequality systems. Further, we apply these results to game theory, giving a generalization of von Neumann's minimax theorem for semi-infinite games.

In Chapter 5, we present two kinds of lexicographic utility representations: one is about lexicographic expected utility, and the other is about lexicographic extensive utility. These two representations are derived in similar manners from the lexicographical separation theorem. The lexicographic expected utility representation given in this thesis is a modification of Hausner's lexicographic expected utility theory, by omitting the existence of irrational-valued lotteries. On the other hand, the lexicographical extensive utility representation is a modification of classical Hahn's (purely mathematical) embedding theorem, by giving a necessary and sufficient condition for the existence of extensive utilities on indivisible items.

In Chapter 6, we summarize our results and suggest future work.

Throughout this thesis, " **Theorem** " signifies our original contribution.

# Chapter 2

# Algebraic Preliminaries

In this chapter we shall define some algebraic concepts which will be needed in the subsequent chapters. In Section 2.1 we define weak orders and lexicographic orders. Also we introduce a new concept called the *lexicographically ordered polynomial ring* (Definition 2.1.5), showing that lexicographic orders can be described by the polynomial ring in which a variable is infinitely small. The ring of lexicographically ordered polynomials is considered as a substructure of the ordered field of hyperreal numbers (Remark 2.1.7). In Section 2.2 we define the notion of convex cones in linear spaces, and discuss the relationship between convex cones and weak orders. Section 2.3 is a counterpart of Section 2.2 for Abelian group theory. Section 2.4 is a brief review of standard linear algebra. For the understanding of the algebraic concepts in this chapter, readers are assumed to have the knowledge of groups, rings, fields, and vector spaces (see e.g. Lang [25]).

## 2.1 Lexicographically Ordered Polynomials

Let us begin with the standard definitions of orderings (see e.g. Roberts [38]). A binary relation $\prec$ on $S$ is called a *weak order* if it satisfies the following conditions:

(i) for all $x, y \in S$, $\quad x \prec y$ implies not $(y \prec x)$,

(ii) for all $x, y, z \in S$, $\quad$ not $(x \prec y)$ and not $(y \prec z)$ imply not $(x \prec z)$.

Let $\prec$ be a weak order on $S$. Its associated *indifference relation* $\sim$ is defined by

$$x \sim y \quad \text{iff} \quad \text{not}\,(x \prec y) \ \text{and} \ \text{not}\,(y \prec x)\,.$$

We also write $x \precsim y$ iff $x \prec y$ or $x \sim y$. A weak order $\prec$ on $S$ is said to be *nontrivial* if there exist $x, y \in S$ such that $x \prec y$.

**Proposition 2.1.1** Let $\prec$ be a weak order on $S$. Then, for all $x, y, z \in S$,

(a) exactly one of $x \prec y$, $\ y \prec x$, $\ x \sim y$ holds,

(b) $x \prec y$ and $y \prec z$ imply $x \prec z$,

(c) $\sim$ is an equivalence relation, i.e. (i) $x \sim x$,
(ii) $x \sim y$ implies $y \sim x$, (iii) $x \sim y$ and $y \sim z$ imply $x \sim z$,

(d) $x \prec y$ and $y \sim z$ imply $x \prec z$; $\quad x \sim y$ and $y \prec z$ imply $x \prec z$,

(e) $\precsim$ satisfies the following conditions:

(i) $x \precsim y$ and $y \precsim z$ imply $x \precsim z$, $\quad$ (ii) $x \precsim y$ or $y \precsim x$.

**Proof.** (a) Since $\prec$ is a weak order, we cannot have both $x \prec y$ and $y \prec x$. If not $(x \prec y)$ and not $(y \prec x)$ then this is equivalent to $x \sim y$.

(b) Suppose there exist $x, y, z \in S$ such that $x \prec y$, $y \prec z$, and not $(x \prec z)$. From $y \prec z$ we have not $(z \prec y)$. Therefore, not $(x \prec z)$ and not $(z \prec y)$ yield not $(x \prec y)$, a contradiction.

(c) (i) For any $x \in S$, we cannot have $x \prec x$: for, if $x \prec x$ then not $(x \prec x)$, a contradiction. Hence not $(x \prec x)$ for all $x \in S$, and therefore $x \sim x$ for all $x \in S$. Both (ii) and (iii) are straightforward from the definition of $\sim$.

(d) Suppose there exist $x, y, z \in S$ such that $x \prec y$, $y \sim z$, and not $(x \prec z)$. Then, $x \sim z$ or $z \prec x$ by (a). If $x \sim z$ then, with $y \sim z$, we have $x \sim y$, a contradiction. If $z \prec x$ then, with $x \prec y$, we have $z \prec y$ by (b), a contradiction. The second half of (d) is similarly proved.

(e) (i) is a consequence of (b), (c), and (d); (ii) is a consequence of (a). □

A weak order $\prec$ is called a *total order* if its associated indifference relation $\sim$ reduces to the identity,[1] that is, $x \sim y$ only if $x = y$. An example of a total order is the *lexicographic order* on $\mathbb{R}^n$ as defined below.

**Definition 2.1.2** Let $<$ be the usual strict order on the set of real numbers $\mathbb{R}$. A vector $\boldsymbol{x} = (x_1, \ldots, x_n)$ in $\mathbb{R}^n$ is said to be *lexicographically less than* $\boldsymbol{y} = (y_1, \ldots, y_n)$ if we have $x_k < y_k$ for the least $k$ such that $x_k \neq y_k$. In other words,

$$\boldsymbol{x} <_L \boldsymbol{y} \quad \text{iff} \quad \boldsymbol{x} \neq \boldsymbol{y} \quad \text{and} \quad x_k < y_k \quad \text{for} \quad k = \min\{\, i \mid x_i \neq y_i \,\},$$

where $<_L$ denotes "lexicographically less than." We also write $\boldsymbol{x} \leqq_L \boldsymbol{y}$ iff $\boldsymbol{x} <_L \boldsymbol{y}$ or $\boldsymbol{x} = \boldsymbol{y}$. □

It can easily be verified that the lexicographic order $<_L$ is a total order on $\mathbb{R}^n$.

Let $\mathbb{R}[t]$ be the set of all polynomials over the real numbers, and let $\langle \mathbb{R}[t], +, \cdot, 0, 1 \rangle$ denote the polynomial ring. We shall show that a lexicographic order on $\mathbb{R}[t]$ can be defined in the same manner as Definition 2.1.2. To begin with, let us recall the standard definition of *ordered rings* (see e.g. Lang [25]).

A ring $\langle R, +, \cdot, 0, 1 \rangle$ is said to be *ordered* if there exists a subset $P$ of $R$, called the set of *positive elements*, satisfying the following properties:

(1) For all $x \in R$, exactly one of $x \in P$ or $x = 0$ or $-x \in P$ holds.

(2) For all $x, y \in R$, $\quad x, y \in P$ implies $x + y,\ x \cdot y \in P$.

We define $x < y$ to mean $y - x \in P$. By (1), for all $x, y \in R$, exactly one of $x < y$ or $x = y$ or $y < x$ holds. By (2), for all $x, y, x', y', z \in R$,

[1] A *total order* $\prec$ on $S$ is also defined by the following conditions: for all $x, y, z \in S$, (i) not $(x \prec x)$ (ii) $x \prec y$ and $y \prec z$ imply $x \prec z$ (iii) $x \prec y$ or $y \prec x$ or $x = y$. It is easy to check that these two definitions are equivalent.

(a) $x < y$ and $x' < y'$ imply $x + x' < y + y'$,

(b) $x < y$ and $z > 0$ imply $x \cdot z < y \cdot z$.

It is easy to verify that $<$ is a total order on $R$.

Conversely, every total order $<$ on $R$ satisfying (a) and (b) defines a set of positive elements: we put $P = \{x \in R \mid x > 0\}$. Thus we say that a total order $<$ on $R$ is *compatible* if it satisfies (a) and (b).

**Definition 2.1.3** Let $p(t) = a_0 + a_1 t + \cdots + a_m t^m$ and $q(t) = b_0 + b_1 t + \cdots + b_n t^n$ be polynomials in $\mathbb{R}[t]$ with $m \leq n$. We shall express $p(t) = a_0 + a_1 t + \cdots + a_n t^n$ with $a_{m+1} = \cdots = a_n = 0$. $p(t)$ is said to be *lexicographically less than* $q(t)$ if we have $a_k < b_k$ for the least $k$ such that $a_k \neq b_k$. In other words,

$$a_0 + a_1 t + \cdots + a_n t^n \ <_L \ b_0 + b_1 t + \cdots + b_n t^n \quad \text{iff} \quad \begin{cases} p(t) \neq q(t) & \text{and} \\ a_k < b_k \ \text{ for } \ k = \min\{\, i \mid a_i \neq b_i \,\}, \end{cases}$$

where $<_L$ denotes "lexicographically less than." □

It can easily be verified that the lexicographic order $<_L$ is a compatible total order on $\mathbb{R}[t]$. It also holds that, with respect to the lexicographic order $<_L$ on $\mathbb{R}[t]$, the monomial $t$ is *infinitely small* in the sense that

$$0 <_L t \quad \text{and} \quad t <_L 1/k \quad \text{for all positive integer } k.$$

We can show its converse:

**Lemma 2.1.4** Let $\langle \mathbb{R}[t], +, \cdot, 0, 1 \rangle$ be the polynomial ring, and let $<$ be a binary relation on $\mathbb{R}[t]$. Suppose that

(i) $<$ is a compatible total order on $\mathbb{R}[t]$,

(ii) the restriction of $<$ to the subring $\mathbb{R}$ is the usual strict ordering of the real numbers,

(iii) $t$ is infinitely small with respect to $<$, i.e. $0 < t$ and $t < 1/k$ for all positive integer $k$.

Then $<$ is the lexicographic order $<_L$ on $\mathbb{R}[t]$.

**Proof.** To derive the conclusion, it is sufficient to show the following equivalence:

$$c_0 + c_1 t + \cdots + c_n t^n < 0 \quad \Longleftrightarrow \quad c_k < 0 \ \text{ for } \ k = \min\{\, i \mid c_i \neq 0 \,\}, \tag{2.1}$$

because, if this is satisfied, the following holds for all $a_0, \ldots, a_n, b_0, \ldots, b_n \in \mathbb{R}$,

$$\begin{aligned} & a_0 + a_1 t + \cdots + a_n t^n < b_0 + b_1 t + \cdots + b_n t^n \\ \Longleftrightarrow \ & (a_0 - b_0) + (a_1 - b_1)t + \cdots + (a_n - b_n)t^n < 0 && \text{(by the compatibility of } <\text{)} \\ \Longleftrightarrow \ & a_k < b_k \ \text{ for } \ k = \min\{\, i \mid a_i \neq b_i \,\} && \text{(by (2.1))} \\ \Longleftrightarrow \ & a_0 + a_1 t + \cdots + a_n t^n <_L b_0 + b_1 t + \cdots + b_n t^n . \end{aligned}$$

(2.1) is equivalent to the following two conditions:

$$\text{if } \ c_k > 0 \ \text{ for } \ k = \min\{\, i \mid c_i \neq 0 \,\}, \quad \text{then } \ c_0 + c_1 t + \cdots + c_n t^n > 0, \tag{2.2}$$

$$\text{if } \ c_k < 0 \ \text{ for } \ k = \min\{\, i \mid c_i \neq 0 \,\}, \quad \text{then } \ c_0 + c_1 t + \cdots + c_n t^n < 0. \tag{2.3}$$

Hence it is enough to show (2.2). We can derive (2.3) in a parallel way with (2.2).

Suppose $p(t) = c_0 + c_1 t + \cdots + c_n t^n$ is a nonzero polynomial with $c_k > 0$ for $k = \min\{ i \mid c_i \neq 0 \}$. We shall show $p(t) > 0$.

First we consider the case $k = 0$, i.e. $c_0 > 0$. If $c_i \geq 0$ for all $i > 0$, then $p(t) > 0$ is trivial. Suppose there exists $c_i < 0$ for some $i > 0$. Let $c = \max\{ |c_i| \mid c_i < 0 \}$. Then there is a sufficiently large integer $M$ such that $0 < 1/M < c_0/nc$. Thus

$$p(t) = c_0 + c_1 t + \cdots + c_n t^n \;>\; c_0 - c(t + \cdots + t^n) \;>\; c_0 - nc/M \;>\; 0 .$$

[ Note that, under the assumptions (i)–(iii), $t^2$ is also infinitely small: for, we have $0 < t^2$ from $0 < t$ with the compatibility of $<$, and $t^2 < (1/m)t < 1/mn$ for all positive integers $m, n$. For the same reason, $t^3$, $t^4$, $\cdots$ are all infinitely small.]

For the case $k > 0$, let $q(t)$ be the polynomial $q(t) = c_k + c_{k+1} t + \cdots + c_n t^{n-k}$ such that $p(t) = q(t) t^k$. Then $c_k > 0$ and hence it is enough to apply the above argument to $q(t)$. □

That is to say, the lexicographic order $<_L$ can be described in terms of the ordered polynomial ring satisfying (i)–(iii) of Lemma 2.1.4.

**Definition 2.1.5** Let $\varepsilon$ be a variable, and let $\mathbb{R}[\varepsilon]$ be the set of all polynomials over $\mathbb{R}$, i.e.

$$\mathbb{R}[\varepsilon] \;=\; \{ r_0 + r_1\varepsilon + \cdots + r_n\varepsilon^n \mid n \in \mathbb{N},\; r_0, r_1, \ldots, r_n \in \mathbb{R} \} .$$

Let $\langle \mathbb{R}[\varepsilon], +, \cdot, 0, 1 \rangle$ denote the polynomial ring . We say that the polynomial ring is *lexicographically ordered* iff there is a relation $<$ on $\mathbb{R}[\varepsilon]$ satisfying the following conditions:

(i) $<$ is a total order on $\mathbb{R}[\varepsilon]$,

(ii) for all $x, y, x', y', z \in \mathbb{R}[\varepsilon]$,

(a) $x < y$ and $x' < y'$ imply $x + x' \,<\, y + y'$,

(b) $x < y$ and $z > 0$ imply $x \cdot z \,<\, y \cdot z$,

(iii) the restriction of $<$ to the subring $\mathbb{R}$ is the usual strict order on the real numbers,

(iv) $0 < \varepsilon$ and $\varepsilon < 1/k$ for all positive integer $k$.

By Lemma 2.1.4, the relation $<$ satisfying (i)–(iv) is the lexicographic order on $\mathbb{R}[\varepsilon]$, i.e.

$$a_0 + a_1\varepsilon + \cdots + a_n\varepsilon^n \;<\; b_0 + b_1\varepsilon + \cdots + b_n\varepsilon^n \quad \text{iff} \quad \begin{cases} (a_0, a_1, \ldots, a_n) \neq (b_0, b_1, \ldots, b_n) & \text{and} \\ a_k < b_k \;\text{ for }\; k = \min\{ i \mid a_i \neq b_i \} . \end{cases}$$

Such polynomials are called *lexicographically ordered polynomials.* □

Throughout this thesis, we denote by $\mathbb{R}[\varepsilon]$ the set of lexicographically ordered polynomials (as in Definition 2.1.5), assuming that there is a total order $<$ on it satisfying the conditions (i)–(iv) of Definition 2.1.5. Also, in this thesis, the notation $r_0 + r_1\varepsilon + \cdots + r_n\varepsilon^n$ always stands for a lexicographically ordered polynomial in $\mathbb{R}[\varepsilon]$.

**Definition 2.1.6** Let $n$ be a positive integer. We denote by $\mathbb{R}[\varepsilon]_n$ the set of lexicographically ordered polynomials whose degrees are less than $n$, i.e.

$$\mathbb{R}[\varepsilon]_n \;=\; \{\, r_1 + r_2\varepsilon + \cdots + r_n\varepsilon^{n-1} \mid r_1, \cdots, r_n \in \mathbb{R} \,\}.$$

Note that a polynomial $r_1 + r_2\varepsilon + \cdots + r_n\varepsilon^{n-1}$ can be identified with the lexicographically ordered vector $(r_1, r_2, \ldots, r_n)$ as was defined in Definition 2.1.2. □

Our notation $r_0 + r_1\varepsilon + \cdots + r_n\varepsilon^n$ can be understood in a parallel way with the *decimal number system*. The decimal number system is indeed a lexicographical order; readers should compare 2.13, 3.1 and 3.13 say. (Recall that 2.13 is identified with the real number $2 + 1 \times \frac{1}{10} + 3 \times (\frac{1}{10})^2$, and so on.)

**Remark 2.1.7** In our notation $r_0 + r_1\varepsilon + \cdots + r_n\varepsilon^n$ we use the expression "$\varepsilon$" for the variable. This is because the lexicographically ordered polynomial ring $\mathbb{R}[\varepsilon]$ is a proper subring of the *hyperreal numbers*.

The ordered field of hyperreal numbers is usually constructed as an ultrapower of the real numbers, and hence it has such properties as (1) including $\mathbb{R}$ as a subfield, (2) containing an infinitesimal $\varepsilon$, (3) satisfying the transfer principle, i.e. satisfying the same first-order sentences as $\mathbb{R}$. For more details, see e.g. Goldblatt [13]. It is easy to verify that the lexicographically ordered polynomial ring $\mathbb{R}[\varepsilon]$ can be embedded into the ordered field of hyperreal numbers, identifying the variable $\varepsilon$ with an infinitesimal.

Hammond [15] introduced the smallest subfield containing both $\mathbb{R}$ and an infinitesimal $\varepsilon$, which is a simplified substructure of the field of hyperreal numbers. The lexicographically ordered polynomial ring $\mathbb{R}[\varepsilon]$ is, therefore, a bit smaller substructure than Hammond's one. □

## 2.2 Convex Cones and Weakly–Ordered Spaces

Let $\mathbb{F}$ be an ordered field (e.g. the rational numbers $\mathbb{Q}$, or the real numbers $\mathbb{R}$), and let $V$ be a vector space over $\mathbb{F}$. A subset $C$ of $V$ is said to be *convex* if $\lambda x + (1-\lambda)y \in C$ whenever $x, y \in C$ and $\lambda \in \mathbb{F}$ with $0 < \lambda < 1$. Let $S$ be an arbitrary subset of $V$ and let $x \in V$. We say that $x$ is a *convex combination of* $S$ if there exist finite $y_1, \ldots, y_k \in S$ and $\lambda_1, \ldots, \lambda_k \in \mathbb{F}$ with $\lambda_i > 0$ for $i = 1, \ldots, k$ and $\sum_{i=1}^k \lambda_i = 1$ such that $x = \lambda_1 y_1 + \cdots + \lambda_k y_k$. The *convex hull* of $S$ is the set of all convex combinations of $S$.

Let $P$ be a subset of $V$. We say that $P$ is a *convex cone* if for all $x, y \in V$ and all $\lambda \in \mathbb{F}$,

(i) $x, y \in P$ implies $x + y \in P$, (ii) $x \in P$ and $\lambda > 0$ imply $\lambda x \in P$.

(It is easy to check that a convex cone $P$ is a convex set.) A *positive cone* $P$ is a convex cone in which $\mathbf{0}$ is not contained, i.e. (iii) $\mathbf{0} \notin P$ (where $\mathbf{0}$ denotes the origin of $V$). We denote by $\bar{P}$ the complement of $P$, that is, $\bar{P} = V \setminus P$. We write $-P$ to mean $\{-x \mid x \in P\}$. For a positive cone $P$, its associated *indifference subspace* $I$ is defined by $I = V \setminus (P \cup -P)$. Note that the indifference subspace is nonempty, since it contains at least $\mathbf{0}$.

**Definition 2.2.1** Let $V$ be a vector space over $\mathbb{F}$. Suppose there is a weak order $\prec$ on $V$. We say that a weak order $\prec$ is *compatible with the vector space structure of* $V$, or briefly, *compatible*, if for all $x, y, z \in V$ and all $\lambda \in \mathbb{F}$,

(i) $x \prec y$ implies $x + z \prec y + z$,

(ii) $x \prec y$ and $\lambda > 0$ imply $\lambda x \prec \lambda y$.

Such a vector space $V$ is called a *weakly–ordered vector space*. □

We shall show that a weakly–ordered vector space can be described by a positive cone and its convex complement:

**Lemma 2.2.2** Let $V$ be a vector space over $\mathbb{F}$, and suppose there is a subset $P$ of $V$ such that $P$ is a positive cone and $\bar{P}$ is a convex cone. Define a binary relation $\prec$ on $V$ by

$$x \prec y \quad \text{iff} \quad y - x \in P$$

Then $\prec$ is a compatible weak order on $V$.

**Proof.** First we show (i) and (ii) in Definition 2.2.1. (i) If $x \prec y$, then $y - x = (y + z) - (x + z) \in P$, therefore $x + z \prec y + z$. (ii) Suppose $x \prec y$ and $\lambda > 0$. This means $y - x \in P$ and $\lambda > 0$. Hence $\lambda(y - x) \in P$ and therefore $\lambda x \prec \lambda y$.

Next we show that $\prec$ is a weak order: Suppose there exist $x, y \in V$ such that $x \prec y$ and $y \prec x$. Then $y - x \in P$ and $x - y \in P$, hence $(y - x) + (x - y) = \mathbf{0} \in P$, a contradiction. Thus, for all $x, y \in V$, $x \prec y$ implies not$(y \prec x)$. Now, suppose not$(x \prec y)$ and not$(y \prec z)$. This means $y - x \in \bar{P}$ and $z - y \in \bar{P}$. Hence $(y - x) + (z - y) = z - x \in \bar{P}$. Therefore not$(x \prec z)$. □

The converse of Lemma 2.2.2 can also be shown:

**Lemma 2.2.3** Let $V$ be a vector space over $\mathbb{F}$, and suppose there is a compatible weak order $\prec$ on $V$. Let $P = \{ x \in V \mid x \succ \mathbf{0} \}$. Then $P$ is a positive cone and $\bar{P}$ is a convex cone.

**Proof.** ($P$ is a positive cone.) It is clear that $P$ does not contain $\mathbf{0}$. Suppose $x \succ \mathbf{0}$ and $y \succ \mathbf{0}$. This implies $x + y \succ y$ and $y \succ \mathbf{0}$. Hence, by Proposition 2.1.1, we have $x + y \succ \mathbf{0}$. Suppose $x \succ \mathbf{0}$ and $\lambda > 0$. Then, obviously $\lambda x \succ \mathbf{0}$.

($\bar{P}$ is a convex cone.) Suppose not$(x \succ \mathbf{0})$ and not$(y \succ \mathbf{0})$. The first condition implies not$(x + y \succ y)$: for, if $x + y \succ y$ then $(x + y) - y \succ y - y$, which means $x \succ \mathbf{0}$, a contradiction. Thus, both not$(x + y \succ y)$ and not$(y \succ \mathbf{0})$ yield not$(x + y \succ \mathbf{0})$. Now, suppose not$(x \succ \mathbf{0})$ and $\lambda > 0$. If $\lambda x \succ \mathbf{0}$, then, by multiplying $1/\lambda$, we have $x \succ \mathbf{0}$, a contradiction. Therefore not$(\lambda x \succ \mathbf{0})$. □

That is to say, Lemma 2.2.2 and Lemma 2.2.3 show that there is one-to-one correspondence between a compatible weak order on $V$ and a positive cone with its convex complement.

**Lemma 2.2.4** Let $V$ be a vector space over $\mathbb{F}$, and suppose there is a subset $P$ of $V$ such that $P$ is a positive cone and $\bar{P}$ is a convex cone. Let $I$ be the associated indifference subspace. Then,

(a) $I$ is a subspace of $V$, i.e. for all $x, y \in V$ and all $\lambda \in \mathbb{F}$,
(i) $x, y \in I$ implies $x + y \in I$ ; (ii) $x \in I$ implies $\lambda x \in I$,

(b) for all $x \in V$, exactly one of $x \in P$, $x \in I$, and $x \in -P$ holds,

(c) for all $x, y \in V$, $x \in P$ and $y \in P \cup I$ imply $x + y \in P$.

**Proof.** (a) First we note that (ii) is trivial in case $\lambda = 0$. By the definition, $I = \bar{P} \cap \overline{(-P)} = \bar{P} \cap -\bar{P}$. (Hence, if $x \in I$ then $-x \in I$.) Since $\bar{P}$ is a convex cone, $-\bar{P}$ is also a convex cone; hence their intersection $I$ is also a convex cone. Therefore (i) is obvious. It remains to show that (ii) holds for all $x \in V$ and all $\lambda \neq 0$. But, if $\lambda > 0$ then (ii) is obvious; if $\lambda < 0$ then both $x \in I$ and $-\lambda > 0$ yield $-\lambda x \in I$, which implies $\lambda x \in I$.

(b) It is enough to show that $P$, $I$, and $-P$ are mutually disjoint. $P \cap I = \emptyset$ and $-P \cap I = \emptyset$ are clear from the definition of $I$. Suppose $P \cap -P \neq \emptyset$. Let $x \in P \cap -P$. Then $x \in P$ and $-x \in P$, hence $x + (-x) = \mathbf{0} \in P$, a contradiction. Therefore $P \cap -P = \emptyset$.

(c) It is enough to show that, for all $x, y \in V$, $x \in P$ and $y \in I$ imply $x + y \in P$. Suppose there exist $x, y \in V$ such that $x \in P$, $y \in I$ and $x+y \notin P$. Then $x+y \in -P \cup I$ by (b). If $x + y \in -P$, then this means $-(x + y) \in P$, and therefore, with $x \in P$, we have $-y \in P$, contradicting to $y \in I$. If $x + y \in I$, then, with $y \in I$, we have $x \in I$ by (a), contradicting to $x \in P$. □

## 2.3 Weakly–Ordered Groups

As we have seen, a weakly–ordered vector space can be described by a positive cone and its convex complement. In this section we show that a similar description can also be applied to weakly–ordered Abelian groups, which we shall define below.

**Definition 2.3.1** Let $\langle S, +, 0\rangle$ be an Abelian group, and suppose there is a weak order $\prec$ on $S$. We say that a weak order $\prec$ is *compatible with addition*, or briefly, *compatible*, if

$$\text{for all } x, y, z \in S, \quad x \prec y \text{ implies } x + z \prec y + z .$$

Such an Abelian group $\langle S, +, 0\rangle$ is called a *weakly–ordered Abelian group*. □

Let $\langle S, +, 0\rangle$ be an Abelian group. Consider a pair of subsets $P$, $I$ ($\subseteq S$) satisfying the following conditions:

C1 $\langle I, +, 0\rangle$ is a subgroup of $\langle S, +, 0\rangle$, i.e. for all $x, y \in S$,
(i) $x, y \in I$ implies $x + y \in I$ ; (ii) $x \in I$ implies $-x \in I$,

C2 for all $x \in S$, exactly one of $x \in P$, $x \in I$, and $x \in -P$ holds,

C3 for all $x, y \in S$, $x \in P$ and $y \in P \cup I$ imply $x + y \in P$.

Note that these conditions are variations of (a), (b), and (c) in Lemma 2.2.4.

**Lemma 2.3.2** Let $\langle S, +, 0\rangle$ be an Abelian group, and suppose there is a pair of subsets $P$, $I$ $(\subseteq S)$ satisfying C1–C3. Define a binary relation $\prec$ on $S$ by

$$x \prec y \quad \text{iff} \quad y - x \in P\,.$$

Then $\prec$ is a compatible weak order on $\langle S, +, 0\rangle$.

**Proof.** The compatibility is obvious: if $x \prec y$, then $y - x = (y + z) - (x + z) \in P$, therefore $x + z \prec y + z$. We shall show that $\prec$ is a weak order on $S$. It is easy to verify that $\bar{P} = -P \cup I$ by C2, and for all $x, y \in S$,

(i) $x \in P$ implies $-x \notin P$ (by C2),

(ii) $x \in -P \cup I$ and $y \in -P \cup I$ imply $x + y \in -P \cup I$ (by C1 and C3).

By (i), for all $x, y \in S$, $x \prec y$ implies not $(y \prec x)$. By (ii), for all $x, y \in S$, not $(x \prec y)$ and not $(y \prec z)$ imply not $(x \prec z)$. □

The converse of Lemma 2.3.2 can also be shown:

**Lemma 2.3.3** Let $\prec$ be a compatible weak order on $\langle S, +, 0\rangle$. Define $P$, $I$ $(\subseteq S)$ as

$$P = \{\, x \in S \mid 0 \prec x \,\}, \qquad I = \{\, x \in S \mid 0 \sim x \,\}\,.$$

Then the pair $P$, $I$ satisfies C1–C3.

**Proof.** Left to readers. □

Thus Lemma 2.3.2 and Lemma 2.3.3 show that, for an Abelian group $\langle S, +, 0\rangle$, there is one-to-one correspondence between a compatible weak order on it and a pair of subsets $P$, $I$ satisfying C1–C3.

**Lemma 2.3.4** Let $\langle S, +, 0\rangle$ be an Abelian group, and $P$, $I$ be subsets of $S$ satisfying C1–C3.

(1) For any $x \in S$, if $nx \in P$ for some positive integer $n$, then $x \in P$.
($nx$ denotes $x + \cdots + x$ for $n$ times)

(2) For any $x \in S$, if $nx \in I$ for some positive integer $n$, then $x \in I$.

**Proof.** (1) If $x \notin P$, then $x \in -P \cup I$ by C2, and hence $-x \in P \cup I$ by C1 (ii). Applying C3 or C1 (i) repeatedly, $n(-x) = -nx \in P \cup I$. Therefore $-nx \notin -P$ by C2, and thus $nx \notin P$.
(2) The proof is similar to (1). □

Let $\mathbb{F}$ be an ordered field, and let $V$ be a vector space over $\mathbb{F}$. In Section 2.2 we have seen that a compatible weak order on $V$ can be described by a positive cone and its convex complement. We shall show that, in case $\mathbb{F}$ is the rational numbers $\mathbb{Q}$, a compatible weak order on $V$ can also be described by a pair of subsets $P$, $I$ satisfying C1–C3.

**Lemma 2.3.5** Let $V$ be a vector space over the rational numbers $\mathbb{Q}$, and let $P$, $I$ be subsets of $V$ satisfying C1–C3 (where $V$ is considered as an additive group $\langle S, +, 0\rangle$ ). Then, $P$ is a positive cone and $\bar{P}$ is a convex cone.
( Conversely, suppose $P$ is a positive cone and $\bar{P}$ is a convex cone. Let $I$ be the associated indifference subspace of $P$. Then $P$ and $I$ satisfy C1–C3. )

**Proof.** Note that the latter statement follows from Lemma 2.2.4.
Let $P$, $I$ be subsets of $V$ satisfying C1–C3.

( $P$ is a positive cone. ) It is enough to show that, for all $x \in V$ and all $\lambda \in \mathbb{Q}$, $x \in P$ and $\lambda > 0$ imply $\lambda x \in P$. Suppose $x \in P$ and $\lambda > 0$. Then, there are positive integers $m, n$ such that $\lambda = m/n$. By applying C3 repeatedly, $x \in P$ yields $mx \in P$. Hence, by Lemma 2.3.4 (1), we get $(m/n)\, x \in P$.

( $\bar{P}$ is a convex cone. ) By C2, we have $\bar{P} = -P \cup I$. We show that for all $x, y \in V$ and all $\lambda \in \mathbb{Q}$,

(1) $x, y \in -P \cup I$ implies $x + y \in -P \cup I$,

(2) $x \in -P \cup I$ and $\lambda > 0$ imply $\lambda x \in -P \cup I$.

For (1), suppose $x, y \in -P \cup I$. Then $-x, -y \in P \cup I$ by C1 (ii). Hence $-(x+y) \in P \cup I$ by C3 or C1 (i). Therefore $x + y \in -P \cup I$ by C1 (ii).
For (2), suppose $x \in -P \cup I$ and $\lambda > 0$. Then $-x \in P \cup I$ by C1 (ii), and also there are positive integers $m, n$ such that $\lambda = m/n$. By applying C3 or C1 (i) repeatedly, $-x \in P \cup I$ yields $m(-x) \in P \cup I$. Hence, by Lemma 2.3.4, we get $-(m/n)\, x \in P \cup I$. Therefore $(m/n)\, x \in -P \cup I$ by C1 (ii). □

Lemma 2.3.5 may not hold for another ordered field than $\mathbb{Q}$. Here we give an illustrative example. Let $\mathbb{Q}(\sqrt{2}) = \{\, x + y\sqrt{2} \mid x, y \in \mathbb{Q} \,\}$. Then $\mathbb{Q}(\sqrt{2})$ is an ordered subfield of $\mathbb{R}$. We consider $\mathbb{Q}(\sqrt{2})$ as a 1-dimensional vector space over $\mathbb{Q}(\sqrt{2})$. Let us define a pair of subsets $P_0$, $I_0$ ( $\subseteq \mathbb{Q}(\sqrt{2})$ ) by

$$P_0 \;=\; \{\; x + y\sqrt{2} \;\mid\; x - y\sqrt{2} \;>\; 0 \;\}, \qquad I_0 \;=\; \{\, 0 \,\}.$$

One can easily verifies that $P_0$ and $I_0$ satisfy C1–C3. However, $P_0$ is not a positive cone: for, we have $1 \in P_0$ and $\sqrt{2} > 0$, but $\sqrt{2} \notin P_0$.

## 2.4 Linear Spaces

Let $\mathbb{F}$ be an ordered field, and let $V$ be a vector space over $\mathbb{F}$. A function $u$ on $V$ is said to be *linear* if (i) $u(x + y) = u(x) + u(y)$ for all $x, y \in V$, (ii) $u(\lambda x) = \lambda u(x)$ for all $x \in V$ and all $\lambda \in \mathbb{F}$.

Here we recall some standard results in linear algebra (see e.g. Lang [25]). Let $F_n$ denote the $n$-dimensional vector space over $\mathbb{F}$ with the basis $\boldsymbol{e_1}, \ldots, \boldsymbol{e_n}$, i.e.

$$F_n = \{\, \lambda_1 \boldsymbol{e_1} + \cdots + \lambda_n \boldsymbol{e_n} \mid \lambda_1, \cdots, \lambda_n \in \mathbb{F} \,\}.$$

The *inner product* of $\boldsymbol{\lambda} = \lambda_1 \boldsymbol{e_1} + \cdots + \lambda_n \boldsymbol{e_n}$ and $\boldsymbol{\mu} = \mu_1 \boldsymbol{e_1} + \cdots + \mu_n \boldsymbol{e_n}$ is defined by

$$\boldsymbol{\lambda} \cdot \boldsymbol{\mu} \;=\; \lambda_1 \mu_1 + \lambda_2 \mu_2 + \cdots + \lambda_n \mu_n .$$

A set of nonzero vectors $\boldsymbol{\delta_1}, \boldsymbol{\delta_2}, \ldots, \boldsymbol{\delta_k}$ is called an *orthogonal system* if $\boldsymbol{\delta_i} \cdot \boldsymbol{\delta_j} = 0$ for all $i, j$ with $i \neq j$.

**Proposition 2.4.1** Let $\boldsymbol{\delta_1}, \boldsymbol{\delta_2}, \ldots, \boldsymbol{\delta_k}$ be an orthogonal system and let $\boldsymbol{\lambda} = \sum_{i=1}^{k} \lambda_i \, \boldsymbol{\delta_i}$ $(\lambda_i \in \mathbb{F})$. Then

$$\lambda_j = \frac{\boldsymbol{\lambda} \cdot \boldsymbol{\delta_j}}{\boldsymbol{\delta_j} \cdot \boldsymbol{\delta_j}} \quad \text{for } j = 1, \ldots, k,$$

and hence the orthogonal system $\boldsymbol{\delta_1}, \boldsymbol{\delta_2}, \ldots, \boldsymbol{\delta_k}$ is linearly independent.

**Proof.**

$$\boldsymbol{\lambda} \cdot \boldsymbol{\delta_j} \;=\; \sum_{i=1}^{k} \lambda_i \, \boldsymbol{\delta_i} \cdot \boldsymbol{\delta_j} \;=\; \lambda_j \, \boldsymbol{\delta_j} \cdot \boldsymbol{\delta_j} \,,$$

and if $\boldsymbol{\lambda} = 0$ then $0 = \lambda_j \, \boldsymbol{\delta_j} \cdot \boldsymbol{\delta_j}$ therefore $\lambda_j = 0$ for $j = 1, \ldots, k$. □

**Proposition 2.4.2** Let $H_k$ be a $k$-dimensional subspace of $F_n$ for $0 < k < n$ (i.e. there are linearly independent vectors $\boldsymbol{\eta_1}, \ldots, \boldsymbol{\eta_k}$ in $F_n$ such that $H_k = \{ \lambda_1 \boldsymbol{\eta_1} + \cdots + \lambda_k \boldsymbol{\eta_k} \mid \lambda_1, \ldots, \lambda_k \in \mathbb{F} \}$). Then, there is an orthogonal system $\boldsymbol{\delta_1}, \boldsymbol{\delta_2}, \ldots, \boldsymbol{\delta_k}$ in $H_k$ such that

$$H_k \;=\; \{ \lambda_1 \boldsymbol{\delta_1} + \cdots + \lambda_k \boldsymbol{\delta_k} \mid \lambda_1, \ldots, \lambda_k \in \mathbb{F} \}.$$

**Proof.** This can be proved by the following well-known method called "the orthogonalization of Schmidt". It suffices to show that " if $\boldsymbol{\delta_1}, \ldots, \boldsymbol{\delta_l}$ is an orthogonal system in $H_k$ for $l < k$, then there is $\boldsymbol{\delta_{l+1}} \in H_k$ such that $\boldsymbol{\delta_1}, \ldots, \boldsymbol{\delta_l}, \boldsymbol{\delta_{l+1}}$ is also an orthogonal system in $H_k$."

If $l < k$, then there is $\boldsymbol{\delta'} \in H_k$ which is not a linear combination of $\boldsymbol{\delta_1}, \ldots, \boldsymbol{\delta_l}$. Let

$$\boldsymbol{\delta_{l+1}} \;=\; \boldsymbol{\delta'} - \left( \sum_{i=1}^{l} \frac{\boldsymbol{\delta'} \cdot \boldsymbol{\delta_i}}{\boldsymbol{\delta_i} \cdot \boldsymbol{\delta_i}} \, \boldsymbol{\delta_i} \right) .$$

Then $\boldsymbol{\delta_{l+1}} \neq 0$, and

$$\boldsymbol{\delta_{l+1}} \cdot \boldsymbol{\delta_j} \;=\; \boldsymbol{\delta'} \cdot \boldsymbol{\delta_j} - \left( \sum_{i=1}^{l} \frac{\boldsymbol{\delta'} \cdot \boldsymbol{\delta_i}}{\boldsymbol{\delta_i} \cdot \boldsymbol{\delta_i}} \, \boldsymbol{\delta_i} \right) \cdot \boldsymbol{\delta_j} \;=\; \boldsymbol{\delta'} \cdot \boldsymbol{\delta_j} - \boldsymbol{\delta'} \cdot \boldsymbol{\delta_j} \;=\; 0$$

for $j = 1, \ldots, l$. Therefore $\boldsymbol{\delta_1}, \ldots, \boldsymbol{\delta_l}, \boldsymbol{\delta_{l+1}}$ is an orthogonal system. □

# Chapter 3

# Lexicographical Separation in $\mathbb{F}^n$

In this chapter, we shall present our main results of this thesis. Let $\mathbb{F}$ be an ordered field such that $\mathbb{Q} \subseteq \mathbb{F} \subseteq \mathbb{R}$. In Section 3.1 we introduce our main theorem, called the *lexicographical separation theorem* (Theorem 3.1.1), stating that a convex cone and its convex complement in $\mathbb{F}^n$ can be separated by linear functions and the lexicographic order. We also provide an equivalent version of the lexicographical separation theorem (Theorem 3.1.2) in terms of the lexicographically ordered polynomial ring which was introduced in the previous chapter. Further, we show that the lexicographical separation theorem has another interpretation that any compatible weak order on $\mathbb{F}^n$ can be embedded into the lexicographic order on $\mathbb{R}^n$ (Corollary 3.1.3). Lastly, we present a counterpart of the lexicographical separation theorem in Abelian group theory (Theorem 3.1.5). Applications of the lexicographical separation theorem will be discussed in the subsequent chapters. In Section 3.2 and Section 3.3 we give a proof of the lexicographical separation theorem by using the lexicographically ordered polynomial ring.

## 3.1 Lexicographical Separation Theorems

Throughout this chapter, $\mathbb{F}$ stand for an ordered field such that $\mathbb{Q} \subseteq \mathbb{F} \subseteq \mathbb{R}$. Also we denote by $F_n$ the $n$-dimensional vector space over $\mathbb{F}$ with the basis $\boldsymbol{e_1}, \ldots, \boldsymbol{e_n}$, i.e.

$$F_n \;=\; \{\, \lambda_1\boldsymbol{e_1} + \cdots + \lambda_n\boldsymbol{e_n} \mid \lambda_1, \ldots, \lambda_n \in \mathbb{F} \,\}.$$

This chapter will be devoted to prove the following theorem:

**Theorem 3.1.1 (Main Theorem)** *Let $F_n$ be the $n$-dimensional vector space over $\mathbb{F}$, and let $P$ be a nonempty subset of $F_n$. Suppose $P$ is a positive cone and $\bar{P}$ is a convex cone. Then, there exist real-valued linear functions $g_1, \ldots, g_n$ on $F_n$, where $g_1$ is not constantly zero, such that for all $x \in F_n$,*

$$x \in P \quad \text{iff} \quad (\, g_1(x), \ldots, g_n(x)\,) \;>_L\; (\,0, \ldots, 0\,)\,. \tag{3.1}$$

*Moreover, $g_1$ is unique up to a positive scalar multiple,*[1] *that is, if $g_1', \ldots, g_n'$ are other functions with the same property, then there is a real constant $a > 0$ such that*

$$g_1'(x) \;=\; a\, g_1(x) \quad \text{for all } x \in F_n\,. \tag{3.2}$$

□

[1] The other functions $g_2, \ldots, g_n$ are entirely indefinite in almost all cases.

In case $\mathbb{F} = \mathbb{R}$, equivalent versions of this theorem was proved by Hausner and Wendel [18], Klee [23], Martínez-Legaz and Singer [30].

Theorem 3.1.1 can also be stated in the following way:

**Theorem 3.1.2** *Let $F_n$ be the $n$-dimensional vector space over $\mathbb{F}$, and let $P$ be a nonempty subset of $F_n$. Suppose $P$ is a positive cone and $\bar{P}$ is a convex cone. Then there are lexicographically ordered polynomials[2] $q_1, \ldots, q_n \in \mathbb{R}[\varepsilon]$ such that for all $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$,*

$$\lambda_1 q_1 + \cdots + \lambda_n q_n \; > \; 0 \quad \text{iff} \quad \lambda_1 \boldsymbol{e_1} + \cdots + \lambda_n \boldsymbol{e_n} \in P. \tag{3.3}$$

*Further, $q_1, \ldots, q_n$ can be written as the following form:*

$$\begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{pmatrix} = \begin{pmatrix} r_{11} & \cdots & r_{1n} \\ r_{21} & \cdots & r_{2n} \\ \vdots & \ddots & \vdots \\ r_{n1} & \cdots & r_{nn} \end{pmatrix} \begin{pmatrix} 1 \\ \varepsilon \\ \vdots \\ \varepsilon^{n-1} \end{pmatrix} \tag{3.4}$$

*where $\{r_{ij}\}_{1 \le i,j \le n}$ is an $n \times n$ matrix of real numbers such that the first column vector is nonzero, i.e. $(r_{11}, \ldots, r_{n1})^T \neq (0, \ldots, 0)^T$.*
*Moreover, the first column vector is unique up to a positive scalar multiple, i.e. if $\{s_{ij}\}_{1 \le i,j \le n}$ is another matrix with the same property, then there exists a positive real number $a > 0$ such that*

$$\begin{pmatrix} s_{11} \\ \vdots \\ s_{n1} \end{pmatrix} = a \begin{pmatrix} r_{11} \\ \vdots \\ r_{n1} \end{pmatrix}.$$

□

That is to say, in any finite-dimensional vector space over $\mathbb{F}$, a positive cone and its convex complement can be separated by a set of lexicographically ordered polynomials. We will prove Theorem 3.1.2 in the next section. Here we briefly check the equivalence of Theorem 3.1.1 and Theorem 3.1.2:

Suppose there exist real-valued linear functions $g_1, \ldots, g_n$ on $F_n$, where $g_1$ is not constantly zero, such that (3.1) holds for all $x \in F_n$. Let the corresponding $n \times n$ real matrix $\{r_{ij}\}_{1 \le i,j \le n}$ be defined as $r_{ij} := g_j(\boldsymbol{e_i})$ for all $i, j$. Let $q_1, \ldots, q_n \in \mathbb{R}[\varepsilon]$ be defined by (3.4). Then one can easily verifies that $(r_{11}, \ldots, r_{n1})^T \neq (0, \ldots, 0)^T$, and for all $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$,

$$\begin{aligned} & \lambda_1 \boldsymbol{e_1} + \cdots + \lambda_n \boldsymbol{e_n} \in P \\ \iff \quad & (\lambda_1, \ldots, \lambda_n) \begin{pmatrix} r_{11} & \cdots & r_{1n} \\ r_{21} & \cdots & r_{2n} \\ \vdots & \ddots & \vdots \\ r_{n1} & \cdots & r_{nn} \end{pmatrix} >_L (0, \ldots, 0) \qquad \text{(by (3.1))} \\ \iff \quad & (\lambda_1, \ldots, \lambda_n) \begin{pmatrix} r_{11} & \cdots & r_{1n} \\ r_{21} & \cdots & r_{2n} \\ \vdots & \ddots & \vdots \\ r_{n1} & \cdots & r_{nn} \end{pmatrix} \begin{pmatrix} 1 \\ \varepsilon \\ \vdots \\ \varepsilon^{n-1} \end{pmatrix} > (0, \ldots, 0) \\ \iff \quad & \lambda_1 q_1 + \cdots + \lambda_n q_n > 0. \end{aligned}$$

[2]See Definition 2.1.5 for the definition of lexicographically ordered polynomials.

Conversely, suppose there exist $q_1, \ldots, q_n \in \mathbb{R}[\varepsilon]$ such that (3.3) holds for all $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$, where $q_1, \ldots, q_n$ are represented by an $n \times n$ real matrix $\{r_{ij}\}_{1 \le i,j \le n}$ of the form (3.4) such that $(r_{11}, \ldots, r_{n1})^T \neq (0, \ldots, 0)^T$. Let the corresponding linear functions $g_1, \ldots, g_n$ on $F_n$ be defined as

$$g_j(\lambda_1 \boldsymbol{e_1} + \cdots + \lambda_n \boldsymbol{e_n}) := \lambda_1 r_{1j} + \cdots + \lambda_n r_{nj} \qquad \text{for all } \lambda_1, \ldots, \lambda_n \in \mathbb{F}.$$

Then $g_1$ is not constantly zero, and also, by a similar argument to the above, (3.1) holds for all $x \in F_n$.

It is also easy to verify the correspondence of the uniqueness parts of Theorem 3.1.1 and Theorem 3.1.2.

We shall present a corollary of Theorem 3.1.1, which states that every compatible weak order on a finite-dimensional vector space over $\mathbb{F}$ can be represented by linear functions and the lexicographic order:

**Corollary 3.1.3** *Let $F_n$ be the $n$-dimensional vector space over $\mathbb{F}$, and let $\prec$ be a nontrivial weak order on $F_n$. Suppose $\prec$ is compatible with the vector space structure of $F_n$ (as in Definition 2.2.1). Then, there exist real-valued linear functions $g_1, \ldots, g_n$ on $F_n$, where $g_1$ is not constantly zero, such that for all $x, y \in F_n$,*

$$x \prec y \quad \text{iff} \quad (g_1(x), \ldots, g_n(x)) <_L (g_1(y), \ldots, g_n(y)). \tag{3.5}$$

*Moreover, $g_1$ is unique up to a positive scalar multiple, that is, if $g_1', \ldots, g_n'$ are other functions with the same property, then there is a real constant $a > 0$ such that (3.2) holds.*

**Proof.** Let $\prec$ be a nontrivial weak order on $F_n$, and suppose $\prec$ is compatible. Let $P = \{x \in F_n \mid x \succ \mathbf{0}\}$. Then, by Lemma 2.2.3, $P$ is a positive cone and $\bar{P}$ is a convex cone. Since $\prec$ is nontrivial, $P$ is nonempty. By Theorem 3.1.1, there exist real-valued linear functions $g_1, \ldots, g_n$ on $F_n$, where $g_1$ is not constantly zero, such that (3.1) holds for all $x \in F_n$. Therefore, for all $x, y \in F_n$,

$$\begin{aligned} y \succ x \iff y - x \succ \mathbf{0} &\iff (g_1(y-x), \ldots, g_n(y-x)) >_L (0, \ldots, 0) \quad \text{(by (3.1))} \\ &\iff (g_1(y), \ldots, g_n(y)) >_L (g_1(x), \ldots, g_n(x)). \end{aligned}$$

For the uniqueness, let $g_1, \ldots, g_n$ be real-valued linear functions on $F_n$, where $g_1'$ is not constantly zero, such that (3.5) holds for all $x, y \in F_n$. Then, $x \succ \mathbf{0}$ iff $(g_1(x), \ldots, g_n(x)) >_L (0, \ldots, 0)$; hence (3.1) holds for all $x \in F_n$. Therefore the uniqueness part of Corollary 3.1.3 follows from the corresponding one of Theorem 3.1.1. □

**Remark 3.1.4** Martínez-Legaz and Singer [32] discussed compatible weak orders on $\mathbb{R}^n$ in terms of the lexicographic order and linear operators. They proved the lexicographical representation theorem in their Theorem 1.1 from the viewpoint of lexicographical separation of two convex sets in $\mathbb{R}^n$, whose idea is similar to Corollary 3.1.3 in the present paper. In brief, Theorem 1.1 of Martínez-Legaz and Singer [32] can be presented as follows: For any compatible weak order $\prec$ on $\mathbb{R}^n$, there exist unique $r \in \{0, 1, \ldots, n\}$ and a linear operator $g : \mathbb{R}^n \to \mathbb{R}^r$ such that for all $\boldsymbol{y}, \boldsymbol{y'} \in \mathbb{R}^n$,

$$\boldsymbol{y} \prec \boldsymbol{y'} \quad \text{iff} \quad g(\boldsymbol{y}) <_L g(\boldsymbol{y'}).$$

Also they showed the uniqueness of the linear operator. □

Here we shall present another separation theorem for weakly–ordered Abelian groups, which can be derived as a consequence of Theorem 3.1.2. The first half of the following theorem is essentially equivalent to classical Hahn's Embedding Theorem (see Remark 3.1.6); but the latter (uniqueness) part doesn't follow from Hahn's Theorem. Let $\langle G_n, +, 0\rangle$ be the Abelian group generated by $\alpha_1, \ldots, \alpha_n$, i.e.

$$G_n \;=\; \{\, k_1\alpha_1 + \cdots + k_n\alpha_n \mid k_1, \ldots, k_n \in \mathbb{Z}\,\}$$

such that $k_1\alpha_1 + \cdots + k_n\alpha_n = 0$ implies $k_1 = \cdots = k_n = 0$. (In other words, let $\langle G_n, +, 0\rangle$ be the $\mathbb{Z}$-free module with the basis $\alpha_1, \ldots, \alpha_n$.)

**Theorem 3.1.5** *Let $\langle G_n, +, 0\rangle$ be the $\mathbb{Z}$-free module generated by $\alpha_1, \ldots, \alpha_n$, and let $P$, $I$ be nonempty subsets of $G_n$ satisfying C1–C3.*[3] *Then there are lexicographically ordered polynomials $q_1, \ldots, q_n \in \mathbb{R}[\varepsilon]$ such that for all $k_1, \ldots, k_n \in \mathbb{Z}$,*

$$k_1q_1 + \cdots + k_nq_n > 0 \quad \textit{iff} \quad k_1\alpha_1 + \cdots + k_n\alpha_n \in P \tag{3.6}$$

*where the polynomials $q_1, \ldots, q_n$ are represented by an $n \times n$ matrix of real numbers $\{r_{ij}\}_{1 \le i,j \le n}$ of the form (3.4) such that the first column vector is nonzero. Moreover, the first column vector is unique up to a positive scalar multiple.*

**Proof.** Let $Q_n$ be the vector space over $\mathbb{Q}$ with the basis $\boldsymbol{e_1}, \ldots, \boldsymbol{e_n}$, i.e.

$$Q_n = \{\, q_1\boldsymbol{e_1} + \cdots + q_n\boldsymbol{e_n} \mid q_1, \ldots, q_n \in \mathbb{Q}\,\}\,.$$

It is clear that $\langle G_n, +, 0\rangle$ is naturally embedded in $Q_n$, the generator $\alpha_i$ corresponding to the base $\boldsymbol{e_i}$ for $i = 1, \ldots, n$.

Suppose $P$, $I$ is a pair of nonempty subsets of $G_n$ satisfying C1–C3. Define a pair of subsets $P'$, $I'$ in $Q_n$ as follows: for all $\boldsymbol{x} \in Q_n$,

$$\boldsymbol{x} \in P' \quad \text{iff} \quad k\boldsymbol{x} \in P \;\text{ for some positive integer }\; k > 0\,, \tag{3.7}$$

$$\boldsymbol{x} \in I' \quad \text{iff} \quad k\boldsymbol{x} \in I \;\text{ for some positive integer }\; k > 0\,. \tag{3.8}$$

Clearly, these subsets $P'$, $I'$ of $Q_n$ are extensions of the original subsets $P$, $I$ of $G_n$, respectively. Since the subsets $P$, $I$ of $G_n$ satisfy C1, C3, and by Lemma 2.3.4, the following hold for all $\boldsymbol{x}$ in $Q_n$:

$$\begin{array}{l}\text{if } k\boldsymbol{x} \in P \text{ for some positive integer } k > 0 \\ \qquad \text{then } q\boldsymbol{x} \in P \text{ for all positive rational number } q > 0 \text{ such that } q\boldsymbol{x} \in G_n\,,\end{array} \tag{3.9}$$

$$\begin{array}{l}\text{if } k\boldsymbol{x} \in I \text{ for some positive integer } k > 0 \\ \qquad \text{then } q\boldsymbol{x} \in I \text{ for all rational number } q \text{ such that } q\boldsymbol{x} \in G_n\,.\end{array} \tag{3.10}$$

We shall show that the pair of subsets $P'$, $I'$ in $Q_n$ satisfies C1–C3 (where $Q_n$ is regarded as an additive group $\langle S, +, 0\rangle$); the verification of C1 is left to readers.

C2: For all $\boldsymbol{x} \in Q_n$, there exists a positive integer $k > 0$ such that $k\boldsymbol{x} \in G_n$. Since C3 holds for $P$, $I$ in $G_n$, we have exactly one of $k\boldsymbol{x} \in P$, $k\boldsymbol{x} \in I$, and $-k\boldsymbol{x} \in P$. Therefore $\boldsymbol{x} \in P'$ or $\boldsymbol{x} \in I'$ or $-\boldsymbol{x} \in P'$. By (3.9) and (3.10), these conditions are exclusive.

[3] Recall that, to any pair of subsets $P$, $I$ in $G_n$ satisfying C1–C3, there corresponds a compatible weak order on $G_n$. See Section 2.3.

C3: Suppose $\boldsymbol{x} \in P'$ and $\boldsymbol{y} \in P' \cup I'$. Then, by definition, there are positive integers $k$, $m$ such that $k\boldsymbol{x} \in P$ and $m\boldsymbol{y} \in P \cup I$. By (3.9) or (3.10), we can assume $m = k$. Since C2 holds for $P$, $I$ in $G_n$, $k\boldsymbol{x} \in P$ and $k\boldsymbol{y} \in P \cup I$ yields $k\boldsymbol{x} + k\boldsymbol{y} = k(\boldsymbol{x} + \boldsymbol{y}) \in P$. Thus $\boldsymbol{x} + \boldsymbol{y} \in P'$.

Now we are in a position to establish (3.6). Let $P$, $I$ be nonempty subsets of $G_n$ satisfying C1–C3. Then, as we have mentioned above, they are extended to the subsets $P'$, $I'$ of $Q_n$ satisfying C1–C3. By Lemma 2.3.5, $P'$ is a positive cone and $\bar{P'}$ is a convex cone in $Q_n$. Hence, by Theorem 3.1.2, we obtain a desired solution of polynomials $q_1, \ldots, q_n \in \mathbb{R}[\varepsilon]$ satisfying (3.3) for all $\lambda_1, \ldots, \lambda_n \in \mathbb{Q}$, which clearly satisfies (3.6) for all $k_1, \ldots, k_n \in \mathbb{Z}$.

For the uniqueness part, suppose there is a set of polynomials $q_1, \ldots, q_n$ satisfying (3.6) with respect to the subsets $P$, $I$ of $G_n$. We show that $q_1, \ldots, q_n$ also satisfy (3.3) with respect to the extension $P'$, $I'$ in $Q_n$: for all $\lambda_1, \ldots, \lambda_n \in \mathbb{Q}$,

$$\lambda_1 \boldsymbol{e_1} + \cdots + \lambda_n \boldsymbol{e_n} \in P'$$

$$\Longleftrightarrow \quad \begin{cases} (k\lambda_1)\,\alpha_1 + \cdots + (k\lambda_n)\,\alpha_n \in P \\ \text{for some positive integer } k > 0 \text{ such that } k\lambda_1, \ldots, k\lambda_n \in \mathbb{Z} \end{cases} \qquad \text{(by (3.7))}$$

$$\Longleftrightarrow \quad \begin{cases} (k\lambda_1)\,q_1 + \cdots + (k\lambda_n)\,q_n > 0 \\ \text{for some positive integer } k > 0 \text{ such that } k\lambda_1, \ldots, k\lambda_n \in \mathbb{Z} \end{cases} \qquad \text{(by (3.6))}$$

$$\Longleftrightarrow \quad \lambda_1 q_1 + \cdots + \lambda_n q_n > 0 .$$

Therefore the uniqueness part of Theorem 3.1.5 follows from the corresponding one of Theorem 3.1.2. □

**Remark 3.1.6** Theorem 3.1.5 (except for the uniqueness part) can be derived also as a consequence of Hahn's Embedding Theorem. A system $\langle S, \prec, +, 0\rangle$ is called an *ordered Abelian group* if it satisfies the following conditions:

O1 $\langle S, +, 0\rangle$ is an Abelian group,

O2 $\prec$ is a total order on $S$,

O3 for all $x, y, z \in S$, $x \prec y$ implies $x + z \prec y + z$.

Hahn [14] showed that any ordered Abelian group $\langle S, \prec, +, 0\rangle$ can be embedded into a lexicographically ordered vector space over the real numbers. ( For a proof see e.g. Fuchs [11]. ) Moreover, if the group is generated by $n$ generators, then the $n$-dimensional vector space $\mathbb{R}^n$ is sufficient for the embedding. ( For a proof see e.g. Teh [43]. ) The following is an outline of how to obtain the existence part of Theorem 3.1.5 from Hahn's Theorem:

Let $\langle G_n, \prec, +, 0\rangle$ be a weakly–ordered Abelian group generated by $\alpha_1, \ldots, \alpha_n$. It can easily be verified that the associated indifference $\sim$ satisfies the following property for all $x, x', y, y' \in G_n$:

$$\text{if } x \sim x' \text{ and } y \sim y', \quad \text{then } x + y \sim x' + y' \quad \text{and } x \precsim y \Leftrightarrow x' \precsim y' .$$

Let $I = \{ x \in G_n \mid 0 \sim x \}$. Then one can obtain the quotient $\langle G_n/I, \prec, +, 0 \rangle$. Let $\varphi$ be the natural projection:

$$\varphi \,:\, G_n \longrightarrow G_n/I\,.$$

$\langle G_n/I, \prec, +, 0\rangle$ is an ordered Abelian group with $l$ generators for some $l \leq n$. Let $\langle \mathbb{R}^l, <_L, +, 0\rangle$ be the $l$-dimensional lexicographically ordered vector space. By Hahn's Embedding Theorem, there is a mapping

$$\psi \,:\, G_n/I \longrightarrow \mathbb{R}^l$$

satisfying the following conditins for all $x, y \in G_n/I$:

$$x \prec y \quad \text{iff} \quad \psi(x) <_L \psi(x)\,; \qquad \psi(x+y) \;=\; \psi(x) + \psi(y)\,.$$

We shall show that

$$q_i \;:=\; \psi \circ \varphi(\alpha_i) \qquad \text{for } i = 1, \ldots, n$$

is a desired solution: For all $k_1, \ldots, k_n \in \mathbb{Z}$,

$$\begin{aligned} \psi \circ \varphi(k_1\alpha_1 + \cdots + k_n\alpha_n) &= k_1 \cdot (\psi \circ \varphi(\alpha_1)) + \cdots + k_n \cdot (\psi \circ \varphi(\alpha_n)) \\ &= k_1 q_1 + \cdots + k_n q_n\,, \end{aligned}$$

and hence

$$k_1\alpha_1 + \cdots + k_n\alpha_n \succ 0 \quad \text{iff} \quad k_1 q_1 + \cdots + k_n q_n >_L 0\,.$$

Thus, we obtain (3.6) of Theorem 3.1.5. Note that $q_i$ is a vector $(r_{i1}, \ldots, r_{il})$, and hence we need to translate it into a polynomial $r_{i1} + r_{i2}\,\varepsilon + \cdots + r_{il}\,\varepsilon^{l-1}$. □

## 3.2 Proof of Theorem 3.1.2 (for $n = 2$)

The proof of Theorem 3.1.2 will be given by a series of lemmas. We start with the case $n = 2$. Let $F_2$ be the 2-dimensional vector space over $\mathbb{F}$ with the basis $\boldsymbol{e_1}, \boldsymbol{e_2}$, i.e.

$$F_n = \{ \lambda_1 \boldsymbol{e_1} + \lambda_2 \boldsymbol{e_2} \mid \lambda_1, \lambda_2 \in \mathbb{F} \}\,.$$

Let $P$ be a nonempty subset of $F_2$ such that $P$ is a positive cone and $\bar{P}$ is a convex cone, and let $I$ be the associated indifference subspace.

We shall identify an element $\lambda_1 \boldsymbol{e_1} + \lambda_2 \boldsymbol{e_2} \in F_2$ with a point $(\lambda_1, \lambda_2)$ in 2-dimensional space $\mathbb{F}^2$. (Also we assume that $\mathbb{F}^2$ is naturally embedded in $\mathbb{R}^2$.) Accordingly, we shall use the notation "$(\lambda_1, \lambda_2) \in P$" instead of "$\lambda_1 \boldsymbol{e_1} + \lambda_2 \boldsymbol{e_2} \in P$," and so on.

Note that $P$, $I$, and $-P$ are mutually disjoint (by Lemma 2.2.4 (b)) and *anti-symmetric*: for all $x \in \mathbb{F}^2$,

$$x \in P \quad \text{iff} \quad -x \in -P\,, \qquad x \in I \quad \text{iff} \quad -x \in I\,.$$

Our goal is to show that they are divided by a line through the origin (see Figure 3.1). The following lemma is fundamental for our proof.

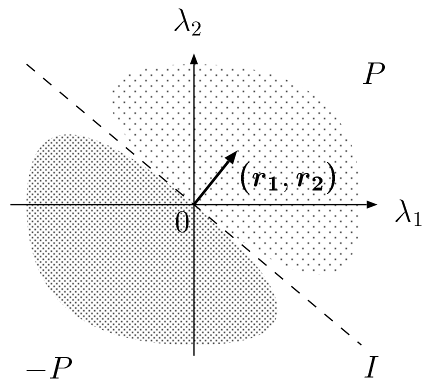


Figure 3.1: Assignment of $P$, $I$, and $-P$

**Lemma 3.2.1** Let $P$ be a nonempty subset of $F_2$. Suppose $P$ is a positive cone and $\bar{P}$ is a convex cone. Then there is a pair of real numbers $(r_1, r_2) \neq (0,0)$ such that for all $\lambda_1, \lambda_2 \in \mathbb{F}$,

$$\begin{array}{lll} \lambda_1 r_1 + \lambda_2 r_2 > 0 & \text{implies} & \lambda_1 \boldsymbol{e_1} + \lambda_2 \boldsymbol{e_2} \in P\,, \\ \lambda_1 r_1 + \lambda_2 r_2 < 0 & \text{implies} & \lambda_1 \boldsymbol{e_1} + \lambda_2 \boldsymbol{e_2} \in -P\,. \end{array} \tag{3.11}$$

Moreover, the pair of real numbers $(r_1, r_2)$ satisfying (3.11) is unique up to a positive scalar multiple.

**Proof.** It is enough to consider the case $\boldsymbol{e_1}, \boldsymbol{e_2} \in P \cup I$: for, if the other cases hold, e.g. $\boldsymbol{e_1} \in P \cup I$ and $\boldsymbol{e_2} \notin P \cup I$, then, by Lemma 2.2.4 (b), $\boldsymbol{e_1}, -\boldsymbol{e_2} \in P \cup I$; hence the following arguments can be applied by changing the bases $\boldsymbol{e_1}, \boldsymbol{e_2}$ into $\boldsymbol{e_1}, -\boldsymbol{e_2}$.

Assume $(1,0), (0,1) \in P \cup I$ (that is $\boldsymbol{e_1}, \boldsymbol{e_2} \in P \cup I$). Then, at least one of $(1,0)$ or $(0,1)$ is in $P$: for, if $(1,0), (0,1) \in I$ then $(\lambda_1, \lambda_2) \in I$ for all $\lambda_1, \lambda_2 \in \mathbb{F}$ by Lemma 2.2.4 (a), contradicting the assumption $P \neq \emptyset$.

We claim that all the points in the first quadrant $\{ (\lambda_1, \lambda_2) \in \mathbb{F}^2 \mid \lambda_1, \lambda_2 > 0 \}$ belong to $P$. This is trivial for the case $(1,0), (0,1) \in P$. If $(1,0) \in I$ (hence $(0,1) \in P$), we have $(\lambda_1, 0) \in I$ for all $\lambda_1 > 0$ by Lemma 2.2.4 (a), and $(0, \lambda_2) \in P$ for all $\lambda_2 > 0$; therefore $(\lambda_1, \lambda_2) \in P$ for all $\lambda_1, \lambda_2 > 0$ by Lemma 2.2.4 (c). The same argument applies to the case $(0,1) \in I$. Accordingly, all the points in the third quadrant $\{ (\lambda_1, \lambda_2) \in \mathbb{F}^2 \mid \lambda_1, \lambda_2 < 0 \}$ belong to $-P$. (See the left-hand side of Figure 3.2.)

We define subsets $\Gamma$, $\Delta$ of $\mathbb{F}$ as follows:

$$\Gamma = \left\{ \frac{\lambda_2}{\lambda_1} \in \mathbb{F} \;\middle|\; \lambda_1 > 0,\ (\lambda_1, \lambda_2) \in -P \cup I \right\}, \quad \Delta = \left\{ \frac{\lambda_2}{\lambda_1} \in \mathbb{F} \;\middle|\; \lambda_1 > 0,\ (\lambda_1, \lambda_2) \in P \right\}.$$

(Namely, a number $\lambda_2/\lambda_1 \in \mathbb{F}$ in these sets is the gradient of a line between the origin $(0,0)$ and a point $(\lambda_1, \lambda_2)$ on the right half-plane of $\mathbb{F}^2$.) Note that $\Gamma \cup \Delta = \mathbb{F}$. It can easily be seen that

(a) if $\Gamma = \emptyset$, then all the points on the right half-plane $\{ (\lambda_1, \lambda_2) \in \mathbb{F}^2 \mid \lambda_1 > 0 \}$ belong to $P$,

(b) if $\Gamma \neq \emptyset$, then $\Gamma$ has the least upper bound $\hat{p}$ in $\mathbb{R}$. (Recall that $\mathbb{Q} \subseteq \mathbb{F} \subseteq \mathbb{R}$.)

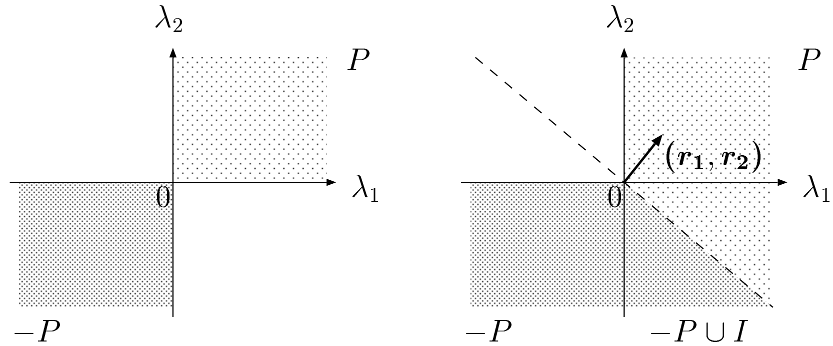


Figure 3.2: Assignment of $P$, $I$, and $-P$

We shall show that, in the case (b), the least upper bound $\hat{p}$ of $\Gamma$ is also the greatest lower bound of $\Delta$ in $\mathbb{R}$. First we show that $\hat{p}$ is a lower bound of $\Delta$. Suppose otherwise: there exists $(\eta_1, \eta_2) \in P$ such that $\eta_1 > 0$ and $\hat{p} > \eta_2/\eta_1$. Since $\hat{p}$ is the least upper bound of $\Gamma$, there exists $(\theta_1, \theta_2) \in -P \cup I$ such that $\theta_1 > 0$ and $\hat{p} > \theta_2/\theta_1 > \eta_2/\eta_1$. Then $\theta_2\eta_1 - \theta_1\eta_2 > 0$, and hence, from the hypothesis $(0,1) \in P \cup I$, we have $(0, \theta_2\eta_1 - \theta_1\eta_2) \in P \cup I$. But this leads to $(\theta_1, \theta_2) \in P$ as follows, contradicting the assumption $(\theta_1, \theta_2) \in -P \cup I$.

$$
\begin{aligned}
(\eta_1, \eta_2) \in P \implies & \theta_1(\eta_1, \eta_2) = (\theta_1\eta_1, \theta_1\eta_2) \in P \\
\implies & (\theta_1\eta_1, \theta_1\eta_2) + (0, \theta_2\eta_1 - \theta_1\eta_2) \in P \quad \text{(by Lemma 2.2.4 (c))} \\
\implies & (\theta_1\eta_1, \theta_2\eta_1) = \eta_1(\theta_1, \theta_2) \in P \\
\implies & (\theta_1, \theta_2) \in P .
\end{aligned}
$$

Let $p^*$ be the greatest lower bound of $\Delta$. It remains to show that $\hat{p} = p^*$. Suppose $\hat{p} < p^*$. Then, by the density of $\mathbb{F}$ in $\mathbb{R}$, there is an $r \in \mathbb{F}$ such that $\hat{p} < r < p^*$. Hence $r \notin \Gamma \cup \Delta$, which contradicts $\Gamma \cup \Delta = \mathbb{F}$.

Thus, in the case (b), we get a real number $\hat{p}$ such that for all $\lambda \in \mathbb{F}$,

$$
\begin{aligned}
\lambda \in \Gamma \quad & \text{implies} \quad \lambda \leq \hat{p}, \\
\lambda \in \Delta \quad & \text{implies} \quad \lambda \geq \hat{p}.
\end{aligned} \tag{3.12}
$$

It is easy to verify that such $\hat{p}$ satisfying (3.12) is unique in $\mathbb{R}$. Note that $\hat{p}$ is nonpositive, i.e. $\hat{p} \leq 0$, since $\Delta$ includes $1/k$ for all positive integer $k > 0$, and hence $\hat{p} \leq 1/k$ for all positive integer $k > 0$. Now, let $r_1, r_2$ be a pair of real numbers satisfying

$$
-r_1/r_2 \;=\; \hat{p}, \quad r_1 \geq 0, \quad r_2 > 0 . \tag{3.13}
$$

That is to say, $(r_1, r_2)$ is a *normal vector* of the line whose gradient is $\hat{p}$. Since $\hat{p}$ is uniquely determined, the vector $(r_1, r_2)$ is uniquely determined up to a positive scalar multiple (see the right-hand side of Figure 3.2). We shall show that for all $\lambda_1, \lambda_2 \in \mathbb{F}$,

$$
(\lambda_1, \lambda_2) \in I \quad \text{implies} \quad \lambda_1 r_1 + \lambda_2 r_2 = 0 . \tag{3.14}
$$

That is to say, the set $I$ must be contained on the line whose gradient is $\hat{p}$. Suppose otherwise: there exists $(\theta_1, \theta_2) \in I$ such that $\theta_1 r_1 + \theta_2 r_2 \neq 0$. By the symmetry of

$I$ with respect to the origin, we can assume $\theta_1 \geq 0$. If $\theta_1 > 0$, then $\theta_2/\theta_1 \in \Gamma$, which implies $\theta_2/\theta_1 \leq p$ by (3.12), and therefore $\theta_1 r_1 + \theta_2 r_2 \leq 0$. Since we assumed $\theta_1 r_1 + \theta_2 r_2 \neq 0$, we have $\theta_1 r_1 + \theta_2 r_2 < 0$. If $\theta_1 = 0$, we can assume $\theta_2 < 0$ by the symmetry of $I$ with respect to the origin, so that $\theta_1 r_1 + \theta_2 r_2 < 0$. Thus, in both cases, we have $\theta_1 r_1 + \theta_2 r_2 < 0$. Let $e := -\theta_1 r_1 - \theta_2 r_2 > 0$. Then, by the density of $\mathbb{F}$ in $\mathbb{R}$, there exists $\eta \in \mathbb{F}$ such that $0 < \eta - p < e/2r_2$. This means $\eta < p$ and hence $(1, \eta) \in P$. By the hypothesis $\theta_1 \geq 0$ we have $\theta_1 + 1 > 0$. Therefore,

$$\begin{aligned}
\frac{\theta_2 + \eta}{\theta_1 + 1} - p &= \frac{\theta_2 + \eta}{\theta_1 + 1} + \frac{r_1}{r_2} \\
&= \frac{(\theta_1 r_1 + \theta_2 r_2) + (r_1 + \eta r_2)}{(\theta_1 + 1) r_2} \\
&= \frac{(-e) + (\eta - p) r_2}{(\theta_1 + 1) r_2} \\
&< \frac{(-e + e/2)}{(\theta_1 + 1) r_2} < 0.
\end{aligned}$$

This implies $(\theta_1+1, \theta_2+\eta) \in -P \cup I$. But we also have $(\theta_1+1, \theta_2+\eta) = (\theta_1, \theta_2) + (1, \eta) \in P$ by Lemma 2.2.4 (c), a contradiction.

Now we are in a position to establish (3.11): there exists a nonzero real vector $(r_1, r_2)$ such that for all $\lambda_1, \lambda_2 \in \mathbb{F}$,

$$\lambda_1 r_1 + \lambda_2 r_2 > 0 \quad \text{implies} \quad (\lambda_1, \lambda_2) \in P, \tag{3.15}$$

$$\lambda_1 r_1 + \lambda_2 r_2 < 0 \quad \text{implies} \quad (\lambda_1, \lambda_2) \in -P, \tag{3.16}$$

and also $(r_1, r_2)$ is uniquely determined up to a positive scalar multiple.

In the case (a), all the points on the left half-plane $\{ (\lambda_1, \lambda_2) \in \mathbb{F}^2 \mid \lambda_1 < 0 \}$ belong to $-P$. Therefore, any vector $(r_1, r_2) = (s, 0)$ with $s > 0$ is a solution of (3.15) and (3.16). One can easily verifies that $(r_1, r_2)$ is uniquely determined up to a positive scalar multiple.

In the case (b), let $(r_1, r_2)$ be the vector which was defined by (3.13). We shall show that this vector $(r_1, r_2)$ satisfies (3.15) and (3.16) for all $\lambda_1, \lambda_2 \in \mathbb{F}$: It is enough to verify (3.15), since (3.16) can be treated in a similar way. Suppose $\lambda_1, \lambda_2 \in \mathbb{F}$ satisfy $\lambda_1 r_1 + \lambda_2 r_2 > 0$. We shall divide the case according to the sign of $\lambda_1$.

($\lambda_1 > 0$) $\lambda_1 r_1 + \lambda_2 r_2 > 0$ implies $r_1/r_2 + \lambda_2/\lambda_1 = \lambda_2/\lambda_1 - \hat{p} > 0$. Hence $\lambda_2/\lambda_1 \in \Delta$ by (3.12). Therefore $(\lambda_1, \lambda_2) \in P$.

($\lambda_1 < 0$) $\lambda_1 r_1 + \lambda_2 r_2 > 0$ implies $r_1/r_2 + \lambda_2/\lambda_1 = \lambda_2/\lambda_1 - \hat{p} = (-\lambda_2)/(-\lambda_1) - \hat{p} < 0$. Hence $(-\lambda_2)/(-\lambda_1) \in \Gamma$ by (3.12), and therefore $(-\lambda_1, -\lambda_2) \in -P \cup I$. Suppose $(-\lambda_1, -\lambda_2) \in I$. Then, by (3.14), this yields $\lambda_1 r_1 + \lambda_2 r_2 = 0$, contradicting $\lambda_1 r_1 + \lambda_2 r_2 > 0$. Hence we must have $(-\lambda_1, -\lambda_2) \in -P$, which means $(\lambda_1, \lambda_2) \in P$.

($\lambda_1 = 0$) $\lambda_1 r_1 + \lambda_2 r_2 > 0$ means $\lambda_2 > 0$. Assume $(\lambda_1, \lambda_2) = (0, \lambda_2) \in -P \cup I$ (which will lead to a contradiction). Let $d$ be a sufficiently large integer such that $-d\lambda_2 < \hat{p}$. Then $(1, -d\lambda_2) \in -P \cup I$: for, if $(1, -d\lambda_2) \in P$ then $-d\lambda_2 \in \Delta$ and hence $-d\lambda_2 \geq \hat{p}$ by (3.12), a contradiction. Thus $(-1, d\lambda_2) \in P \cup I$. Now, the assumption $(0, \lambda_2) \in -P \cup I$ means $(0, -\lambda_2) \in P \cup I$, which implies $(0, -2d\lambda_2) \in P \cup I$. This and $(-1, d\lambda_2) \in P \cup I$ yield $(-1, -d\lambda_2) \in P \cup I$ by Lemma 2.2.4. Hence

$(1, d\lambda_2) \in -P \cup I$. However, we have $(1, d\lambda_2) \in P$ because $P$ includes all the points in the first quadrant, a contradiction. Thus we must have $(\lambda_1, \lambda_2) \in P$.

To show that $(r_1, r_2)$ is unique up to a positive scalar multiple, suppose there is another vector $(r_1', r_2') \neq (0,0)$ satisfying (3.15) and (3.16) for all $\lambda_1, \lambda_2 \in \mathbb{F}$. First we show that $r_2' > 0$. Assume $r_2' \leq 0$. If $r_2' = 0$, then obviously $r_1' > 0$; hence, by (3.15), $\lambda_1 > 0$ implies $(\lambda_1, \lambda_2) \in P$ for all $\lambda_1, \lambda_2 \in \mathbb{F}$. This means $\Gamma = \emptyset$, a contradiction. If $r_2' < 0$, then, by (3.16), there is a sufficiently large $\lambda_2 \in \mathbb{F}$ such that $(1, \lambda_2) \in -P$, which contradicts that all the points in the first quadrant belong to $P$. Thus we have shown $r_2' > 0$. Let $p := -r_1'/r_2'$. By (3.15) and (3.16), for all $\lambda_1, \lambda_2 \in \mathbb{F}$ with $\lambda_1 > 0$,

$$\lambda_2/\lambda_1 > p \text{ implies } (\lambda_1, \lambda_2) \in P; \qquad \lambda_2/\lambda_1 < p \text{ implies } (\lambda_1, \lambda_2) \in -P.$$

Let us consider their contrapositions: for all $\lambda_1, \lambda_2 \in \mathbb{F}$ with $\lambda_1 > 0$,

$$(\lambda_1, \lambda_2) \in -P \cup I \text{ implies } \lambda_2/\lambda_1 \leq p; \qquad (\lambda_1, \lambda_2) \in P \cup I \text{ implies } \lambda_2/\lambda_1 \geq p.$$

Thus $p$ satisfies (3.12) for all $\lambda \in \mathbb{F}$; therefore $p = \hat{p}$. That is, $r_1'/r_2' = r_1/r_2$, and hence there is a positive real number $a > 0$ such that $(r_1', r_2') = a(r_1, r_2)$. □

Let us continue the proof of Theorem 3.1.2 to find a pair of polynomials $q_1, q_2 \in \mathbb{R}[\varepsilon]_2$ satisfying (3.3) for all $\lambda_1, \lambda_2 \in \mathbb{F}$. Suppose we obtained a nonzero real vector $(r_1, r_2)$ satisfying (3.11) in Lemma 3.2.1. We consider the following two different cases:

*Case 1. There is no* $(\lambda_1, \lambda_2) \in P$ *such that* $\lambda_1 r_1 + \lambda_2 r_2 = 0$.

In this case, there is no $(\lambda_1, \lambda_2) \in -P$ such that $\lambda_1 r_1 + \lambda_2 r_2 = 0$. Hence, for all $\lambda_1, \lambda_2 \in \mathbb{F}$, $\lambda_1 r_1 + \lambda_2 r_2 = 0$ implies $(\lambda_1, \lambda_2) \in I$. Therefore, along with (3.11) in Lemma 3.2.1, the following holds for all $(\lambda_1, \lambda_2) \in \mathbb{F}^2$:

$$\begin{aligned} \lambda_1 r_1 + \lambda_2 r_2 > 0 &\quad \text{implies} \quad (\lambda_1, \lambda_2) \in P, \\ \lambda_1 r_1 + \lambda_2 r_2 = 0 &\quad \text{implies} \quad (\lambda_1, \lambda_2) \in I, \\ \lambda_1 r_1 + \lambda_2 r_2 < 0 &\quad \text{implies} \quad (\lambda_1, \lambda_2) \in -P. \end{aligned}$$

Then, one can easily verifies that these implications are actually equivalences. Thus $(q_1, q_2) = (r_1, r_2)$ is a solution of (3.3).

*Case 2. There is* $(\theta_1, \theta_2) \in P$ *such that* $\theta_1 r_1 + \theta_2 r_2 = 0$.

We divide this case according to the sign of $\theta_1 r_2 - \theta_2 r_1$.

We cannot have $\theta_1 r_2 - \theta_2 r_1 = 0$: for, if $\theta_1 r_2 - \theta_2 r_1 = 0$ then, with the hypothesis $\theta_1 r_1 + \theta_2 r_2 = 0$, one can easily obtain $(\theta_1, \theta_2) = (0,0)$, which contradicts $(\theta_1, \theta_2) \in P$.

Suppose $\theta_1 r_2 - \theta_2 r_1 > 0$. We shall show that for all $\lambda_1, \lambda_2 \in \mathbb{F}$,

$$\text{“} \lambda_1 r_1 + \lambda_2 r_2 = 0 \text{ and } \lambda_1 r_2 - \lambda_2 r_1 > 0 \text{”} \quad \text{implies} \quad (\lambda_1, \lambda_2) \in P. \tag{3.17}$$

If $\lambda_1 r_1 + \lambda_2 r_2 = 0$ then, with the hypothesis $\theta_1 r_1 + \theta_2 r_2 = 0$, we have $(\lambda_1, \lambda_2) = \rho(\theta_1, \theta_2)$ for some $\rho \in \mathbb{F}$. If we have also $\lambda_1 r_2 - \lambda_2 r_1 > 0$, then $\rho$ is positive, i.e. $\rho > 0$: because $\lambda_1 r_2 - \lambda_2 r_1 = (\rho\theta_1) r_2 - (\rho\theta_2) r_1 = \rho(\theta_1 r_2 - \theta_2 r_1) > 0$, which implies $\rho > 0$ by the hypothesis $\theta_1 r_2 - \theta_2 r_1 > 0$. Thus we have shown that, if both $\lambda_1 r_1 + \lambda_2 r_2 = 0$ and $\lambda_1 r_2 - \lambda_2 r_1 > 0$ are satisfied, $(\lambda_1, \lambda_2) = \rho(\theta_1, \theta_2)$ for some $\rho > 0$. Therefore the hypothesis $(\theta_1, \theta_2) \in P$ yields $(\lambda_1, \lambda_2) \in P$.

We claim that $(q_1, q_2) = (r_1, r_2) + \varepsilon\,(r_2, -r_1)$ is a solution of (3.3): for all $\lambda_1, \lambda_2 \in \mathbb{F}$,

$$
\begin{aligned}
\lambda_1 q_1 + \lambda_2 q_2 > 0 \;&\Longrightarrow\; (\lambda_1 r_1 + \lambda_2 r_2) + \varepsilon\,(\lambda_1 r_2 - \lambda_2 r_1) > 0 \\
&\Longrightarrow\; \lambda_1 r_1 + \lambda_2 r_2 > 0 \quad \text{or} \\
&\qquad \text{“ } \lambda_1 r_1 + \lambda_2 r_2 = 0 \;\text{ and }\; \lambda_1 r_2 - \lambda_2 r_1 > 0 \text{ ”} \\
&\Longrightarrow\; (\lambda_1, \lambda_2) \in P \qquad\qquad \text{(by (3.11) or (3.17))},
\end{aligned}
$$

and similarly

$$
\begin{aligned}
\lambda_1 q_1 + \lambda_2 q_2 < 0 \quad &\text{implies} \quad (\lambda_1, \lambda_2) \in -P\,, \\
\lambda_1 q_1 + \lambda_2 q_2 = 0 \quad &\text{implies} \quad (\lambda_1, \lambda_2) = (0, 0) \in I\,.
\end{aligned}
$$

These implications are actually equivalences.

In case $\theta_1 r_2 - \theta_2 r_1 < 0$, a similar argument shows $(q_1, q_2) = (r_1, r_2) + \varepsilon\,(-r_2, r_1)$ is a solution of (3.3).

Thus, in both Case 1 and Case 2, the pair of polynomials $(q_1, q_2)$ is written in the matrix form (3.4), in which the first column vector $(r_1, r_2)^T$ is nonzero.

For the uniqueness part of Theorem 3.1.2, suppose there is a pair of polynomials $q_1', q_2' \in \mathbb{R}[\varepsilon]$ satisfying (3.3) for all $\lambda_1, \lambda_2 \in \mathbb{F}$. Suppose also the pair of polynomials $q_1', q_2'$ is represented by a $2 \times 2$ real matrix $\{r_{ij}'\}_{1 \le i,j \le 2}$ of the form (3.4) such that the first column vector $(r_{11}', r_{21}')^T$ is nonzero. Then the following holds for all $\lambda_1, \lambda_2 \in \mathbb{F}$:

$$
\begin{aligned}
&\text{if } \lambda_1 r_{11}' + \lambda_2 r_{21}' > 0 \quad \text{then} \quad \lambda_1 q_1' + \lambda_2 q_2' \;=\; (\lambda_1 r_{11}' + \lambda_2 r_{21}') + \varepsilon\,(\lambda_1 r_{12}' + \lambda_2 r_{22}') \;>\; 0 \\
&\qquad\qquad\qquad\qquad\qquad\;\; \text{hence} \quad \lambda_1 \boldsymbol{e_1} + \lambda_2 \boldsymbol{e_2} \in P\,, \qquad\qquad \text{(by (3.3))} \\
&\text{if } \lambda_1 r_{11}' + \lambda_2 r_{21}' < 0 \quad \text{then} \quad \lambda_1 q_1' + \lambda_2 q_2' < 0 \quad \text{hence} \quad \lambda_1 \boldsymbol{e_1} + \lambda_2 \boldsymbol{e_2} \in -P\,.
\end{aligned}
$$

By Lemma 3.2.1, the vector $(r_{11}', r_{21}')^T$ is unique up to a positive scalar multiple. **Q.E.D.**

## 3.3 Proof of Theorem 3.1.2 (for general $n$)

First we shall prove a generalization of Lemma 3.2.1:

**Lemma 3.3.1** Let $F_n$ be the $n$-dimensional vector space over $\mathbb{F}$ with the basis $\boldsymbol{e_1}, \ldots, \boldsymbol{e_n}$ for some $n \ge 2$, and let $P$ be a nonempty subset of $F_n$. Suppose $P$ is a positive cone and $\bar{P}$ is a convex cone. Then there is a tuple of real numbers $(r_1, \ldots, r_n) \ne (0, \ldots, 0)$ such that for all $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$,

$$
\begin{aligned}
\lambda_1 r_1 + \cdots + \lambda_n r_n > 0 \quad &\text{implies} \quad \lambda_1 \boldsymbol{e_1} + \cdots + \lambda_n \boldsymbol{e_n} \in P\,, \\
\lambda_1 r_1 + \cdots + \lambda_n r_n < 0 \quad &\text{implies} \quad \lambda_1 \boldsymbol{e_1} + \cdots + \lambda_n \boldsymbol{e_n} \in -P\,.
\end{aligned}
\tag{3.18}
$$

Moreover, the tuple of real numbers $(r_1, \ldots, r_n)$ satisfying (3.18) is unique up to a positive scalar multiple.

**Proof.** We have already treated the case $n = 2$ in Lemma 3.2.1. Hence it is enough to verify the remaining cases $n > 2$. For this, we will use the result of Lemma 3.2.1.

Let $F_n$ be the $n$-dimensional vector space over $\mathbb{F}$ for some $n > 2$, and let $P$ be a nonempty subsets of $F_n$ such that $P$ is a positive cone and $\bar{P}$ is a convex cone. We shall denote by $F^{(ij)}$ the 2-dimensional subspace of $F_n$ having the basis $\boldsymbol{e_i}$, $\boldsymbol{e_j}$ $(1 \le i,j \le n$, $i \ne j)$, i.e.

$$F^{(ij)} \;=\; \{\, \lambda_i \boldsymbol{e_i} + \lambda_j \boldsymbol{e_j} \mid \lambda_i, \lambda_j \in \mathbb{F} \,\}\,.$$

One can easily verifies that the restriction of $P$ in the subspace $F^{(ij)}$ is a positive cone; also the restriction of $\bar{P}$ is a convex cone in $F^{(ij)}$. Hence, by the result of Lemma 3.2.1, there is a pair of real numbers $(r_i^{(ij)}, r_j^{(ij)})$ such that for all $\lambda_i, \lambda_j \in \mathbb{F}$,

$$\begin{array}{lll} \lambda_i r_i^{(ij)} + \lambda_j r_j^{(ij)} > 0 & \text{implies} & \lambda_i \boldsymbol{e_i} + \lambda_j \boldsymbol{e_j} \in P\,, \\ \lambda_i r_i^{(ij)} + \lambda_j r_j^{(ij)} < 0 & \text{implies} & \lambda_i \boldsymbol{e_i} + \lambda_j \boldsymbol{e_j} \in -P\,. \end{array} \tag{3.19}$$

( $P$ may be empty in the subspace $F^{(ij)}$; but, in such a case, we put $(r_i^{(ij)}, r_j^{(ij)}) = (0,0)$, which clearly satisfies (3.19) for all $\lambda_i, \lambda_j \in \mathbb{F}$.) By Lemma 3.2.1, the pair of real numbers $(r_i^{(ij)}, r_j^{(ij)})$ satisfying (3.19) is unique up to a positive scalar multiple.

That is to say, for each pair $i$, $j$ $(1 \le i,j \le n,\ i \ne j)$, there is a real vector $(r_i^{(ij)}, r_j^{(ij)})$ satisfying (3.19) for all $\lambda_i, \lambda_j \in \mathbb{F}$. We shall show that there exists $i$ such that $r_i^{(ij)} \ne 0$ for all $j\,(\ne i)$. Suppose otherwise: for all $i$ there exists $\varphi(i)$ such that $(r_i^{(i\varphi(i))}, r_{\varphi(i)}^{(i\varphi(i))}) = (0, s_{\varphi(i)})$; we can choose $\varphi(i)$ to satisfy $s_{\varphi(i)} \ne 0$, because if $(r_i^{(ij)}, r_j^{(ij)}) = (0,0)$ for all $j$ then this means $P$ is empty in $F^{(ij)}$ for all $j$, which yields $I = F_n$ by Lemma 2.2.4 (a) and thus contradicts $P \ne \emptyset$. Now, let us consider the following sequence:

$$1, \quad \varphi(1), \quad \varphi(\varphi(1)), \quad \varphi(\varphi(\varphi(1))), \quad \cdots$$

Since the dimension $n$ is finite, there exists $j$ such that $j = \overbrace{\varphi(\varphi(\cdots(\varphi}^{k\text{ times}}(\,j\,))\cdots))$ for some $k\,(\le n)$. For simplicity, we shall assume

$$\varphi(1) = 2, \quad \varphi(2) = 3, \quad \ldots \quad \varphi(k-1) = k, \quad \varphi(k) = 1\,.$$

Let $d_l := 1$ if $s_l > 0$, and $d_l := -1$ if $s_l < 0$ for $1 \le l \le k$. Then, by (3.19),

$$\left.\begin{array}{rclll} -d_1 r_1^{(12)} + d_2 r_2^{(12)} &=& -d_1 \cdot 0 + d_2 s_2 > 0 & \text{implies} & -d_1 \boldsymbol{e_1} + d_2 \boldsymbol{e_2} \in P \\ -d_2 r_2^{(23)} + d_3 r_3^{(23)} &=& -d_2 \cdot 0 + d_3 s_3 > 0 & \text{implies} & -d_2 \boldsymbol{e_2} + d_3 \boldsymbol{e_3} \in P \\ & & & \cdots & \\ -d_k r_k^{(k1)} + d_1 r_1^{(k1)} &=& -d_k \cdot 0 + d_1 s_1 > 0 & \text{implies} & -d_k \boldsymbol{e_k} + d_1 \boldsymbol{e_1} \in P \end{array}\right\} \Rightarrow \ \mathbf{0} \in P\,,$$

a contradiction. Thus, there is $i$ such that $r_i^{(ij)} \ne 0$ for all $j\,(\ne i)$. We can show, moreover, that either $r_i^{(ij)} > 0$ for all $j\,(\ne i)$ or $r_i^{(ij)} < 0$ for all $j\,(\ne i)$ : for, if we have $r_i^{(ik)} > 0$ and $r_i^{(il)} < 0$ for some $k$, $l$, then

$$\left.\begin{array}{lll} 1 \cdot r_i^{(ik)} + 0 \cdot r_k^{(ik)} > 0 & \text{implies} & 1 \cdot \boldsymbol{e_i} + 0 \cdot \boldsymbol{e_k} \in P \\ -1 \cdot r_i^{(il)} + 0 \cdot r_l^{(il)} > 0 & \text{implies} & -1 \cdot \boldsymbol{e_i} + 0 \cdot \boldsymbol{e_l} \in P \end{array}\right\} \Rightarrow \ \mathbf{0} \in P\,,$$

a contradiction.

Without loss of generality, we shall assume $r_1^{(1j)} > 0$ for all $j \ne 1$. Note that this means $\boldsymbol{e_1} \in P$ by (3.19). Further, by multiplying $1/r_1^{(1j)}$ if necessary, we can write $(r_1^{(1j)}, r_j^{(1j)}) = (1, r_j)$ for $j = 2, \ldots, n$.

We claim that the tuple of real numbers $(1, r_2, \dots, r_n)$ is a solution of (3.18): Suppose $\lambda_1 + \lambda_2 r_2 + \cdots + \lambda_n r_n > 0$ ($\lambda_1, \dots, \lambda_n \in \mathbb{F}$). We shall show $\lambda_1 \boldsymbol{e_1} + \lambda_2 \boldsymbol{e_2} + \cdots + \lambda_n \boldsymbol{e_n} \in P$. To simplify the situation, suppose $\lambda_2, \dots, \lambda_k \geq 0$ and $\lambda_{k+1}, \dots, \lambda_n \leq 0$ for some $k$ ($\leq n$). Then, by the density of $\mathbb{F}$ in $\mathbb{R}$, there are numbers $\theta_2, \dots, \theta_n \in \mathbb{F}$ such that

$$\theta_2 < r_2, \quad \dots \quad \theta_k < r_k, \quad \theta_{k+1} > r_{k+1}, \quad \dots \quad \theta_n > r_n \quad \text{and}$$
$$\lambda_1 + \lambda_2\theta_2 + \cdots + \lambda_n\theta_n > 0\,.$$

The last condition implies

$$(\lambda_1 + \lambda_2\theta_2 + \cdots + \lambda_n\theta_n)\, \boldsymbol{e_1} \in P \tag{3.20}$$

from the hypothesis $\boldsymbol{e_1} \in P$. The rest of the conditions imply

$$\begin{cases} -\theta_2 \cdot 1 + 1 \cdot r_2 > 0\,, \quad \cdots, \quad -\theta_k \cdot 1 + 1 \cdot r_k > 0\,, \\ \qquad -\theta_{k+1} \cdot 1 + 1 \cdot r_{k+1} < 0\,, \quad \cdots, \quad -\theta_n \cdot 1 + 1 \cdot r_n < 0\,, \end{cases}$$

$$\Longrightarrow \begin{cases} -\theta_2 \boldsymbol{e_1} + 1\boldsymbol{e_2} \in P\,, \quad \cdots, \quad -\theta_k \boldsymbol{e_1} + 1\boldsymbol{e_k} \in P\,, \\ \qquad -\theta_{k+1}\boldsymbol{e_1} + 1\boldsymbol{e_{k+1}} \in -P\,, \quad \cdots, \quad -\theta_n \boldsymbol{e_1} + 1\boldsymbol{e_n} \in -P\,, \end{cases} \qquad \text{(by (3.19))}$$

$$\Longrightarrow \begin{cases} \lambda_2(-\theta_2 \boldsymbol{e_1} + 1\boldsymbol{e_2}) \in P \cup I\,, \quad \cdots, \quad \lambda_k(-\theta_k \boldsymbol{e_1} + 1\boldsymbol{e_k}) \in P \cup I\,, \\ \qquad \lambda_{k+1}(-\theta_{k+1}\boldsymbol{e_1} + 1\boldsymbol{e_{k+1}}) \in P \cup I\,, \quad \cdots, \quad \lambda_n(-\theta_n \boldsymbol{e_1} + 1\boldsymbol{e_n}) \in P \cup I\,, \end{cases}$$

$$\Longrightarrow \quad -(\lambda_2\theta_2 + \cdots + \lambda_n\theta_n)\, \boldsymbol{e_1} + \lambda_2 \boldsymbol{e_2} + \cdots + \lambda_n \boldsymbol{e_n} \in P \cup I\,. \qquad \text{(by Lemma 2.2.4)}$$

Hence, with (3.20) and by Lemma 2.2.4 (c), this yields the conclusion

$$\lambda_1 \boldsymbol{e_1} + \lambda_2 \boldsymbol{e_2} + \cdots + \lambda_n \boldsymbol{e_n} \in P\,.$$

In case $\lambda_1 + \lambda_2 r_2 + \cdots + \lambda_n r_n < 0$, we can show similarly $\lambda_1 \boldsymbol{e_1} + \lambda_2 \boldsymbol{e_2} + \cdots + \lambda_n \boldsymbol{e_n} \in -P$.

For the uniqueness part, suppose a tuple of real numbers $(r_1, \dots, r_n) \neq (0, \dots, 0)$ satisfies (3.18) for all $\lambda_1, \dots, \lambda_n \in \mathbb{F}$. We can assume, without loss of generality, that $r_1 \neq 0$. Then, for each $j = 2, \dots, n$, the pair of real numbers $(r_1, r_j)$ satisfies the following conditions for all $\lambda_1, \lambda_j \in \mathbb{F}$,

$$\begin{aligned} \lambda_1 r_1 + \lambda_j r_j > 0 \quad &\text{implies} \quad \lambda_1 \boldsymbol{e_1} + \lambda_j \boldsymbol{e_j} \in P\,, \\ \lambda_1 r_1 + \lambda_j r_j < 0 \quad &\text{implies} \quad \lambda_1 \boldsymbol{e_1} + \lambda_j \boldsymbol{e_j} \in -P\,. \end{aligned} \tag{3.21}$$

By Lemma 3.2.1, such a pair $(r_1, r_j)$ is unique up to a positive scalar multiple. Now, suppose there is another tuple of real numbers $(r'_1, \dots, r'_n) \neq (0, \dots, 0)$ satisfying (3.18) for all $\lambda_1, \dots, \lambda_n \in \mathbb{F}$. In case $r'_1 = 0$, there exists $j$ such that $r'_j \neq 0$. Then, the pair of real numbers $(r'_1, r'_j)$ ($= (0, r'_j)$) also satisfies (3.21) for all $\lambda_1, \lambda_j \in \mathbb{F}$; hence there must be a positive real number $c > 0$ such that $(0, r'_j) = c(r_1, r_j)$, which is impossible because $r_1 \neq 0$. Therefore we have $r'_1 \neq 0$. Thus, for each $j = 2, \dots, n$, the pair of real numbers $(r'_1, r'_j)$ satisfies (3.21) for all $\lambda_1, \lambda_j \in \mathbb{F}$. Hence there is a positive real number $a > 0$ such that $(r'_1, r'_j) = a(r_1, r_j)$ for all $j = 2, \dots, n$. Thus $(r'_1, \dots, r'_n) = a(r_1, \dots, r_n)$. □

Theorem 3.1.2 can be proved by inductive use of Lemma 3.3.1. The following is the idea of the proof: Let $F_n$ be the $n$-dimensional vector space over $\mathbb{F}$, and let $P$ be a nonempty subset of $F_n$ such that $P$ is a positive cone and $\bar{P}$ is a convex cone. Then Lemma 3.3.1 shows that, intuitively, there is a hyperplane by which $P$ and $-P$ are divided. This hyperplane is a $k$-dimensional subspace of $F_n$ for some $k < n$, and hence can be regarded as $F_k$. It is easy to see that the restrictions of $P$ and $\bar{P}$ in this subspace $F_k$ are also a positive cone and a convex cone, respectively. Then Lemma 3.3.1 shows that, again, there is a hyperplane by which $P$ and $-P$ are divided .... Repeating this division, we will obtain a desired hyperplane by which $P$, $I$, and $-P$ are completely divided.

Thus, we shall prove Theorem 3.1.2 by induction on $n$.

**(Base)**

Let $F_1$ be the 1-dimensional vector space over $\mathbb{F}$ with the basis $\boldsymbol{e_1}$, and let $P$ be a nonempty subset of $F_1$ such that $P$ is a positive cone and $\bar{P}$ is a convex cone. Then, by Lemma 2.2.4 (b), $F_1$ is divided into three parts $P$, $I$, and $-P$.

We cannot have $\boldsymbol{e_1} \in I$: for, if $\boldsymbol{e_1} \in I$ then $I = F_1$ by Lemma 2.2.4 (a), contradicting $P \neq \emptyset$.

Suppose $\boldsymbol{e_1} \in P$. Then, for all $\lambda_1 \in \mathbb{F}$,

$$\begin{array}{lll} \lambda_1 > 0 & \text{implies} & \lambda_1 \boldsymbol{e_1} \in P\,, \\ \lambda_1 = 0 & \text{implies} & \lambda_1 \boldsymbol{e_1} \;=\; \boldsymbol{0} \in I\,, \\ \lambda_1 < 0 & \text{implies} & \lambda_1 \boldsymbol{e_1} \in -P\,. \end{array}$$

Since $P$, $I$, and $-P$ are mutually disjoint, the above implications are actually equivalences. Thus $q_1 = 1$ is a solution of (3.3). One can easily verifies that any positive real number $q_1 > 0$ is a solution of (3.3).

If $\boldsymbol{e_1} \in -P$, one can show similarly that any negative real number $q_1 < 0$ is a solution of (3.3).

**(Induction step)**

Let $F_n$ be the $n$-dimensional vector space over $\mathbb{F}$ with the basis $\boldsymbol{e_1}, \ldots, \boldsymbol{e_n}$ for some $n \geq 2$, and let $P$ be a nonempty subset of $F_n$ such that $P$ is a positive cone and $\bar{P}$ is a convex cone. Let $I$ be the associated indifference subspace. By Lemma 3.3.1, there is a tuple of real numbers $(r_1, \ldots, r_n) \neq (0, \ldots, 0)$ such that for all $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$,

$$\begin{array}{lll} \lambda_1 r_1 + \cdots + \lambda_n r_n > 0 & \text{implies} & \lambda_1 \boldsymbol{e_1} + \cdots + \lambda_n \boldsymbol{e_n} \in P\,, \\ \lambda_1 r_1 + \cdots + \lambda_n r_n < 0 & \text{implies} & \lambda_1 \boldsymbol{e_1} + \cdots + \lambda_n \boldsymbol{e_n} \in -P\,. \end{array} \tag{3.22}$$

Let us define a subspace $H^*$ of $F_n$ by

$$H^* \;=\; \{\, \lambda_1 \boldsymbol{e_1} + \cdots + \lambda_n \boldsymbol{e_n} \in F_n \mid \lambda_1 r_1 + \cdots + \lambda_n r_n = 0 \,\}. \tag{3.23}$$

Then $H^*$ is a $k$-dimensional subspace of $F_n$ for some $k < n$.

Suppose $H^*$ is 0-dimensional, i.e. $H^*$ contains only the origin $\boldsymbol{0}$. This means

$$\lambda_1 r_1 + \cdots + \lambda_n r_n = 0 \quad \text{implies} \quad \lambda_1 \boldsymbol{e_1} + \cdots + \lambda_n \boldsymbol{e_n} \;=\; \boldsymbol{0} \in I\,.$$

Then, with the implications of (3.22), one can easily verifies that these implications turn out to be equivalences. Thus $(r_1, \ldots, r_n)$ is a solution of (3.3).

Suppose $H^*$ is $k$-dimensional for some $k > 0$. Then, by Proposition 2.4.2, there exists an orthogonal system $\boldsymbol{\delta_1}, \ldots, \boldsymbol{\delta_k}$ in $H^*$:

$$\boldsymbol{\delta_1} \;=\; d_{11} \boldsymbol{e_1} + \cdots + d_{1n} \boldsymbol{e_n}\,, \quad \cdots, \quad \boldsymbol{\delta_k} \;=\; d_{k1} \boldsymbol{e_1} + \cdots + d_{kn} \boldsymbol{e_n}\,.$$

Note that, by Proposition 2.4.1, any $\boldsymbol{\lambda} \in H^*$ can be written as

$$\boldsymbol{\lambda} \;=\; \left(\frac{\boldsymbol{\lambda}\cdot\boldsymbol{\delta_1}}{\boldsymbol{\delta_1}\cdot\boldsymbol{\delta_1}}\right)\boldsymbol{\delta_1} \;+\;\cdots\;+\; \left(\frac{\boldsymbol{\lambda}\cdot\boldsymbol{\delta_k}}{\boldsymbol{\delta_k}\cdot\boldsymbol{\delta_k}}\right)\boldsymbol{\delta_k}\,. \tag{3.24}$$

Let $P^*$, $I^*$ be the restrictions of $P$, $I$ in the subspace $H^*$, respectively, i.e.

$$P^* \;=\; P \cap H^*. \qquad I^* \;=\; I \cap H^*. \tag{3.25}$$

It is easy to verify that $P^*$ is a positive cone and $\bar{P}^*$ is a convex cone in $H^*$.

If $P^*$ is empty,

$$\lambda_1 r_1 + \cdots + \lambda_n r_n = 0 \quad \text{implies} \quad \lambda_1\boldsymbol{e_1} + \cdots + \lambda_n\boldsymbol{e_n} \in I\,.$$

Then, with the implications of (3.22), one can easily verifies that these implications turn out to be equivalences. Thus $(r_1,\ldots,r_n)$ is a solution of (3.3).

If $P^*$ is nonempty, then, by the induction hypothesis, there are polynomials $p_1,\ldots,p_k \in \mathbb{R}[\varepsilon]_k$ such that for all $\lambda_1,\ldots,\lambda_k \in \mathbb{F}$,

$$\lambda_1 p_1 + \cdots + \lambda_k p_k \;>\; 0 \quad \text{iff} \quad \lambda_1\boldsymbol{\delta_1} + \cdots + \lambda_k\boldsymbol{\delta_k} \;\in P^*. \tag{3.26}$$

From (3.26), it also holds that for all $\lambda_1,\ldots,\lambda_k \in \mathbb{F}$,

$$\begin{aligned} \lambda_1 p_1 + \cdots + \lambda_k p_k \;&<\; 0 \quad \text{iff} \quad \lambda_1\boldsymbol{\delta_1} + \cdots + \lambda_k\boldsymbol{\delta_k} \;\in -P^*,\\ \lambda_1 p_1 + \cdots + \lambda_k p_k \;&=\; 0 \quad \text{iff} \quad \lambda_1\boldsymbol{\delta_1} + \cdots + \lambda_k\boldsymbol{\delta_k} \;\in I^*. \end{aligned} \tag{3.27}$$

Now we claim that the following set of polynomials $q_1,\ldots,q_n \in \mathbb{R}[\varepsilon]$ is a solution of (3.3):

$$\begin{pmatrix} q_1\\ \vdots\\ q_n \end{pmatrix} = \begin{pmatrix} r_1\\ \vdots\\ r_n \end{pmatrix} + \varepsilon \times \left( \frac{p_1}{\boldsymbol{\delta_1}\cdot\boldsymbol{\delta_1}} \begin{pmatrix} d_{11}\\ \vdots\\ d_{1n} \end{pmatrix} + \cdots + \frac{p_k}{\boldsymbol{\delta_k}\cdot\boldsymbol{\delta_k}} \begin{pmatrix} d_{k1}\\ \vdots\\ d_{kn} \end{pmatrix} \right)$$

One can easily verifies that the polynomials $q_1,\ldots,q_n$ are represented by an $n\times n$ real matrix of the form (3.4) such that the first column vector is nonzero. We shall show that $\lambda_1 q_1 + \cdots + \lambda_n q_n > 0$ implies $\lambda_1\boldsymbol{e_1} + \cdots + \lambda_n\boldsymbol{e_n} \in P$ for all $\lambda_1,\ldots,\lambda_n \in \mathbb{F}$:

$$\begin{aligned}
& \lambda_1 q_1 + \cdots + \lambda_n q_n > 0\\
\iff\; & (\,\lambda_1 r_1 + \cdots + \lambda_n r_n\,) \;+\; \varepsilon \times \left( \left(\tfrac{\boldsymbol{\lambda}\cdot\boldsymbol{\delta_1}}{\boldsymbol{\delta_1}\cdot\boldsymbol{\delta_1}}\right) p_1 \;+\;\cdots\;+\; \left(\tfrac{\boldsymbol{\lambda}\cdot\boldsymbol{\delta_k}}{\boldsymbol{\delta_k}\cdot\boldsymbol{\delta_k}}\right) p_k \right) \;>\; 0\\
& \hfill (\text{where } \boldsymbol{\lambda} = \lambda_1\boldsymbol{e_1} + \cdots + \lambda_n\boldsymbol{e_n}\,)\\
\iff\; & \begin{cases} \lambda_1 r_1 + \cdots + \lambda_n r_n > 0 \quad \text{or}\\ \lambda_1 r_1 + \cdots + \lambda_n r_n = 0 \;\text{ and }\; \left(\tfrac{\boldsymbol{\lambda}\cdot\boldsymbol{\delta_1}}{\boldsymbol{\delta_1}\cdot\boldsymbol{\delta_1}}\right) p_1 \;+\;\cdots\;+\; \left(\tfrac{\boldsymbol{\lambda}\cdot\boldsymbol{\delta_k}}{\boldsymbol{\delta_k}\cdot\boldsymbol{\delta_k}}\right) p_k \;>\; 0 \end{cases}\\
\iff\; & \begin{cases} \lambda_1 r_1 + \cdots + \lambda_n r_n > 0 \quad \text{or}\\ \lambda_1\boldsymbol{e_1} + \cdots + \lambda_n\boldsymbol{e_n} \in H^* \;\text{ and }\; \left(\tfrac{\boldsymbol{\lambda}\cdot\boldsymbol{\delta_1}}{\boldsymbol{\delta_1}\cdot\boldsymbol{\delta_1}}\right)\boldsymbol{\delta_1} \;+\;\cdots\;+\; \left(\tfrac{\boldsymbol{\lambda}\cdot\boldsymbol{\delta_k}}{\boldsymbol{\delta_k}\cdot\boldsymbol{\delta_k}}\right)\boldsymbol{\delta_k} \;\in\; P^* \end{cases}\\
& \hfill (\text{by (3.23) and (3.26)})\\
\iff\; & \lambda_1 r_1 + \cdots + \lambda_n r_n > 0 \quad \text{or} \quad \text{“}\,\boldsymbol{\lambda} \in H^* \text{ and } \boldsymbol{\lambda} \in P^*\,\text{”} \qquad (\text{by (3.24)})\\
\Longrightarrow\; & \lambda_1\boldsymbol{e_1} + \cdots + \lambda_n\boldsymbol{e_n} \in P\,. \qquad (\text{by (3.22) or (3.25)})
\end{aligned}$$

Similarly, with (3.27), one can show that for all $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$,

$$\begin{aligned} \lambda_1 q_1 + \cdots + \lambda_n q_n < 0 \quad &\text{implies} \quad \lambda_1 \boldsymbol{e_1} + \cdots + \lambda_n \boldsymbol{e_n} \in -P\,, \\ \lambda_1 q_1 + \cdots + \lambda_n q_n = 0 \quad &\text{implies} \quad \lambda_1 \boldsymbol{e_1} + \cdots + \lambda_n \boldsymbol{e_n} \in I\,. \end{aligned}$$

Hence, these implications are actually equivalences.

For the uniqueness part of Theorem 3.1.2, suppose a set of polynomials $q'_1, \ldots, q'_n \in \mathbb{R}[\varepsilon]$ satisfies (3.3) for all $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$. Suppose also the set of polynomials $q'_1, \ldots, q'_n$ is represented by an $n \times n$ real matrix $\{r'_{ij}\}_{1 \le i,j \le n}$ of the form (3.4) such that the first column vector $(r'_{11}, \ldots, r'_{n1})^T$ is nonzero. Then the following holds for all $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$:

if $\lambda_1 r'_{11} + \cdots + \lambda_n r'_{n1} > 0$

then $\lambda_1 q'_1 + \cdots + \lambda_n q'_n = (\lambda_1 r'_{11} + \cdots + \lambda_n r'_{n1}) + \varepsilon\varphi(\varepsilon) > 0$

( where $\varphi(\varepsilon) = s_1 + s_2\varepsilon + \cdots + s_{n-1}\varepsilon^{n-2}$ for some reals $s_i$ )

hence $\lambda_1 \boldsymbol{e_1} + \cdots + \lambda_n \boldsymbol{e_n} \in P\,,$ (by (3.3))

if $\lambda_1 r'_{11} + \cdots + \lambda_2 r'_{n1} < 0$

then $\lambda_1 q'_1 + \cdots + \lambda_2 q'_n < 0$ hence $\lambda_1 \boldsymbol{e_1} + \cdots + \lambda_n \boldsymbol{e_n} \in -P\,.$

By Lemma 3.3.1, the vector $(r'_{11}, \ldots, r'_{n1})^T$ is unique up to a positive scalar multiple.

**Q.E.D.**

# Chapter 4

# Applications to Linear Inequality Systems

In the following chapters we shall provide some applications of the lexicographical separation theorem which was proved in the previous chapter. Here we shall deal with the topics of infinite systems of linear inequalities. In Section 4.1 the main result of this chapter is presented (Theorem 4.1.2), giving a necessary and sufficient condition for the existence of solutions to infinite systems of strict linear inequalities, where the solutions are allowed to be infinitely small. The result is a generalization of the well-known theorem of the alternatives for finite linear inequality systems. We also give a sufficient condition for the existence of real-valued solutions to infinite systems of linear inequalities (Theorem 4.1.5). Further, we obtain a Farkas type theorem (Theorem 4.1.6) for lexicographical inequality systems. In Section 4.2 we apply the results obtained in Section 4.1 to game theory, giving a generalization of von Neumann's minimax theorem for semi-infinite games.

## 4.1 Infinite Systems of Linear Inequalities

Let $P$ be an nonempty subset of $\mathbb{R}^n$. Consider the following system of strict linear inequalities:

$$\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n \ > \ 0 \quad \text{for all} \ \ (\lambda_1, \lambda_2, \ldots, \lambda_n) \in P . \tag{4.1}$$

It is well known that, in case $P$ is a finite set, there are several equivalent versions of linear existence theorems which give solutions to (4.1), such as the theorem of the alternatives, Farkas' lemma, Motzkin's lemma, Gordan's lemma, and so on. Here we present a version of these theorems:

**Proposition 4.1.1** (Theorem of the Alternatives) Let $P$ be a finite subset of $\mathbb{R}^n$. Then, $\mathbf{0}$ is not a convex combination[1] of $P$ if and only if the system (4.1) has solutions $x_1, x_2, \ldots, x_n$ in $\mathbb{R}$. □

For a proof see Proposition 1.1.3. Readers should consult Gale [12], Skala [42] for other standard results on finite linear inequality systems.

[1]We say that $\boldsymbol{x}$ is a *convex combination of* $S$ if there exist finite $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k \in S$ and $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ with $\lambda_i > 0$ for $i = 1, \ldots, k$ and $\sum_{i=1}^{k} \lambda_i = 1$ such that $\boldsymbol{x} = \lambda_1 \boldsymbol{y}_1 + \cdots + \lambda_k \boldsymbol{y}_k$.

Proposition 4.1.1 cannot however be directly generalized for the infinite cases. Let

$$P' \;=\; \{\,(1,0),\,(0,1),\,(-1,1),\,(-2,1),\,\ldots,\,(-k,1),\,\ldots\,\}\,.$$

One can easily verifies that $(0,0)$ is not a convex combination of $P'$. Consider the corresponding system of linear inequalities:

$$x > 0\,, \quad y > 0\,, \quad \text{and} \quad -kx + y > 0 \;\text{ for all positive integer } k\,. \tag{4.2}$$

The above system (4.2) has no real-valued solutions: for, (4.2) means "$0 < x < (1/k)y$ for all positive integer $k$," which is impossible in $\mathbb{R}$.

Let $(x,y) = (\varepsilon,1)$ be an imaginary solution to (4.2); that is to say, $0 < \varepsilon$ and $\varepsilon < 1/k$ for all positive integer $k$. We call $\varepsilon$ an *infinitesimal*. Geometrically, this means that there exists a line through $\mathbf{0}$ whose gradient is $-\varepsilon$ such that $P'$ lies entirely "above" the line (see Figure 4.1).

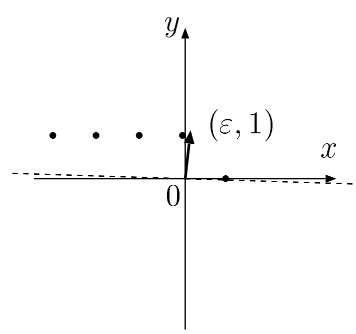


Figure 4.1: Geometrical Interpretation of $\varepsilon$

We shall show that Proposition 4.1.1 can be generalized for the infinite cases if we adopt such an infinitesimal $\varepsilon$ as a solution to the system of strict linear inequalities:

**Theorem 4.1.2** *Let $P$ be an nonempty subset of $\mathbb{R}^n$. Then, $\mathbf{0}$ is not a convex combination of $P$ if and only if the system*

$$\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n \;>\; 0 \quad \textit{for all} \;\; (\lambda_1, \lambda_2, \ldots, \lambda_n) \,\in\, P \tag{4.3}$$

*has solutions $x_1, x_2, \ldots, x_n$ in $\mathbb{R}[\varepsilon]_n$.* [2]

**Proof.** (**if**) Suppose (4.3) has solutions $x_1, x_2, \ldots, x_n$ in $\mathbb{R}[\varepsilon]_n$, and let $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$. Suppose also $\mathbf{0}$ is a convex combination of $P$, i.e. there exist $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k \in P$ and $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ with $\lambda_i > 0$ for $i = 1, \ldots, k$ such that $\mathbf{0} = \lambda_1 \boldsymbol{y}_1 + \cdots + \lambda_k \boldsymbol{y}_k$. Then

$$0 \;=\; \boldsymbol{x} \cdot \mathbf{0} \;=\; \boldsymbol{x} \cdot \Big(\sum_{k=1}^{m} \lambda_k\, \boldsymbol{y}_k\Big) \;=\; \sum_{k=1}^{m} \lambda_k\, (\boldsymbol{x} \cdot \boldsymbol{y}_k) \;>\; 0\,,$$

a contradiction.

[2]For the definition of $\mathbb{R}[\varepsilon]_n$, see Definition 2.1.5 and Definition 2.1.6.

(**only if**) Let $P'$ be a subset of $\mathbb{R}^n$, and suppose $\mathbf{0}$ is not a convex combination of $P'$. We shall construct a subset $P$ of $\mathbb{R}^n$ in such a way that $P \supseteq P'$, $P$ is a positive cone, and $\bar{P}$ is a convex cone in $\mathbb{R}^n$. Once the existence of such $P$ is confirmed, by Theorem 3.1.2 there exist polynomials $q_1, \ldots, q_n \in \mathbb{R}[\varepsilon]_n$ such that (3.3) holds for all $(\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$, and hence $(x_1, \ldots, x_n) = (q_1, \ldots, q_n)$ is a desired solution to (4.3).

By Zorn's lemma, there exists a maximal subset $P$ of $\mathbb{R}^n$ (with respect to inclusion) such that

$$\text{(i)} \quad P \supseteq P', \qquad \text{(ii)} \quad \mathbf{0} \text{ is not a convex combination of } P. \tag{4.4}$$

Then, the following conditions hold for $P$:

(1) $\quad \mathbf{0} \notin P$,

(2) $\quad$ for all $\boldsymbol{x} \in \mathbb{R}^n$, $\quad \boldsymbol{x} \in P$ or $\boldsymbol{x} = \mathbf{0}$ or $-\boldsymbol{x} \in P$,

(3) $\quad$ for all $\boldsymbol{x} \in \mathbb{R}^n$, $\quad \boldsymbol{x} \in P$ and $\lambda > 0$ imply $\lambda \boldsymbol{x} \in P$,

(4) $\quad$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, $\quad \boldsymbol{x} \in P$ and $\boldsymbol{y} \in P$ imply $\boldsymbol{x} + \boldsymbol{y} \in P$.

(1) is trivial. We shall verify (2), (3), and (4) as follows.

(2): Let $\boldsymbol{x}$ be a nonzero vector in $\mathbb{R}^n$. We shall show that either $\boldsymbol{x}$ or $-\boldsymbol{x}$ is in $P$. Suppose otherwise: $\boldsymbol{x} \notin P$ and $-\boldsymbol{x} \notin P$. Then, by the maximality of $P$ with respect to (4.4), $\mathbf{0}$ is a convex combination of $P \cup \{\boldsymbol{x}\}$, and also $\mathbf{0}$ is a convex combination of $P \cup \{-\boldsymbol{x}\}$. Hence, there exist $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_l, \boldsymbol{q}_1, \ldots, \boldsymbol{q}_m \in P$ and $\lambda_0, \ldots, \lambda_l, \mu_0, \ldots, \mu_m \in \mathbb{R}$ such that

$$\lambda_0 \boldsymbol{x} + \sum_{k=1}^{l} \lambda_k \boldsymbol{p}_k = \mathbf{0}, \qquad \lambda_i > 0 \quad \text{for } i = 0, \ldots, l, \quad \text{and} \quad \textstyle\sum_{i=0}^{l} \lambda_i = 1,$$

$$\mu_0 (-\boldsymbol{x}) + \sum_{k=1}^{m} \mu_k \boldsymbol{q}_k = \mathbf{0}, \qquad \mu_j > 0 \quad \text{for } j = 0, \ldots, m, \quad \text{and} \quad \textstyle\sum_{i=0}^{m} \mu_i = 1.$$

Therefore $\mu_0 (\sum_{k=1}^{l} \lambda_k \boldsymbol{p}_k) + \lambda_0 (\sum_{k=1}^{m} \mu_k \boldsymbol{q}_k) = \mathbf{0}$. Let $\xi := \mu_0 (\sum_{k=1}^{l} \lambda_k) + \lambda_0 (\sum_{k=1}^{m} \mu_k)$. Then

$$\frac{1}{\xi} \mu_0 \Big( \sum_{k=1}^{l} \lambda_k \boldsymbol{p}_k \Big) + \frac{1}{\xi} \lambda_0 \Big( \sum_{k=1}^{m} \mu_k \boldsymbol{q}_k \Big) = \mathbf{0},$$

which means that $\mathbf{0}$ is a convex combination of $P$, a contradiction.

(3): Suppose $\boldsymbol{x} \in P$, $\lambda > 0$, and $\lambda \boldsymbol{x} \notin P$. Then, by the maximality of $P$ with respect to (4.4), $\mathbf{0}$ is a convex combination of $P \cup \{ \lambda \boldsymbol{x} \}$. Hence, there exist $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_l \in P$ and $\mu_0, \ldots, \mu_l \in \mathbb{R}$ such that

$$\mu_0 (\lambda \boldsymbol{x}) + \sum_{k=1}^{l} \mu_k \boldsymbol{p}_k = \mathbf{0}, \qquad \mu_i > 0 \quad \text{for } i = 0, \ldots, l, \quad \text{and} \quad \textstyle\sum_{i=0}^{l} \mu_i = 1.$$

Let $\xi := \mu_0 \lambda + \sum_{k=1}^{l} \mu_k$. Then $\dfrac{1}{\xi} \Big( \mu_0 \lambda \boldsymbol{x} + \displaystyle\sum_{k=1}^{l} \mu_k \boldsymbol{p}_k \Big) = \mathbf{0}$. This means that $\mathbf{0}$ is a convex combination of $P$, a contradiction.

(4): Suppose there exist $\boldsymbol{x}$, $\boldsymbol{y}$ such that $\boldsymbol{x} \in P$, $\boldsymbol{y} \in P$, and $\boldsymbol{x} + \boldsymbol{y} \notin P$. Then, by (2), we have either $\boldsymbol{x} + \boldsymbol{y} = \mathbf{0}$ or $-\boldsymbol{x} - \boldsymbol{y} \in P$; both cases contradict the hypothesis that $\mathbf{0}$ is not a convex combination of $P$.

Thus, by (1), (3), and (4), $P$ is a positive cone in $\mathbb{R}^n$. Now, by (1) and (2), we have exactly one of $\boldsymbol{x} \in P$, $\boldsymbol{x} = \boldsymbol{0}$, and $\boldsymbol{x} \in -P$ for all $\boldsymbol{x} \in \mathbb{R}^n$. Hence $\bar{P} = -P \cup \{\boldsymbol{0}\}$. Therefore $\bar{P}$ is a convex cone in $\mathbb{R}^n$. □

**Example 4.1.3** Let $P_1$ ($\subseteq \mathbb{R}^2$) be

$$P_1 \;=\; \{(1,-1), (-1,2), (-2,3), \ldots, (-k, k+1), \ldots\}.$$

One can easily verifies that $(0,0)$ is not a convex combination of $P_1$, but the corresponding system

$$x - y > 0\,, \quad -x + 2y > 0\,, \quad -2x + 3y > 0\,, \;\ldots\,, \quad -kx + (k+1)y > 0\,, \;\ldots$$

has no *real-valued* solution. By Theorem 4.1.2, the above system has solutions $x, y$ in $\mathbb{R}[\varepsilon]_2$: for example, $(x,y) = (1+\varepsilon\,,\,1)$, where $\varepsilon$ is an infinitesimal. □

**Example 4.1.4** Let $P_2$ ($\subseteq \mathbb{R}^3$) be

$$\begin{aligned} P_2 \;=\; \{(1,0,0), (0,1,0), (0,0,1),\} \;&\cup\; \{(1,-1,0), (1,-2,0), \ldots, (1,-k,0), \ldots\} \\ &\cup\; \{(0,1,-1), (0,1,-2), \ldots, (0,1,-k), \ldots\}. \end{aligned}$$

It is not difficult to verify that $(0,0,0)$ is not a convex combination of $P_2$. By Theorem 4.1.2, the corresponding system

$$\begin{aligned} x > 0\,,\; y > 0\,,\; z > 0\,, \quad & x - y > 0\,,\; x - 2y > 0\,, \;\ldots\,,\; x - ky > 0\,, \;\ldots\,, \\ & y - z > 0\,,\; y - 2z > 0\,, \;\ldots\,,\; y - kz > 0\,, \;\ldots \end{aligned}$$

has solutions $x, y, z$ in $\mathbb{R}[\varepsilon]_3$: for example, $(x,y,z) = (1, \varepsilon, \varepsilon^2)$. □

Theorem 4.1.2 has the following corollary, which gives a sufficient condition for the existence of real-valued solutions to linear inequality systems. Note that this can also be derived as a consequence of the well-known supporting hyperplane lemma (see e.g. Rockafellar [39]).

**Theorem 4.1.5** *Let $P$ be an nonempty subset of $\mathbb{R}^n$. If $\boldsymbol{0}$ is not a convex combination of $P$, then there exist $r_1, r_2, \ldots, r_n \in \mathbb{R}$ such that $(r_1, \ldots, r_n) \neq (0, \ldots, 0)$ and*

$$\lambda_1 r_1 + \lambda_2 r_2 + \cdots + \lambda_n r_n \;\geq\; 0 \qquad \textit{for all } (\lambda_1, \lambda_2, \ldots, \lambda_n) \in P\,.$$

**Proof.** Let $P'$ be an nonempty subset of $\mathbb{R}^n$, and suppose $\boldsymbol{0}$ is not a convex combination of $P'$. As in the proof of Theorem 4.1.2, we can construct a subset $P$ of $\mathbb{R}^n$ in such a way that $P \supseteq P'$, $P$ is a positive cone, and $\bar{P}$ is a convex cone in $\mathbb{R}^n$. By Theorem 3.1.2, there exist polynomials $q_1, \ldots, q_n \in \mathbb{R}[\varepsilon]$ such that (3.3) holds for all $(\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$; moreover, the polynomials $q_1, \ldots, q_n$ are written as

$$q_i \;=\; r_{i1} + r_{i2}\varepsilon + \cdots + r_{in}\varepsilon^{n-1} \quad (r_{i1}, \ldots, r_{in} \in \mathbb{R}) \qquad \text{for } i = 1, \ldots, n\,,$$

such that $(r_{11}, \ldots, r_{n1}) \neq (0, \ldots, 0)$. Therefore, by (3.3),

$$\lambda_1 q_1 + \cdots + \lambda_n q_n \;>\; 0 \qquad \text{for all } (\lambda_1, \ldots, \lambda_n) \in P\,. \tag{4.5}$$

Now, it is easy to verify that for all $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$,

$$\begin{aligned} \lambda_1 q_1 + \cdots + \lambda_n q_n \; > \; 0 \quad \text{iff} \quad & (\lambda_1 r_{11} + \cdots + \lambda_n r_{n1}) \\ & + \varepsilon\,(\lambda_1 r_{12} + \cdots + \lambda_n r_{n2}) \\ & + \cdots + \varepsilon^{n-1}(\lambda_1 r_{1n} + \cdots + \lambda_n r_{nn}) \; > \; 0\,, \end{aligned}$$

and hence $\lambda_1 q_1 + \cdots + \lambda_n q_n > 0$ implies $\lambda_1 r_{11} + \cdots + \lambda_n r_{n1} \geq 0$. Therefore, by (4.5),

$$\lambda_1 r_{11} + \cdots + \lambda_n r_{n1} \;\geq\; 0 \qquad \text{for all } (\lambda_1, \ldots, \lambda_1) \in P\,.$$

Thus $(r_{11}, \ldots, r_{n1})$ is a desired solution. □

Martínez-Legaz [28] presented a generalization of Farkas' lemma (see Proposition 1.1.4) for lexicographical consequences of linear inequality systems. Here we shall give another generalization of Farkas' lemma for lexicographical inequality systems. Let $S$ be a subset of $\mathbb{R}^n$. We denote by $\mathrm{cone}(S)$ the smallest convex cone containing $S$ and $\mathbf{0}$; that is to say, $\mathrm{cone}(S)$ consists of all the finite sums $\sum_{i=1}^{k} \lambda_i \boldsymbol{s}_i$ such that $\lambda_i \geq 0$ and $\boldsymbol{s}_i \in S$ for $i = 1, \ldots, k$.

**Theorem 4.1.6** *Let $S$ be an nonempty subset of $\mathbb{R}^n$, and let $\boldsymbol{c} \in \mathbb{R}^n$. Then, $\boldsymbol{c} \in \mathrm{cone}(S)$ if and only if for all $n \times n$ real matrices $A$,*

$$\text{“}\, A\boldsymbol{x} \leqq_L \mathbf{0} \text{ for all } \boldsymbol{x} \in S\,\text{”} \quad \text{implies} \quad A\boldsymbol{c} \leqq_L \mathbf{0}\,. \quad {}^{3}$$

**Proof.** (**only if**) Let $A$ be an $n \times n$ real matrix, and let $\boldsymbol{c} \in \mathbb{R}^n$. Suppose $\boldsymbol{c} \in \mathrm{cone}(S)$ and $A\boldsymbol{x} \leqq_L \mathbf{0}$ for all $\boldsymbol{x} \in S$. Then, $\boldsymbol{c} \in \mathrm{cone}(S)$ means there exist finite $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_k \in S$ and $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ with $\lambda_i \geq 0$ for $i = 1, \ldots, k$ such that $\boldsymbol{c} = \sum_{i=1}^{k} \lambda_i \boldsymbol{s}_i$. Therefore

$$A\boldsymbol{c} \;=\; A\Big(\sum_{i=1}^{k} \lambda_i \boldsymbol{s}_i\Big) \;=\; \sum_{i=1}^{k} \lambda_i\,(A\boldsymbol{s}_i) \;\leqq_L\; \mathbf{0}\,.$$

(**if**) It is well known by the standard results of linear algebra that, to any $n \times n$ matrix $A$, there corresponds a linear operator $g : \mathbb{R}^n \to \mathbb{R}^n$ such that $g(\boldsymbol{x}) = A\boldsymbol{x}$ for all $\boldsymbol{x} \in \mathbb{R}^n$. Hence it is enough to show that, if $\boldsymbol{c} \notin \mathrm{cone}(S)$, then there is a linear operator $g$ on $\mathbb{R}^n$ such that

$$g(\boldsymbol{x}) \leqq_L \mathbf{0} \quad \text{for all } \boldsymbol{x} \in S \qquad \text{and} \qquad g(\boldsymbol{c}) >_L \mathbf{0}\,. \tag{4.6}$$

Suppose $\boldsymbol{c} \notin \mathrm{cone}(S)$. We shall construct a subset $P$ of $\mathbb{R}^n$ in such a way that $\boldsymbol{c} \in P$, $P \cap \mathrm{cone}(S) = \emptyset$, $P$ is a positive cone in $\mathbb{R}^n$, and $\bar{P}$ is a convex cone in $\mathbb{R}^n$. Once the existence of such $P$ is confirmed, by Theorem 3.1.1 there exist real-valued linear functions $g_1, \ldots, g_n$ on $\mathbb{R}^n$ such that for all $\boldsymbol{x} \in \mathbb{R}^n$,

$$x \in P \quad \text{iff} \quad (\,g_1(x), \ldots, g_n(x)\,)^T \;>_L\; (\,0, \ldots, 0\,)^T,$$

and hence $g(\boldsymbol{x}) = (\,g_1(\boldsymbol{x}), \ldots, g_n(\boldsymbol{x})\,)^T$ is a desired solution of (4.6).

By Zorn's lemma, there exists a maximal subset $P$ of $\mathbb{R}^n$ (with respect to inclusion) such that

$$\text{(i)} \quad \boldsymbol{c} \in P\,, \qquad \text{(ii)} \quad P \cap \mathrm{cone}(S) = \emptyset\,, \qquad \text{(iii)} \quad P \text{ is a convex cone.} \tag{4.7}$$

[3] Here, the elements of $\mathbb{R}^n$ are considered as column vectors; and also the lexicographic ordering $\leqq_L$ of the column vectors is considered in the same manner as the row vectors (see Definition 2.1.2).

Note that $P_0 = \{ \lambda \boldsymbol{c} \mid \lambda > 0 \}$ satisfies (i)–(iii), and hence these conditions are consistent. Now we have $\mathbf{0} \notin P$ by (ii), and hence, with (iii), $P$ is a positive cone. It remains to show that $\bar{P}$ is a convex cone:

$$\text{(1)} \qquad \text{for all } \boldsymbol{x} \in \mathbb{R}^n, \quad \boldsymbol{x} \in \bar{P} \text{ and } \lambda > 0 \text{ imply } \lambda \boldsymbol{x} \in \bar{P},$$

$$\text{(2)} \qquad \text{for all } \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n, \quad \boldsymbol{x} \in \bar{P} \text{ and } \boldsymbol{y} \in \bar{P} \text{ imply } \boldsymbol{x} + \boldsymbol{y} \in \bar{P}.$$

(1) is an immediate consequence of (iii). To verify (2), suppose there exist $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$ such that $\boldsymbol{a} \in \bar{P}$, $\boldsymbol{b} \in \bar{P}$, and $\boldsymbol{a} + \boldsymbol{b} \in P$. By the maximality of $P$ with respect to (4.7), $\boldsymbol{a} \notin P$ means that there exist $\boldsymbol{p} \in P$ and $\lambda \in \mathbb{R}$ with $\lambda > 0$ such that

$$\lambda \boldsymbol{a} + \boldsymbol{p} \in \operatorname{cone}(S).$$

Similarly, $\boldsymbol{b} \notin P$ means that there exist $\boldsymbol{q} \in P$ and $\mu \in \mathbb{R}$ with $\mu > 0$ such that

$$\mu \boldsymbol{b} + \boldsymbol{q} \in \operatorname{cone}(S).$$

Hence, $\lambda\mu (\boldsymbol{a}+\boldsymbol{b}) + \mu \boldsymbol{p} + \lambda \boldsymbol{q} \in \operatorname{cone}(S)$. But, by (iii) and the hypotheses $\boldsymbol{a}+\boldsymbol{b}, \boldsymbol{p}, \boldsymbol{q} \in P$, we have $\lambda\mu (\boldsymbol{a}+\boldsymbol{b}) + \mu \boldsymbol{p} + \lambda \boldsymbol{q} \in P$. Therefore $P \cap \operatorname{cone}(S) \neq \emptyset$, contradicting (ii). □

## 4.2 A Minimax Theorem for Semi-Infinite Games

In this section we consider semi-infinite games, i.e. zero-sum two-person games in which one of the players has infinitely many strategies, and prove a minimax theorem using the result of the previous section. (For the fundamentals of game theory see e.g. Owen [37].) The minimax theorem for finite zero-sum two-person games was first proved by von Neumann [44], and its generalizations for semi-infinite games was introduced by Wald [46], Blackwell and Girshick [2]. In these theorems, boundedness of payoff functions are assumed. Here we present a new minimax theorem with no assumption of boundedness. The result can also be derived as a consequence of Sion's minimax theorem [41] which was proved in a more general setting, but our proof given here is rather elementary.

Consider a zero-sum two-person game in which one of the players has infinitely many strategies. Let $I = \{1, 2, \ldots, m\}$ be a set of pure strategies for player I, and let $\Xi$ be an arbitrary set of pure strategies for player II. The game is denoted by $[s_{i\xi}]_{i \in I, \xi \in \Xi}$ ($s_{i\xi} \in \mathbb{R}$), where $s_{i\xi}$ is the outcome of the game when $i \in I$ is player I's choice and $\xi \in \Xi$ is player II's. Let

$$S = \{ (s_{1\xi}, s_{2\xi}, \ldots, s_{m\xi}) \mid \xi \in \Xi \}.$$

The game is played as follows: Player II selects $s = (s_1, s_2, \ldots, s_m) \in S$, and simultaneously player I selects a coordinate $i$. The outcome of the game is the payoff $s_i$ to player I from player II. Such a game is called an *S-game* (following Blackwell and Girshick [2]). If $\Xi$ is countable, an $S$-game can be represented by an $m \times \infty$ matrix, such as

$$\begin{pmatrix} 1 & 1/2 & 1/3 & 1/4 & 1/5 & \cdots \\ 0 & 1 & 2 & 3 & 4 & \cdots \end{pmatrix}$$

where each row represents a strategy of player I, and each column a strategy of player II.

**Definition 4.2.1** A *mixed strategy* for each player is a probability distribution with finite support on his set of pure strategies.

In other words, a mixed strategy for player I is an $m$-dimensional vector $x = (x_1, \ldots, x_m)$ satisfying $x_i \geq 0$ for $i = 1, \ldots, m$ and

$$\sum_{i=1}^{m} x_i \; = \; 1 \, .$$

For player II, a mixed strategy is a mapping $y : \Xi \to [0,1]$ with the following properties:

$$\text{(i)} \quad \sum_{\xi \in \Xi} y(\xi) = 1, \quad \text{(ii)} \;\; \text{there is a finite } \Theta \subseteq \Xi \text{ such that } y(\xi) = 0 \;\text{ for all } \xi \in \Xi \setminus \Theta \, .$$

We shall denote by $X$ the set of all mixed strategies for player I, and let $Y$ denote the set of player II's mixed strategies. Also, for $y \in Y$ and $\xi \in \Xi$, we denote $y(\xi)$ by $y_\xi$ in the following. □

Suppose that players I and II are playing an $S$-game $[s_{i\xi}]_{i \in I, \xi \in \Xi}$. If player I chooses the mixed strategy $x = (x_1, \ldots, x_m) \in X$ and player II chooses $y \in Y$, then the expected payoff will be

$$S(x,y) \; = \; \sum_{i=1}^{m} \sum_{\xi \in \Xi} x_i \, s_{i\xi} \, y_\xi \, .$$

Player I's expected gain-floor, assuming he uses $x$, will be

$$v(x) \; = \; \inf_{y \in Y} S(x,y) \, ,$$

Player I should choose $x$ so as to maximize $v(x)$, i.e. so as to obtain

$$v_{\text{I}} \; = \; \sup_{x \in X} \inf_{y \in Y} S(x,y) \, .$$

Similarly, if player II chooses $y$ he will obtain the expected loss-ceiling

$$v(y) \; = \; \sup_{x \in X} S(x,y) \, ,$$

and he should choose $y$ so as to obtain

$$v_{\text{II}} \; = \; \inf_{y \in Y} \sup_{x \in X} S(x,y) \, .$$

( Note that $v_{\text{I}}, v_{\text{II}} \in \mathbb{R} \cup \{-\infty, \infty\}$. ) These numbers $v_{\text{I}}$ and $v_{\text{II}}$ are called the *values* of the game for I and II, respectively.

It can easily be proved that, for any function $F(x,y)$ defined on any cartesian product $X \times Y$,

$$\sup_{x \in X} \inf_{y \in Y} F(x,y) \;\leq\; \inf_{y \in Y} \sup_{x \in X} F(x,y) \, .$$

Hence we have

$$v_{\text{I}} \;\leq\; v_{\text{II}} \, .$$

It's just that player I's gain-floor does not exceed II's loss-ceiling. Now, we can actually prove the following theorem, which is a generalization of the well-known minimax theorem:

**Theorem 4.2.2** *For any $S$-game,*

$$v_{\mathrm{I}} \;=\; v_{\mathrm{II}}\,.$$

**Proof.** We shall prove this theorem by way of the method of separating two convex sets by a hyperplane. (The method was first given by von Neumann and Morgenstern [45] for the finite case.) Let $[s_{i\xi}]_{i\in I,\xi\in\Xi}$ be an $S$-game. Let

$$S \;=\; \{\,(s_{1\xi}\,, s_{2\xi}\,, \ldots, s_{m\xi}\,) \;\mid\; \xi\in\Xi\,\}\,.$$

Then $S$ can be considered as a set of $m$-diminsional vectors in $\mathbb{R}^m$. Let $\boldsymbol{e_1},\ldots,\boldsymbol{e_m}$ be the basis of $\mathbb{R}^m$, i.e.

$$\boldsymbol{e_1} \;=\; (1,0,\ldots,0)\,, \qquad \cdots \qquad \boldsymbol{e_m} \;=\; (0,0,\ldots,1)\,.$$

Let $P = S \cup \{\,\boldsymbol{e_1},\ldots,\boldsymbol{e_m}\,\}$.

By Theorem 4.1.5, at least one of (1) and (2) must hold:

(1) There exist a finite number of vectors $\boldsymbol{p_1},\ldots,\boldsymbol{p_k}$ in $P$ and real numbers $u_1,\ldots,u_k$ such that

$$\begin{aligned} u_i \;&>\; 0 \qquad \text{for } i=1,\ldots,k\,,\\ \sum_{i=1}^{k} u_i\,\boldsymbol{p}_i \;&=\; 0\,. \end{aligned}$$

(2) There exists a vector $\boldsymbol{r}=(r_1,\ldots,r_m)$ such that $(r_1,\ldots,r_m)\neq(0,\ldots,0)$ and

$$\boldsymbol{r}\cdot\boldsymbol{p} \;\geq\; 0 \qquad \text{for all } \boldsymbol{p}\in P\,.$$

Suppose (1) holds. If all $\boldsymbol{p_1},\ldots,\boldsymbol{p_k}$ belong to the basis $\{\,\boldsymbol{e_1},\ldots,\boldsymbol{e_m}\,\}$, it will follow that $\boldsymbol{0}$ is a convex combination of $\boldsymbol{e_1},\ldots,\boldsymbol{e_m}$, which is obviously impossible because they are independent vectors. Hence at least one of $\boldsymbol{p_1},\ldots,\boldsymbol{p_k}$ is in $S$. Hence, (1) means there exist a finite number of vectors

$$\begin{aligned} \boldsymbol{s_{\xi_1}} \;&=\; (s_{1\xi_1},\ldots,s_{m\xi_1})\,,\\ &\;\;\vdots\\ \boldsymbol{s_{\xi_n}} \;&=\; (s_{1\xi_n},\ldots,s_{m\xi_n}) \end{aligned}$$

in $S$ and real numbers $u_1,\ldots,u_n$, $v_1,\ldots,v_m$ such that

$$\begin{aligned} u_l \;&>\; 0 \qquad \text{for } l=1,\ldots,n\,,\\ v_k \;&\geq\; 0 \qquad \text{for } k=1,\ldots,m\,,\\ \sum_{l=1}^{n} u_l\,\boldsymbol{s_{\xi_l}} \;+\; \sum_{k=1}^{m} v_k\,\boldsymbol{e_k} \;&=\; 0\,. \end{aligned}$$

The last equation can also be written as

$$\sum_{l=1}^{n} u_l\,s_{i\xi_l} \;+\; v_i \;\;=\;\; 0 \qquad \text{for } i=1,\ldots,m\,. \tag{4.8}$$

We define a mapping $y : \Xi \to [0,1]$ by

$$\begin{aligned} y_{\xi_l} &= u_l \Big/ \sum_{l=1}^{n} u_l && \text{for } l = 1, \ldots, n\,, \\ y_\xi &= 0 && \text{for } \xi \in \Xi \setminus \{\xi_1, \ldots, \xi_n\}\,. \end{aligned}$$

Then one can easily verifies that $y$ is a mixed strategy for player II. By (4.8), we have

$$\sum_{\xi \in \Xi} y_\xi\, s_{i\xi} = \sum_{l=1}^{n} y_{\xi_l}\, s_{i\xi_l} = -v_i \Big/ \sum_{l=1}^{n} u_l \leq 0 \qquad \text{for } i = 1, \ldots, m\,.$$

Hence $v(y) \leq 0$: because

$$v(y) = \sup_{x \in X} S(x, y) = \sup_{x \in X} \Big( \sum_{i=1}^{m} x_i \sum_{\xi \in \Xi} y_\xi\, s_{i\xi} \Big) \leq 0\,.$$

Therefore $v_{\text{II}} \leq 0$.

Suppose, instead, that (2) holds. As the unit vectors $\boldsymbol{e_1}, \ldots, \boldsymbol{e_m}$ are in $P$, we have

$$\boldsymbol{r} \cdot \boldsymbol{e_i} = r_i \geq 0 \qquad \text{for } i = 1, \ldots, m\,,$$

and hence, with the hypothesis $(r_1, \ldots, r_m) \neq (0, \ldots, 0)$, it follows that $\sum r_i > 0$. Let

$$x_i = r_i \Big/ \sum r_i \qquad \text{for } i = 1, \ldots, m\,.$$

Then $\boldsymbol{x} = (x_1, \ldots, x_m)$ satisfies

$$\begin{aligned} x_i &\geq 0\,, && \text{for } i = 1, \ldots, m\,, \\ \sum_{i=1}^{m} x_i &= 1\,, \end{aligned}$$

and $\boldsymbol{x} \cdot \boldsymbol{s} \geq 0$ for all $\boldsymbol{s} \in S$, i.e.

$$\sum_{i=1}^{m} x_i\, s_{i\xi} \geq 0\,, \qquad \text{for all } \xi \in \Xi\,.$$

Hence $v(\boldsymbol{x}) \geq 0$: because

$$v(\boldsymbol{x}) = \inf_{y \in Y} S(\boldsymbol{x}, y) = \inf_{y \in Y} \Big( \sum_{\xi \in \Xi} y_\xi \sum_{i=1}^{m} x_i\, s_{i\xi} \Big) \geq 0\,.$$

Therefore $v_{\text{I}} \geq 0$.

We have shown, therefore, that it is not possible to have $v_{\text{I}} < 0 < v_{\text{II}}$. Let us suppose, now, that the $S$-game is changed into the $S'$-game $[s'_{i\xi}]_{i \in I, \xi \in \Xi}$ such that

$$s'_{i\xi} = s_{i\xi} + k\,.$$

It can easily be checked that for any $x, y$,

$$S'(x,y) \;=\; S(x,y) \;+\; k\,.$$

Hence

$$v_{\text{I}}(S') \;=\; v_{\text{I}}(S) \;+\; k\,,$$
$$v_{\text{II}}(S') \;=\; v_{\text{II}}(S) \;+\; k\,.$$

(For convenience sake, we assume $\infty = \infty + k$ and $-\infty = -\infty + k$ for all $k \in \mathbb{R}$.) As it is not possible that

$$v_{\text{I}}(S') \;<\; 0 \;<\; v_{\text{II}}(S')\,,$$

it is also not possible that

$$v_{\text{I}}(S) \;<\; -k \;<\; v_{\text{II}}(S)\,.$$

Thus we cannot have $v_{\text{I}}(S) < v_{\text{II}}(S)$ since $k$ is arbitrary. But we have already seen that $v_{\text{I}}(S) \le v_{\text{II}}(S)$. Therefore

$$v_{\text{I}}(S) \;=\; v_{\text{II}}(S)\,.$$

□

The common value $v = v_{\text{I}} = v_{\text{II}}$ is called the *value* of the semi-infinite game. It should be noted that, different from finite games, the existence of the value of a semi-infinite game does not necessarily imply the existence of an equilibrium pair of mixed strategies: for, there may not exist such a mixed strategy for player II that gives him the infimum value $v_{\text{II}}$.

# Chapter 5

# Applications in Utility Theory

In this chapter, we shall deal with the topic of lexicographic utility theory. In Section 5.1 we give a modified version of lexicographic expected utility theory, taking only rational-valued probabilities into account (Theorem 5.1.2). In Section 5.2 we consider the problem of extensive measurement on discrete spaces. We obtain a lexicographical representation theorem for preferences on indivisible items, giving a necessary and sufficient condition for extensive utility representation (Theorem 5.2.1). These results will be derived as consequences of the lexicographical separation theorem in Chapter 3.

## 5.1 Lexicographic Expected Utility on Rational-Valued Lotteries

In this section we shall show that a similar form of the lexicographic expected utility representation considered by Hausner [17] is valid even if we omit the existence of irrational-valued lotteries.

The well-known expected utility theory for decision making under risk, as developed by von Neumann and Morgenstern [45], is based on the assumptions that a preference relation over the lotteries is a complete preorder satisfying an independence condition and a continuity property. The role of the continuity property in single person decision theory is merely technical, namely, to ensure the existence of a real-valued utility function. As Hausner [17] pointed out, dropping this assumption modifies the classical expected utility theory by allowing lexicographic utility functions.

Hausner derived the lexicographic representation of a utility space by embedding the utility space into an ordered vector space over $\mathbb{R}$; and by using the fact, as was proved by Hausner and Wendel [18], that any ordered vector space over $\mathbb{R}$ can be represented as a lexicographic function space. His derivation also suggests that, when a utility space is finitely generated, the dimension of a lexicographic utility function is less than that of the utility space. Finite-dimensionality of a utility function was first axiomatized by Fishburn [7] [10], who introduced a hierarchical axiom which ensures the existence of a finite-dimensinal utility function lexicographically preserving an order on a mixture space. Recently, Nakamura [33] [34] presents a preference-based hierarchical axiom as an alternative to Fishburn's.

All these results are founded on the domain of lotteries, that is, the domain of real-valued probability distributions on an outcome set. In this thesis we restrict our attention to rational-valued probabilities (in order not to consider situations in which an event occurs with an irrational-valued probability), and show that the lexicographic expected utility can also be founded on the domain of rational-valued lotteries.

Throughout this section we shall assume that $S_n = \{\alpha_1, \ldots, \alpha_n\}$ is a finite set of items for some $n > 1$. A *lottery* is a probability distribution on $S_n$ and is denoted by $p_1\,\alpha_1 + \cdots + p_n\,\alpha_n$ where $p_i \in \mathbb{R}$ is a probability of $\alpha_i$, so that $p_i \geq 0$ and $\sum p_i = 1$. We denote by $\Delta^*(S_n)$ the set of all lotteries. Given $x, y \in \Delta^*(S_n)$ and $\theta \in [0,1]$, the *convex combination* $\theta x + (1-\theta)y$ is well defined (see Section 1.1) and also belongs to $\Delta^*(S_n)$.

A weak order $\prec$ on $\Delta^*(S_n)$ satisfies the *independence condition* if $x \prec y$ implies $\theta x + (1-\theta)z \prec \theta y + (1-\theta)z$ for all $z \in \Delta^*(S_n)$ and all $\theta \in (0,1)$. The independence condition for the associated indifference relation $\sim$ is defined similarly. As Hausner [17] and Fishburn [9] pointed out, these independence conditions for a weak order $\prec$ on $\Delta^*(S_n)$ are necessary and sufficient for the existence of a finite-dimensional linear utility function on $\Delta^*(S_n)$ whose lexicographic order preserves the ordering $\prec$:

**Proposition 5.1.1** (Hausner [17], Fishburn [9]) Let $\prec$ be a weak order on $\Delta^*(S_n)$ and let $\sim$ be the associated indifference relation. Then $\prec$ and $\sim$ satisfy the independence conditions if and only if there exists an $(n-1)$-dimensional utility function $U = (u_1, \ldots, u_{n-1})$ on $\Delta^*(S_n)$ such that for all $x, y \in \Delta^*(S_n)$,

$$x \prec y \quad \text{iff} \quad U(x) <_L U(y), \tag{5.1}$$

$$U(\theta x + (1-\theta)y) = \theta\, U(x) + (1-\theta)\, U(y) \qquad \text{for all } \theta \in [0,1]. \tag{5.2}$$

□

That is to say, the assumptions of Proposition 5.1.1 ensure the existence of vector-valued utilities $\boldsymbol{a_1}, \ldots, \boldsymbol{a_n} \in \mathbb{R}^{n-1}$ of the items $\alpha_1, \ldots, \alpha_n \in S_n$, respectively, whose expected value is to be maximized lexicographically: for, if we have a utility function $U : \Delta^*(S_n) \to \mathbb{R}^{n-1}$ satisfying (5.1) and (5.2) then the vector-valued utility $\boldsymbol{a_i}$ of each item $\alpha_i$ can be defined by $\boldsymbol{a_i} := U(\alpha_i)$ for $i = 1, \ldots, n$ (where $\alpha_i$ denotes the lottery in which $\alpha_i$ occurs with probability one), so that for all lotteries $\sum_{i=1}^n \lambda_i \alpha_i$, $\sum_{i=1}^n \mu_i \alpha_i$ in $\Delta^*(S_n)$,

$$\sum_{i=1}^n \lambda_i \alpha_i \prec \sum_{i=1}^n \mu_i \alpha_i \quad \text{iff} \quad \sum_{i=1}^n \lambda_i\, \boldsymbol{a_i} <_L \sum_{i=1}^n \mu_i\, \boldsymbol{a_i}. \tag{5.3}$$

It is also known that the vector-valued utilities can be replaced by single-valued utilities by assuming a suitable Archimedean condition.

It should be pointed out that, from the perspective of Proposition 5.1.1, the existence of such utilities are ultimately founded on the domain of lotteries $\Delta^*(S_n)$; hence it is tacitly assumed that the domain of lotteries includes "irrational-valued lotteries," such as $\frac{\sqrt{2}}{3}\,\alpha_1 + (1 - \frac{\sqrt{2}}{3})\,\alpha_2$, which assign irrational-valued probabilities to some items. In our real life, however, it is difficult at once to imagine such a gamble in which an event occurs with an irrational-valued probability.[1] Thus we will restrict ourselves to the domain of "rational-valued lotteries," that is, lotteries in which all events occur with rational-valued

[1] Kaneko [22] presented a repetition method which yields an irrational-valued probability. But this method is not possible to carry out when the irrational-valued probability is not calculable.

probabilities, and consider axioms for weak orders on the domain of such rational-valued lotteries.

A *rational lottery* is a lottery $p_1\,\alpha_1+\cdots+p_n\,\alpha_n$ in which all probabilities are rational-valued, that is, $p_i \in \mathbb{Q}$ and $p_i \geq 0$ for all $1 \leq i \leq n$, with $\sum_{i=1}^{n} p_i = 1$. We denote by $\Delta(S_n)$ the set of all rational lotteries. It is easy to verify that, for all $x, y \in \Delta(S_n)$ and all $0 \leq \theta \leq 1$ in $\mathbb{Q}$, the convex combination $\theta x + (1-\theta)y$ also belongs to $\Delta(S_n)$.

Let $\prec$ be a weak order on $\Delta(S_n)$, and let $\sim$ be its associated indifference relation. The independence conditions for $\prec$ and $\sim$ on $\Delta(S_n)$ is defined in a parallel way to $\Delta^*(S_n)$: for all $x, y, z \in \Delta(S_n)$ and all $0 < \theta < 1$ in $\mathbb{Q}$,

I1 if $x \prec y$ then $\theta x + (1-\theta)z \;\prec\; \theta y + (1-\theta)z$,

I2 if $x \sim y$ then $\theta x + (1-\theta)z \;\sim\; \theta y + (1-\theta)z$.

Now we state the main result of this section as follows, which is just a counterpart of Proposition 5.1.1 on the domain of rational lotteries:

**Theorem 5.1.2** *Let $\prec$ be a nontrivial weak order on $\Delta(S_n)$ and let $\sim$ be the associated indifference relation. Then $\prec$ and $\sim$ satisfy the independence conditions (I1 and I2, respectively) if and only if there are real-valued functions $u_1, \ldots, u_{n-1}$ on $\Delta(S_n)$, where $u_1$ is not a constant function, such that for all $x, y \in \Delta(S_n)$,*

$$x \prec y \qquad \text{iff} \qquad (\,u_1(x), \ldots, u_{n-1}(x)\,) \;<_L\; (\,u_1(y), \ldots, u_{n-1}(y)\,) \tag{5.4}$$

*and, for each $1 \leq j \leq n-1$ and all $0 \leq \theta \leq 1$ in $\mathbb{Q}$,*

$$u_j(\,\theta x + (1-\theta)y\,) \;=\; \theta\, u_j(x) \;+\; (1-\theta)\, u_j(y)\,. \tag{5.5}$$

*Moreover, $u_1$ is unique up to a positive affine transformation,*[2] *that is, if $v_1, \ldots, v_{n-1}$ are other functions with the same property, then there exist real constants $c > 0$ and $d$ such that*

$$v_1(x) \;=\; c\, u_1(x) + d \quad \text{for all } x \in \Delta(S_n)\,.$$

□

Thus, by a similar argument below Proposition 5.1.1, the assumptions of Theorem 5.1.2 also ensure the existence of vector-valued utilities $\boldsymbol{a_1}, \ldots, \boldsymbol{a_n} \in \mathbb{R}^{n-1}$ of the items $\alpha_1, \ldots, \alpha_n \in S_n$, respectively, satisfying (5.3) for all $\sum_{i=1}^{n} \lambda_i \alpha_i$, $\sum_{i=1}^{n} \mu_i \alpha_i$ in $\Delta(S_n)$. In other words, when all the items are guaranteed to occur with rational-valued probabilities, we do not need to assume the existence of "irrational-valued lotteries."

**Remark 5.1.3** It should be noted that the assumptions of Theorem 5.1.2 do not ensure the existence of *rational-valued* functions $u_1, \ldots, u_{n-1}$ on $\Delta(S_n)$ satisfying (5.4) and (5.5). Here we give an illustrative example: Let $\prec$ be a binary relation on $\Delta(S_3)$ defined by

$$p_1\,\alpha_1 + p_2\,\alpha_2 + p_3\,\alpha_3 \;\prec\; q_1\,\alpha_1 + q_2\,\alpha_2 + q_3\,\alpha_3 \quad \text{iff} \quad -p_1 + \sqrt{2}\,p_3 \;<\; -q_1 + \sqrt{2}\,q_3\,. \tag{5.6}$$

[2] The other functions $u_2, \ldots, u_{n-1}$ are entirely indefinite in almost all cases.

It is easy to check that $\prec$ is a weak order on $\Delta(S_3)$ satisfying the independent conditions. Also

$$\begin{aligned} \alpha_2 &\prec p\,\alpha_1 + (1-p)\,\alpha_3 \qquad \text{if} \quad 0 \le p < \tfrac{\sqrt{2}}{\sqrt{2}+1}\,, \\ \alpha_2 &\succ p\,\alpha_1 + (1-p)\,\alpha_3 \qquad \text{if} \quad \tfrac{\sqrt{2}}{\sqrt{2}+1} < p \le 1\,. \end{aligned} \tag{5.7}$$

Suppose there exist rational-valued functions $u_1, \ldots, u_{n-1}$ on $\Delta(S_n)$ satisfying (5.4) and (5.5). This means that there exist utilities $(a_1, a_2), (b_1, b_2), (c_1, c_2) \in \mathbb{Q}^2$ of the items $\alpha_1, \alpha_2, \alpha_3 \in S_3$, respectively, satisfying

$$\begin{aligned} &p_1\,\alpha_1 + p_2\,\alpha_2 + p_3\,\alpha_3 \prec q_1\,\alpha_1 + q_2\,\alpha_2 + q_3\,\alpha_3 \quad \text{iff} \\ &\quad p_1\,(a_1, a_2) + p_2\,(b_1, b_2) + p_3\,(c_1, c_2) <_L q_1\,(a_1, a_2) + q_2\,(b_1, b_2) + q_3\,(c_1, c_2)\,. \end{aligned} \tag{5.8}$$

Then we get a contradiction as follows. By (5.6), we have $(a_1, a_2) <_L (b_1, b_2) <_L (c_1, c_2)$. First we shall show that $a_1 = b_1 = c_1$. Assume $a_1 < c_1$. Then, $b_1 \neq \frac{\sqrt{2}}{\sqrt{2}+1}\,a_1 + \frac{1}{\sqrt{2}+1}\,c_1$ because the right-hand side is an irrational number. Consider the case $b_1 < \frac{\sqrt{2}}{\sqrt{2}+1}\,a_1 + \frac{1}{\sqrt{2}+1}\,c_1$. By the density of $\mathbb{Q}$ in $\mathbb{R}$, there is a $d_1 \in \mathbb{Q}$ such that

$$b_1 < d_1\,, \tag{5.9}$$

$$d_1 < \tfrac{\sqrt{2}}{\sqrt{2}+1}\,a_1 + \tfrac{1}{\sqrt{2}+1}\,c_1\,. \tag{5.10}$$

Note that $a_1 \le b_1 \le d_1 \le c_1$. By (5.9), we have $(b_1, b_2) <_L (d_1, d_2)$ for any $d_2 \in \mathbb{Q}$, and hence $(b_1, b_2) <_L \frac{c_1-d_1}{c_1-a_1}\,(a_1, a_2) + \frac{d_1-a_1}{c_1-a_1}\,(c_1, c_2)$. Therefore, by (5.8), $\alpha_2 \prec \frac{c_1-d_1}{c_1-a_1}\,\alpha_1 + \frac{d_1-a_1}{c_1-a_1}\,\alpha_3$, and hence, by (5.7), we must have $0 \le \frac{c_1-d_1}{c_1-a_1} < \frac{\sqrt{2}}{\sqrt{2}+1}$. However, by (5.10), we obtain $\frac{\sqrt{2}}{\sqrt{2}+1} < \frac{c_1-d_1}{c_1-a_1}$, a contradiction. The other case $b_1 > \frac{\sqrt{2}}{\sqrt{2}+1}\,a_1 + \frac{1}{\sqrt{2}+1}\,c_1$ can be treated in a parallel way with the above. Thus we have shown that $c_1 = b_1 = a_1$. Similarly we can show that $c_2 = b_2 = a_2$. But this contradicts the hypothesis $(a_1, a_2) <_L (b_1, b_2) <_L (c_1, c_2)$. □

The "if" part of Theorem 5.1.2 is easy to check, so we omit the proof. The "only if" part and the uniqueness part of Theorem 5.1.2 will be proved by a sequence of lemmas.

Hausner [17] derived the lexicographic representation of utility spaces by the following method: he first show that any utility space can be embedded into an ordered vector space over $\mathbb{R}$, and then he uses the fact, which was proved by Hausner and Wendel [18], that any ordered vector space over $\mathbb{R}$ can be represented as a lexicographic function space. Following Hausner, we first show that any weak order on $\Delta(S_n)$ can be embedded into an ordered vector space over $\mathbb{Q}$. Let

$$V(S_n) = \{\, \lambda_1\alpha_1 + \cdots + \lambda_n\alpha_n \mid \lambda_1, \ldots, \lambda_n \in \mathbb{Q} \,\}\,,$$

$$H(S_n) = \{\, \lambda_1\alpha_1 + \cdots + \lambda_n\alpha_n \mid \lambda_1, \ldots, \lambda_n \in \mathbb{Q} \text{ and } \textstyle\sum_{i=1}^{n} \lambda_i = 0 \,\}\,.$$

$V(S_n)$ is the $n$-dimensional vector space over $\mathbb{Q}$ which is generated by $\{\alpha_1, \ldots, \alpha_n\}$, and $H(S_n)$ is an $(n-1)$-dimensional subspace of $V(S_n)$. Also, $\Delta(S_n)$ is naturally embedded in $V(S_n)$. It is easy to verify that, for any $x, y \in \Delta(S_n)$, we have $y - x \in H(S_n)$.

**Lemma 5.1.4** *For any $z \in H(S_n)$, there exist $\lambda > 0$ and $x, y \in \Delta(S_n)$ such that $z = \lambda(y - x)$.*

**Proof.** Suppose $z$ is in $H(S_n)$, and let $x = \frac{1}{n}\alpha_1 + \cdots + \frac{1}{n}\alpha_n \in \Delta(S_n)$ be fixed. Let $y = x + \mu z$ for some $\mu \in \mathbb{Q}$ with $\mu > 0$. Then, if $\mu$ is sufficiently small, we have $y \in \Delta(S_n)$. Let $\lambda = 1/\mu$, and we get the conclusion. □

Let $L(S_n) = \{\lambda\alpha_1 + \cdots + \lambda\alpha_n \mid \lambda \in \mathbb{Q}\}$. Then $L(S_n)$ is a 1-dimensional subspace of $V(S_n)$. The following lemma shows that $V(S_n)$ is the direct sum of $H(S_n)$ and $L(S_n)$:

**Lemma 5.1.5** *For any $x \in V(S_n)$, there are unique $x_H \in H(S_n)$ and $x_L \in L(S_n)$ such that $x = x_H + x_L$.*

**Proof.** Let $x = \lambda_1\alpha_1 + \cdots + \lambda_n\alpha_n \in V(S_n)$, and let $\lambda = \sum_{i=1}^{n} \lambda_i$. We define $x_L \in L(S_n)$ as $x_L = \frac{\lambda}{n}\alpha_1 + \cdots + \frac{\lambda}{n}\alpha_n$. Let $x_H = x - x_L$. Then $x_H$ belongs to $H(S_n)$. The uniqueness of $x_L$ and $x_H$ is obvious. □

By Lemma 5.1.5, we can define the *projection* $p_H$ from $V(S_n)$ to $H(S_n)$ as $p_H(x) = x_H$ for all $x \in V(S_n)$. It is easy to check that $p_H$ is a linear mapping.

**Lemma 5.1.6** *Let $\prec$ be a weak order on $\Delta(S_n)$. Then, for all $x, y, z \in \Delta(S_n)$, $x \precsim y$ and $y \prec z$ imply $x \prec z$.*

**Proof.** By Proposition 2.1.1 (b), (d). □

**Lemma 5.1.7** *Let $\prec$ be a weak order on $\Delta(S_n)$ satisfying I1 and I2. Then, for all $x, y, z, w \in \Delta(S_n)$ and all $0 < \theta < 1$ in $\mathbb{Q}$, $x \precsim y$ and $w \prec z$ imply $\theta x + (1-\theta)w \prec \theta y + (1-\theta)z$.*

**Proof.** Suppose $x \precsim y$ and $w \prec z$. Then, by I1 and I2, $x \precsim y$ implies $\theta x + (1-\theta)w \precsim \theta y + (1-\theta)w$. By I1, $w \prec z$ implies $\theta y + (1-\theta)w \prec \theta y + (1-\theta)z$. Hence, by Lemma 5.1.6, we get $\theta x + (1-\theta)w \prec \theta y + (1-\theta)z$. □

**Lemma 5.1.8** *Let $\prec$ be a weak order on $\Delta(S_n)$ satisfying I1 and I2. Then, for all $x, y, z, w \in \Delta(S_n)$ and all $0 < \theta < 1$ in $\mathbb{Q}$, not$(y \prec x)$ and not$(z \prec w)$ imply not$(\theta y + (1-\theta)z \prec \theta x + (1-\theta)w)$.*

**Proof.** By Proposition 2.1.1 (a), it is enough to show that, for all $x, y, z, w \in \Delta(S_n)$ and all $0 < \theta < 1$ in $\mathbb{Q}$, $x \precsim y$ and $w \precsim z$ imply $\theta x + (1-\theta)w \precsim \theta y + (1-\theta)z$. If either $x \prec y$ or $w \prec z$ (or both) holds, then this is clear from Lemma 5.1.7. Suppose $x \sim y$ and $w \sim z$. Then, by I2, we have $\theta x + (1-\theta)w \sim \theta y + (1-\theta)w \sim \theta y + (1-\theta)z$. □

**Lemma 5.1.9** *Suppose $\prec$ is a weak order on $\Delta(S_n)$ satisfying I1 and I2. Let*

$$P = \{\lambda(y-x) \in H(S_n) \mid \lambda \in \mathbb{Q},\ x, y \in \Delta(S_n) \text{ with } \lambda > 0 \text{ and } x \prec y\}.$$

*Then, for all $x, y \in \Delta(S_n)$, $x \prec y$ iff $y - x \in P$.*

**Proof.** Recall that if $x, y \in \Delta(S_n)$ then $y - x \in H(S_n)$. Hence, if $x \prec y$ then, by the definition, $y - x \in P$. For the converse, suppose $y - x \in P$ and not$(x \prec y)$. By the definition of $P$, there are $\lambda > 0$ and $z, w \in \Delta(S_n)$ with $z \prec w$ such that $y - x = \lambda(w - z)$. Let $\theta = 1/(1+\lambda)$. Then we have $0 < \theta < 1$ and $\theta y + (1-\theta)z = \theta x + (1-\theta)w$. Now, not$(x \prec y)$ yields $y \precsim x$. Hence, by Lemma 5.1.7, $y \precsim x$ and $z \prec w$ imply $\theta y + (1-\theta)z \prec \theta x + (1-\theta)w$ $(= \theta y + (1-\theta)z)$. This is impossible because $\prec$ is asymmetric and hence there is no $v \in \Delta(S_n)$ such that $v \prec v$. □

**Lemma 5.1.10** *Suppose $\prec$ is a nontrivial weak order on $\Delta(S_n)$ satisfying I1 and I2, and let $P$ be as in Lemma 5.1.9. Then, $P$ is a nonempty positive cone in $H(S_n)$, and $\bar{P}$ is a convex cone in $H(S_n)$.*

**Proof.** ($P$ is a nonempty positive cone in $H(S_n)$.) Since $\prec$ is nontrivial, it is easy to see that $P$ is nonempty. Also $P$ does not contain the origin, because $\prec$ is asymmetric and hence there is no $x \in \Delta(S_n)$ such that $x \prec x$. It remains to show that $P$ is a convex cone. It is clear, from the definition of $P$, that $x \in P$ and $\lambda > 0$ imply $\lambda x \in P$. Now we shall show that $x, y \in P$ implies $x+y \in P$. Let $x, y \in P$. Then, by the definition of $P$, there are $\lambda, \mu > 0$ and $z, w, u, v \in \Delta(S_n)$ with $z \prec w$ and $u \prec v$ such that $x = \lambda(w-z)$ and $y = \mu(v-u)$. Let $\theta = \lambda/(\lambda+\mu)$. Then we have $0 < \theta < 1$. Hence, by Lemma 5.1.7, $z \prec w$ and $u \prec v$ yield $\theta z+(1-\theta)u \prec \theta w+(1-\theta)v$. Therefore $\theta(w-z)+(1-\theta)(v-u) \in P$. This means $\frac{1}{\lambda+\mu}(x+y) = \frac{\lambda}{\lambda+\mu}(w-z)+\frac{\mu}{\lambda+\mu}(v-u) = \theta(w-z)+(1-\theta)(v-u) \in P$; thus we obtain $x+y \in P$.

($\bar{P}$ is a convex cone in $H(S_n)$.) First we shall show that $x \in \bar{P}$ and $\lambda > 0$ imply $\lambda x \in \bar{P}$. Suppose $\lambda x \notin \bar{P}$ and $\lambda > 0$. This means $\lambda x \in P$ and $1/\lambda > 0$. Hence $x \in P$ (since $P$ is a convex cone). Therefore $x \notin \bar{P}$. Now we shall show that $x, y \in \bar{P}$ implies $x+y \in \bar{P}$. Let $x, y \in \bar{P}$. Then, by Lemma 5.1.4, there are $\lambda, \mu > 0$ and $z, w, u, v \in \Delta(S_n)$ such that $x = \lambda(w-z)$ and $y = \mu(v-u)$. Moreover, not ($z \prec w$) and not ($u \prec v$): for otherwise, by the definition of $P$, we have $x \in P$ or $y \in P$, a contradiction. Let $\theta = \lambda/(\lambda+\mu)$. Then we have $0 < \theta < 1$. Hence, by Lemma 5.1.8, not ($z \prec w$) and not ($u \prec v$) yield not ($\theta z + (1-\theta)u \prec \theta w + (1-\theta)v$). By Lemma 5.1.9, we obtain $\theta(w-z) + (1-\theta)(v-u) \in \bar{P}$. Since $\frac{1}{\lambda+\mu}(x+y) = \frac{\lambda}{\lambda+\mu}(w-z) + \frac{\mu}{\lambda+\mu}(v-u) = \theta(w-z)+(1-\theta)(v-u) \in \bar{P}$, we get $x+y \in \bar{P}$. □

Now we are in a position to prove Theorem 5.1.2 by way of the lexicographical separation theorem (Theorem 3.1.1).

**Proof of the "only if" part of Theorem 5.1.2 :** Suppose $\prec$ is a nontrivial weak order on $\Delta(S_n)$ satisfying I1 and I2. Let $P$ be defined as in Lemma 5.1.9. Then, by Lemma 5.1.10, $P$ is a nonempty positive cone in $H(S_n)$, and also $\bar{P}$ is a convex cone in $H(S_n)$. By Theorem 3.1.1, there are real-valued linear functions $g_1, \ldots, g_{n-1}$ on $H(S_n)$, where $g_1$ is not constantly zero, such that for all $x \in H(S_n)$,

$$x \in P \quad \text{iff} \quad (g_1(x), \ldots, g_{n-1}(x)) >_L (0, \ldots, 0). \tag{5.11}$$

Let $p_H$ be the projection from $V(S_n)$ to $H(S_n)$ (as was defined below Lemma 5.1.5). Note that $p_H(x) = x$ for all $x \in H(S_n)$, and the composite functions $g_1 \circ p_H, \ldots, g_{n-1} \circ p_H$ are real-valued linear functions on $V(S_n)$. We see that $g_1 \circ p_H, \ldots, g_{n-1} \circ p_H$ satisfy (5.4) for all $x, y \in \Delta(S_n)$:

$$\begin{aligned}
& y \succ x \\
\Longleftrightarrow\ & y - x \in P && \text{(by Lemma 5.1.9)} \\
\Longleftrightarrow\ & (g_1(y-x), \ldots, g_{n-1}(y-x)) >_L (0, \ldots, 0) && \text{(by (5.11))} \\
\Longleftrightarrow\ & (g_1 \circ p_H(y-x), \ldots, g_{n-1} \circ p_H(y-x)) >_L (0, \ldots, 0) \\
\Longleftrightarrow\ & (g_1 \circ p_H(y), \ldots, g_{n-1} \circ p_H(y)) >_L (g_1 \circ p_H(x), \ldots, g_{n-1} \circ p_H(x))
\end{aligned}$$

Also $g_1 \circ p_H$ is not constant on $\Delta(S_n)$: for, if $g_1 \circ p_H(x) = g_1 \circ p_H(y)$ for all $x, y \in \Delta(S_n)$, $g_1 \circ p_H(y-x) = g_1(y-x) = 0$ for all $x, y \in \Delta(S_n)$, and hence, by Lemma 5.1.4,

$g_1(z) = 0$ for all $z \in H(S_n)$, contradicting the hypothesis that $g_1$ is not constantly zero. Since $g_j \circ p_H$ is linear on $V(S_n)$, it satisfies (5.5) for all $x, y \in \Delta(S_n)$ and all $0 \leq \theta \leq 1$ in $\mathbb{Q}$. Thus the restrictions of $g_1 \circ p_H, \ldots, g_{n-1} \circ p_H$ to $\Delta(S_n)$ are desired solutions of the existence part of Theorem 5.1.2.

**Proof of the uniqueness part of Theorem 5.1.2 :** Suppose $u_1, \ldots, u_{n-1}$ are real-valued functions on $\Delta(S_n)$ satisfying (5.4) and (5.5) for all $x, y \in \Delta(S_n)$ and all $0 \leq \theta \leq 1$ in $\mathbb{Q}$. We also assume that $u_1$ is not constant on $\Delta(S_n)$. Let us define linear functions $g_1, \ldots, g_{n-1}$ on $V(S_n)$ as follows: for $j = 1, \ldots, n-1$,

$$g_j(\lambda_1\alpha_1 + \cdots + \lambda_n\alpha_n) := \lambda_1 u_j(\alpha_1) + \cdots + \lambda_n u_j(\alpha_n) \quad \text{for all } \lambda_1, \ldots, \lambda_n \in \mathbb{Q}. \quad (5.12)$$

Then $g_j(x) = u_j(x)$ for all $x \in \Delta(S_n)$; that is, $g_j$ is a (unique) linear extension of $u_j$ to $V(S_n)$. We can see that $g_1$ is not constantly zero on $H(S_n)$: for, by the assumption that $u_1$ is not constant on $\Delta(S_n)$, there must be some $\alpha_k$ and $\alpha_l$ such that $u_1(\alpha_k) < u_1(\alpha_l)$, hence $\alpha_l - \alpha_k \in H(S_n)$ and $g_1(\alpha_l - \alpha_k) = u_1(\alpha_l) - u_1(\alpha_k) > 0$.

Let $P$ be defined as in Lemma 5.1.9. We shall show that $g_1, \ldots, g_{n-1}$ satisfy (5.11) for all $x \in H(S_n)$. First suppose $x \in P$; we will show $(g_1(x), \ldots, g_{n-1}(x)) >_L (0, \ldots, 0)$. By the definition of $P$, there are $\lambda > 0$ and $y, z \in \Delta(S_n)$ with $y \prec z$ such that $x = \lambda(z-y)$. Then $g_j(x) = \lambda g_j(z) - \lambda g_j(y) = \lambda u_j(z) - \lambda u_j(y)$ for $j = 1, \ldots, n-1$, and also $(u_1(y), \ldots, u_{n-1}(y)) <_L (u_1(z), \ldots, u_{n-1}(z))$ by the assumption (5.4). Therefore

$$\begin{aligned} (g_1(x), \ldots, g_{n-1}(x)) &= \lambda[(u_1(z), \ldots, u_{n-1}(z)) - (u_1(y), \ldots, u_{n-1}(y))] \\ &>_L (0, \ldots, 0). \end{aligned}$$

Suppose $x \notin P$, (that is, $x \in H(S_n) \setminus P$). Then, by Lemma 5.1.4, there are $\lambda > 0$ and $y, z \in \Delta(S_n)$ such that $x = \lambda(z-y)$. Moreover, not $(y \prec z)$ (for otherwise $x$ must be in $P$). Therefore, a similar argument to the above leads to $(g_1(x), \ldots, g_{n-1}(x)) \leqq_L (0, \ldots, 0)$. Thus we have shown that the subset $P$ of $H(S_n)$ and the restrictions of $g_1, \ldots, g_{n-1}$ to $H(S_n)$ satisfy the conditions of the existence part of Theorem 3.1.1.

Suppose there are other functions $v_1, \ldots, v_{n-1}$ on $\Delta(S_n)$ with the same property as $u_1, \ldots, u_{n-1}$. We define linear functions $g'_1, \ldots, g'_{n-1}$ on $V(S_n)$ in a similar way to (5.12) for $j = 1, \ldots, n-1$; hence $g'_j(x) = v_j(x)$ for all $x \in \Delta(S_n)$. By a similar argument to the above, the restrictions of $g'_1, \ldots, g'_{n-1}$ to $H(S_n)$ also satisfy the conditions of the existence part of Theorem 3.1.1. Hence, by the uniqueness part of Theorem 3.1.1, there is a real constant $a > 0$ such that $g'_1(x) = a\, g_1(x)$ for all $x \in H(S_n)$. This means $g'_1(y-z) = a\, g_1(y-z)$ for all $y, z \in \Delta(S_n)$. Therefore, for all $y, z \in \Delta(S_n)$,

$$\begin{aligned} v_1(y) - v_1(z) = g'_1(y) - g'_1(z) &= g'_1(y-z) \\ &= a\, g_1(y-z) \\ &= a\, g_1(y) - a\, g_1(z) = a\, u_1(y) - a\, u_1(z). \end{aligned}$$

Let $z$ be fixed for some $z_0 \in \Delta(S_n)$, and let $b = v_1(z_0) - a\, u_1(z_0)$. Then, for all $y \in \Delta(S_n)$,

$$v_1(y) = a\, u_1(y) + [v_1(z_0) - a\, u_1(z_0)] = a\, u_1(y) + b.$$

**Q.E.D.**

**Remark 5.1.11** Let $\mathbb{F}$ be an ordered field such that $\mathbb{Q} \subseteq \mathbb{F} \subseteq \mathbb{R}$. (For example, $\mathbb{F}$ may be the field of algebraic numbers.) It is easy to check that the above proof remains valid even if $\mathbb{Q}$ is replaced by $\mathbb{F}$. Therefore the lexicographic expected utility representation can also be founded on the domain of $\mathbb{F}$-valued lotteries. □

## 5.2 Extensive Utility on Indivisible Items

In this section, we shall give a necessary and sufficient condition for lexicographic extensive measurement. ( See Section 1.1 for the notion of "extensive measurement." ) . Our formulation will be suitable to describe a preference on indivisible items.

Throughout this section we assume that $S_n = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is a finite set of indivisible items for some $n \geq 1$. Let $\mathbb{N}$ be the set of all natural numbers (including zero). A *consumption plan* of the items is denoted by $k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n$ with $k_1, k_2, \ldots, k_n \in \mathbb{N}$, where each natural number $k_i$ is a consumption of $\alpha_i$. A consumption plan is also denoted by $\sum_{i=1}^{n} k_i\alpha_i$. ( Intuitively, the expression $k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n$ means that a consumer consumes $k_i$ pieces of $\alpha_i$ for each $i = 1, \ldots, n$.) We denote by $\Omega(S_n)$ the set of all consumption plans, i.e.

$$\Omega(S_n) \;=\; \{\, k_1\alpha_1 \,+\, k_2\alpha_2 \,+\, \cdots \,+\, k_n\alpha_n \mid k_1, k_2 \cdots, k_n \in \mathbb{N} \,\} .$$

The additive operation + is naturally applied to consumption plans :

$$(k_1\alpha_1 + \cdots + k_n\alpha_n) \,+\, (l_1\alpha_1 + \cdots + l_n\alpha_n) \;\stackrel{\text{def}}{=}\; (k_1 + l_1)\alpha_1 \,+\, \cdots \,+\, (k_n + l_n)\alpha_n .$$

[ This is the reason why we use the additive notation $k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n$ to denote a consumption plan, instead of $(k_1, k_2, \ldots, k_n)$ . ] Let 0 denote $0\alpha_1 + 0\alpha_2 + \cdots + 0\alpha_n$. Thus we established the system $\langle\, \Omega(S_n) \,,\, + \,,\, 0 \,\rangle$, which can be considered as the nonnegative part of an Abelian group.

Let $\prec$ be a preference relation on the set of consumption plans $\Omega(S_n)$, where $x \prec y$ means "$y$ is preferred to $x$." Consider the following conditions :

A1 $\quad \prec$ is a weak order on $\Omega(S_n)$,

A2 $\quad$ for all $x, y, z \in \Omega(S_n)$, $\quad x \prec y \quad$ implies $\quad x + z \;\prec\; y + z$,

A3 $\quad$ for all $x, y, z \in \Omega(S_n)$, $\quad x \sim y \quad$ implies $\quad x + z \;\sim\; y + z$.

( See Section 2.1 for the definitions of a weak order $\prec$ and its associated indifference relation $\sim$.)

The condition A1 means that a preference relation is an ordering. The monotonicity condition A2 asserts that a preference relation should be preserved by adding the same items to both sides of the relation. For example, suppose there are two kinds of items $(\alpha_1, \alpha_2) =$ ( orange , apple ) and also we have a preference

$$2 \cdot \text{orange} \;\prec\; 1 \cdot \text{apple} \;,$$

meaning that we like an apple better than two oranges. Then, by A2, the following preferences are inferred :

$$\begin{array}{llcl} \Rightarrow & 3 \cdot \text{orange} & \prec & 1 \cdot \text{apple} \,+\, 1 \cdot \text{orange}\,, \\ \Rightarrow & 1 \cdot \text{apple} \,+\, 3 \cdot \text{orange} & \prec & 2 \cdot \text{apple} \,+\, 1 \cdot \text{orange}\,, \\ \Rightarrow & & \cdots & \end{array}$$

by adding an orange to both sides, then adding an apple, and so on. The meaning of A3 can be described in a similar way.

The following theorem states that the scheme of conditions A1–A3 is necessary and sufficient for extensive measurement of preferences. ( See Remark 5.2.3 for the verification that the following theorem actually gives extensive measurement ) .

**Theorem 5.2.1** *A preference relation $\prec$ on $\Omega(S_n)$ satisfies A1–A3 if and only if there are lexicographically ordered polynomials* [3] $q_1, \ldots, q_n \in \mathbb{R}[\varepsilon]$ *such that for all* $k_1, \ldots, k_n, l_1, \ldots, l_n \in \mathbb{N}$,

$$\sum_{i=1}^{n} k_i \alpha_i \prec \sum_{i=1}^{n} l_i \alpha_i \quad \text{iff} \quad \sum_{i=1}^{n} k_i q_i < \sum_{i=1}^{n} l_i q_i . \tag{5.13}$$

*Further,* $q_1, \ldots, q_n$ *can be written as the following form:*

$$\begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{pmatrix} = \begin{pmatrix} r_{11} & \cdots & r_{1n} \\ r_{21} & \cdots & r_{2n} \\ \vdots & \ddots & \vdots \\ r_{n1} & \cdots & r_{nn} \end{pmatrix} \begin{pmatrix} 1 \\ \varepsilon \\ \vdots \\ \varepsilon^{n-1} \end{pmatrix} \tag{5.14}$$

*where* $\{r_{ij}\}_{1 \le i,j \le n}$ *is an* $n \times n$ *matrix of real numbers such that either* (1) $r_{ij} = 0$ *for all* $i, j$, *or* (2) *the first column vector is nonzero, i.e.* $(r_{11}, \ldots, r_{n1})^T \neq (0, \ldots, 0)^T$.
*Moreover, the first column vector is unique up to a positive scalar multiple, i.e. if* $\{s_{ij}\}_{1 \le i,j \le n}$ *is another matrix with the same property, then there exists a positive real number* $a > 0$ *such that*

$$\begin{pmatrix} s_{11} \\ \vdots \\ s_{n1} \end{pmatrix} = a \begin{pmatrix} r_{11} \\ \vdots \\ r_{n1} \end{pmatrix} .$$

□

That is to say, the conditions A1–A3 ensure the existence of utilities $q_1, \ldots, q_n$ of the items $\alpha_1, \ldots, \alpha_n$, respectively, which can be added one another freely. The proof of this theorem will be given later.

It should be remarked that the utilities $q_1, \ldots, q_n$ obtained by Theorem 5.2.1 may not be real-valued, but may include an infinitely small number $\varepsilon$. [At first glance, it seems absurd to get such an infinitely small number $\varepsilon$ in our perspective. The readers who are not familiar with an infinitely small number should compare Remark 5.2.3 where the above theorem will be restated in terms of the lexicographic order on $\mathbb{R}^n$.] The following example tells us the reason why an infinitely small number $\varepsilon$ get into our consideration:

**Example 5.2.2** Let $(\alpha_1, \alpha_2) = (\text{sesame}, \text{diamond})$, and let $\prec$ be a preference relation defined by

$$k_1\alpha_1 + k_2\alpha_2 \prec l_1\alpha_1 + l_2\alpha_2 \quad \text{iff} \quad k_2 < l_2 \quad \text{or} \quad \text{“} k_2 = l_2 \text{ and } k_1 < l_1 . \text{”}$$

It is not difficult to verify that the above preference satisfies the conditions A1–A3. In essence, we have

$$0 \prec \alpha_1 \quad \text{and} \quad k\alpha_1 \prec \alpha_2 \quad \text{for all } k > 0 ,$$

where 0 means the zero consumption $0\,\alpha_1 + 0\,\alpha_2$, the item $\alpha_1$ is identified with $1\,\alpha_1 + 0\,\alpha_2$, and the item $\alpha_2$ is identified with $0\,\alpha_1 + 1\,\alpha_2$. That is to say, a grain of sesame is better than nothing, and a diamond is better than any finite grains of sesame. Such a

[3] To put it simply, $\mathbb{R}[\varepsilon]$ is the smallest ring containing both $\mathbb{R}$ and $\varepsilon$, where $\varepsilon$ is an “infinitely small number”, i.e. $0 < \varepsilon$ and $\varepsilon < 1/k$ for all positive integer $k$. See Definition 2.1.5

preference cannot be represented by real numbers extensively,[4] because there is no pair of real numbers $q_1, q_2$ such that $0 < q_1$ and $kq_1 < q_2$ for all $k > 0$. Nevertheless, by Theorem 5.2.1 there exist extensive utilities $q_1, q_2$ in $\mathbb{R}[\varepsilon]$: for example, $(q_1, q_2) = (\varepsilon\,,\, 1)$. In other words, a grain of sesame has infinitely small utility in comparison with a diamond. □

**Remark 5.2.3** The existence part of Theorem 5.2.1 can be stated also in the following way: *A preference relation $\prec$ on $\Omega(S_n)$ satisfies A1–A3 if and only if there exist real-valued functions $u_1, \ldots, u_n$ on $\Omega(S_n)$ such that for all $x, y \in \Omega(S_n)$,*

$$x \prec y \quad \text{iff} \quad (\, u_1(x), \ldots, u_n(x)\,) \;<_L\; (\, u_1(y), \ldots, u_n(y)\,)\,, \tag{5.15}$$

$$u_j(x+y) \;=\; u_j(x) \,+\, u_j(y) \quad \text{for each } j = 1, \ldots n\,. \tag{5.16}$$

(This is also a consequence of Hahn's Theorem. See Remark 3.1.6.) That is to say, the conditions A1–A3 for a preference relation $\prec$ on $\Omega(S_n)$ are necessary and sufficient for the existence of a finite-dimensional utility function on $\Omega(S_n)$ whose lexicographic order preserves the relation $\prec$; moreover, the utility function preserves the addition $+$. Therefore the scheme of conditions A1–A3 is necessary and sufficient for "lexicographical" extensive measurement of preferences.

Let us briefly check the equivalence of the existence part of Theorem 5.2.1 and the above statement. Suppose there exist $q_1, \ldots, q_n \in \mathbb{R}[\varepsilon]$ such that (5.13) holds for all $k_1, \ldots, k_n,\, l_1, \ldots, l_n \in \mathbb{N}$, where $q_1, \ldots, q_n$ are represented by an $n \times n$ real matrix $\{r_{ij}\}_{1 \le i,j \le n}$ of the form (5.14). Let the corresponding functions $u_1, \ldots, u_n$ on $\Omega(S_n)$ be defined as

$$u_j\,(k_1\alpha_1 + \cdots + k_n\alpha_n) \;:=\; k_1 r_{1j} + \cdots + k_n r_{nj} \qquad \text{for all } k_1, \ldots, k_n \in \mathbb{N}.$$

Then, one can easily verifies that for all $x, y \in \Omega(S_n)$,

$$u_j(x+y) \;=\; u_j(x) \,+\, u_j(y) \quad \text{for each } j = 1, \ldots n\,.$$

Also, for all $k_1, \ldots, k_n,\, l_1, \ldots, l_n \in \mathbb{N}$,

$$\begin{aligned}
& k_1\alpha_1 + \cdots + k_n\alpha_n \;\prec\; l_1\alpha_1 + \cdots + l_n\alpha_n \\
\iff\;& k_1 q_1 + \cdots + k_n q_n \;\prec\; l_1 q_1 + \cdots + l_n q_n \qquad \text{(by (5.13))} \\
\iff\;& (k_1, \ldots, k_n) \begin{pmatrix} r_{11} & \cdots & r_{1n} \\ r_{21} & \cdots & r_{2n} \\ \vdots & \ddots & \vdots \\ r_{n1} & \cdots & r_{nn} \end{pmatrix} \begin{pmatrix} 1 \\ \varepsilon \\ \vdots \\ \varepsilon^{n-1} \end{pmatrix} \\
& \qquad < \;(l_1, \ldots, l_n) \begin{pmatrix} r_{11} & \cdots & r_{1n} \\ r_{21} & \cdots & r_{2n} \\ \vdots & \ddots & \vdots \\ r_{n1} & \cdots & r_{nn} \end{pmatrix} \begin{pmatrix} 1 \\ \varepsilon \\ \vdots \\ \varepsilon^{n-1} \end{pmatrix} \\
\iff\;& (k_1, \ldots, k_n) \begin{pmatrix} r_{11} & \cdots & r_{1n} \\ r_{21} & \cdots & r_{2n} \\ \vdots & \ddots & \vdots \\ r_{n1} & \cdots & r_{nn} \end{pmatrix} <_L (l_1, \ldots, l_n) \begin{pmatrix} r_{11} & \cdots & r_{1n} \\ r_{21} & \cdots & r_{2n} \\ \vdots & \ddots & \vdots \\ r_{n1} & \cdots & r_{nn} \end{pmatrix}
\end{aligned}$$

[4] Of course, there can be non-extensive ways to represent such a preference by real numbers.

$$\iff \quad (\, u_1(k_1\alpha_1 + \cdots + k_n\alpha_n)\,,\,\ldots\,,\, u_n(k_1\alpha_1 + \cdots + k_n\alpha_n)\,)$$
$$<_L \;\; (\, u_1(l_1\alpha_1 + \cdots + l_n\alpha_n)\,,\,\ldots\,,\, u_n(l_1\alpha_1 + \cdots + l_n\alpha_n)\,)\,.$$

Conversely, suppose there exist real-valued functions $u_1, \ldots, u_n$ on $\Omega(S_n)$ such that (5.15) and (5.16) hold for all $x, y \in \Omega(S_n)$. Let the corresponding $n \times n$ real matrix $\{r_{ij}\}_{1 \le i,j \le n}$ be defined as $r_{ij} := u_j(\alpha_i)$ for all $i, j$. Let $q_1, \ldots, q_n \in \mathbb{R}[\varepsilon]$ be defined by (5.14). Then, by a similar argument to the above, (5.13) holds for all $k_1, \ldots, k_n,\, l_1, \ldots, l_n \in \mathbb{N}$. □

**Example 5.2.4** Suppose there is only one-type of item $S_1 = \{\, \alpha_1 \,\}$. Then $\Omega(S_1) = \{\, 0, \alpha_1,\, 2\alpha_1,\, \ldots,\, k\alpha_1,\, \ldots \,\}$. Consider the following conditions:

C1 $\prec$ is a weak order on $\Omega(S_1)$,

C2 exactly one of (a), (b), and (c) holds:

$$\begin{array}{ll} (a) & 0 \prec \alpha_1 \prec 2\alpha_1 \prec \cdots \prec k\alpha_1 \prec \cdots \\ (b) & 0 \sim \alpha_1 \sim 2\alpha_1 \sim \cdots \sim k\alpha_1 \sim \cdots \\ (c) & 0 \succ \alpha_1 \succ 2\alpha_1 \succ \cdots \succ k\alpha_1 \succ \cdots \end{array}$$

Let $\prec$ be a preference relation on $\Omega(S_1)$. We shall show that $\prec$ satisfies A1–A3 if and only if $\prec$ satisfies C1 and C2. Suppose $\prec$ satisfies A1–A3. Then, by A1, $\prec$ is a weak order on $\Omega(S_1)$. By Proposition 2.1.1 (a), exactly one of $0 \prec \alpha_1$, $0 \sim \alpha_1$, and $0 \succ \alpha_1$ holds. Hence, by using A2 or A3 repeatedly, exactly one of (a), (b), and (c) holds. Conversely, suppose $\prec$ satisfies C1 and C2. Then obviously A1–A3 hold.

Therefore, in the case $n = 1$, Theorem 5.2.1 can be reduced to the following statement: *A weak order $\prec$ on $\Omega(S_1)$ satisfies one of (a), (b), and (c) if and only if there exists a real number $q_1$ such that (5.13) holds for all $k_1, \ldots, k_n,\, l_1, \ldots, l_n \in \mathbb{N}$; moreover, $q_1$ is unique up to a positive scalar multiple.* It can easily be checked that if (a) holds then $q_1 = 1$ is a solution of (5.13). Similarly, if (b) holds then $q_1 = 0$ is a solution; if (c) holds then $q_1 = -1$ is a solution.

Also, a similar argument to Remark 5.2.3 follows: A weak order $\prec$ on $\Omega(S_1)$ satisfies one of (a), (b), and (c) if and only if there exists a real-valued function $u_1$ on $\Omega(S_1)$ such that for all $x, y \in \Omega(S_1)$,

$$x \prec y \quad \text{iff} \quad u_1(x) \;<\; u_1(y)\;; \qquad u_1(x+y) \;=\; u_1(x) \,+\, u_1(y)\,. \tag{5.17}$$

The left-hand side of (5.17) means that the real-valued function $u_1$ represents the ordering $\prec$, and the right-hand side of (5.17) ensures that $u_1$ is extensive. One can easily find such an extensive utility function $u_1$.

We note that a preference relation $\prec$ on $\Omega(S_1)$ satisfying A1–A3 (and hence C1 and C2) also admits a non-extensive representation. For example, let $\prec$ be a weak order on $\Omega(S_1)$ satisfying (a), i.e.

$$0 \prec \alpha_1 \prec 2\alpha_1 \prec \cdots \prec k\alpha_1 \prec \cdots.$$

Then there exist several different kinds of utility functions on $\Omega(S_1)$, such as

$$\text{(i)} \quad u_1\,(\,k\alpha_1\,) \;=\; k\,, \qquad\qquad \text{(ii)} \quad u_2\,(\,k\alpha_1\,) \;=\; \log k\,,$$

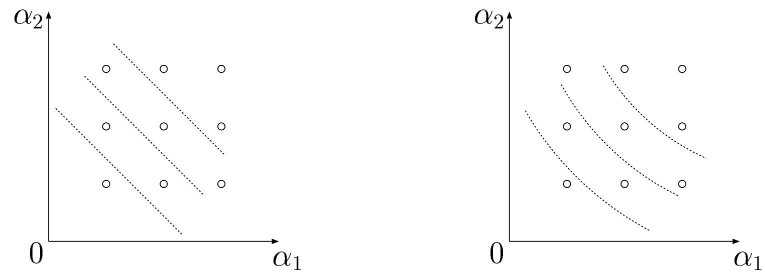


Figure 5.1: Indifference Curves

which preserve the ordering $\prec$. The former function $u_1$ is an extensive utility function, in the sense that the right-hand side of (5.17) holds, while the latter $u_2$ is not extensive. In economics, the latter function $u_2$ is considered as an example of *the principle of diminishing marginal utility,* which means that the utility of an additional item will diminish as one obtains more items. □

A similar discussion to Example 5.2.4 can also be presented for the case $n = 2$ in terms of indifference curves: if there is a preference relation on the consumption space $\Omega(S_2)$ satisfying A1–A3, it is *possible* to assign indifference curves by straight lines whose gradients are fixed everywhere. (See the left-hand side of Figure 5.1. The gradient of the lines is nothing other than the marginal rate of substitution $q_2/q_1$, the numbers $q_1$ and $q_2$ being the extensive utilities of $\alpha_1$ and $\alpha_2$, respectively. Note that the gradient may not be real-valued.) Since our formulation of A1–A3 presumes that the items $\alpha_1, \alpha_2$ are indivisible, the consumption space is discrete, i.e. each consumption plan is an isolated point on the space, and therefore the meaning of indifference curves turns out: "any consumption plan above an indifference curve is preferable to any consumption plan below the curve."

In our discrete formulation, there exist several different ways to assign such indifference curves. Under the conditions A1–A3, indifference curves can be assigned by straight lines (as we mentioned above), but also not by straight lines; for example, indifference curves can be downwards convex as illustrated in the right-hand side of Figure 5.1. This means that the conditions A1–A3 do not necessarily violate *the principle of diminishing marginal rate of substitution* (see Hicks [20]) as is usually presumed in economics.

As we mentioned above, the conditions A1–A3 do not always violates the principle of diminishing marginal rate of substitution, in the sense that indifference curves can be downwards convex. Nevertheless, in the true sense of the word, the principle means that "combination is preferred to monotonicity" as illustrated in the following example:

$$\text{" orange} \prec \text{apple, "} \quad \text{but} \quad \text{" orange + apple} \succ \text{apple + apple. "}$$

In other words, we should restate the principle of diminishing marginal rate of substitution in our discrete terms as follows: *For any different items $\alpha_1$, $\alpha_2$ with $\alpha_1 \precsim \alpha_2$, there exists a positive integer $k > 0$ such that $\alpha_1 + k\alpha_2 \succ (k+1)\alpha_2$.* Under such principle, the condition A2 or A3 must be violated. Hence, by Theorem 5.2.1, utilities of the items cannot be extensive; that is to say, the sum of the utilities of the items may be different from the utility of the sum of the items.

Therefore the conditions A1–A3 are usually not satisfied in preferences. The role of the conditions A1–A3 is, rather, to make precise the distinction between possibility and impossibility of extensive utility representation: if A1–A3 hold, then, by Theorem 5.2.1, it is possible to have extensive utilities of items (while it is also possible to have non-extensive utilities); if the conditions A1–A3 are violated, it is impossible to have extensive utilities of items.

Now we shall prove Theorem 5.2.1 by way of Theorem 3.1.5. Let $\langle G_n, +, 0\rangle$ be the Abelian group generated by $\alpha_1, \dots, \alpha_n$, i.e.

$$G_n = \{ k_1\alpha_1 + \cdots + k_n\alpha_n \mid k_1, \dots, k_n \in \mathbb{Z} \}$$

such that $k_1\alpha_1 + \cdots + k_n\alpha_n = 0$ implies $k_1 = \cdots = k_n = 0$. (In other words, $\langle G_n, +, 0\rangle$ is the $\mathbb{Z}$*-free module* with the basis $\alpha_1, \dots, \alpha_n$.) In the following, the set of consumption plans $\Omega(S_n)$ is identified with the nonnegative part of the $\mathbb{Z}$-free module $\langle G_n, +, 0\rangle$.

**Lemma 5.2.5** *For any $z \in G_n$, there exist $x, y \in \Omega(S_n)$ such that $z = y - x$.*

**Proof.** This is trivial. □

**Lemma 5.2.6** *Let $\prec$ be a preference relation on $\Omega(S_n)$ satisfying A1–A3. Then, for all $x, y, z, w \in \Omega(S_n)$, $x \prec y$ and $z \precsim w$ imply $x + z \prec y + w$.* [5]

**Proof.** Suppose $x \prec y$ and $z \precsim w$. Then, by A2, $x \prec y$ implies $x + z \prec y + z$. Also, by A2 or A3, $z \precsim w$ implies $y + z \precsim y + w$. Hence, by Proposition 2.1.1 (b) (d), we get $x + z \prec y + w$. □

**Lemma 5.2.7** *Let $\prec$ be a preference relation on $\Omega(S_n)$ satisfying A1–A3. Let*

$$P = \{ y - x \in G_n \mid x, y \in \Omega(S_n),\ x \prec y \}.$$

*Then, for all $x, y \in \Omega(S_n)$, $x \prec y$ iff $y - x \in P$.*

**Proof.** If $x \prec y$ then, by the definition, $y - x \in P$. Conversely, suppose $y - x \in P$. This means that there exist $x', y' \in \Omega(S_n)$ with $x' \prec y'$ such that $y - x = y' - x'$. Suppose also not$(x \prec y)$. Then, by Proposition 2.1.1 (a), we have $y \precsim x$. Hence, by Lemma 5.2.6, $x' \prec y'$ and $y \precsim x$ yield $x' + y \prec y' + x$ $(= x' + y)$. This is a contradiction (because $\prec$ is a weak order and hence there is no $v \in \Omega(S_n)$ such that $v \prec v$). □

**Lemma 5.2.8** *Let $\prec$ be a preference relation on $\Omega(S_n)$ satisfying A1–A3, and let $\sim$ be its associated indifference relation. Let $P$ be as in Lemma 5.2.7, and let $I = \{ y - x \in G_n \mid x, y \in \Omega(S_n),\ x \sim y \}$. Then, the pair of subsets $P$, $I$ $(\subseteq G_n)$ satisfy C1–C3.*[6]

**Proof.** C1: Left to readers.
C2: By Lemma 5.2.5, for all $z \in G_n$ there exist $x, y \in \Omega(S_n)$ such that $z = y - x$. By Proposition 2.1.1 (a), we have $x \prec y$ or $x \sim y$ or $y \prec x$; therefore $z \in P$ or $z \in I$ or $z \in -P$. It remains to show that $P$, $I$, and $-P$ are mutually disjoint. First

[5]Recall that $z \precsim w$ iff $z \prec w$ or $z \sim w$. See Section 2.1.
[6]See Section 2.3 for the definition of C1–C3.

we show that $P \cap -P = \emptyset$. Suppose there is $z \in P \cap -P$. This means that there exist $x_1, y_1, x_2, y_2 \in G_n$ with $x_1 \prec y_1$ and $x_2 \prec y_2$ such that $z = y_1 - x_1 = x_2 - y_2$. By Lemma 5.2.6, $x_1 \prec y_1$ and $x_2 \prec y_2$ yield $x_1 + x_2 \prec y_1 + y_2$ $(= x_1 + x_2)$. This is a contradiction (because $\prec$ is a weak order and hence there is no $v \in \Omega(S_n)$ such that $v \prec v$). Similarly we can prove $P \cap I = \emptyset$ and $-P \cap I = \emptyset$.

C3: Suppose $x \in P$ and $y \in P \cup I$. This means that there exist $z_1, w_1, z_2, w_2 \in G_n$ with $z_1 \prec w_1$ and $z_2 \precsim w_2$ such that $x = w_1 - z_1$ and $y = w_2 - z_2$. By Lemma 5.2.6, $z_1 \prec w_1$ and $z_2 \precsim w_2$ yield $z_1 + z_2 \prec w_1 + w_2$. Therefore $x + y = (w_1 + w_2) - (z_1 + z_2) \in P$. □

**Proof of Theorem 5.2.1:** The "if" part is obvious. For the "only if" part, suppose there is a preference relation $\prec$ on $\Omega(S_n)$ satisfying A1–A3. First we consider the trivial case, i.e. there exist no $x, y \in \Omega(S_n)$ such that $x \prec y$. Then, one can easily verifies that $q_1 = q_2 = \cdots = q_n = 0$ is the (unique) solution of (5.13). In the following, therefore, we shall assume that $\prec$ is nontrivial, i.e. there exist $x, y \in \Omega(S_n)$ such that $x \prec y$.

As we mentioned above, the set of consumption plans $\Omega(S_n)$ is identified with the nonnegative part of the $\mathbb{Z}$-free module $\langle G_n, +, 0 \rangle$. We define subsets $P$, $I$ of $G_n$ by

$$\begin{aligned} P &= \{ y - x \in G_n \mid x, y \in \Omega(S_n),\ x \prec y \}, \\ I &= \{ y - x \in G_n \mid x, y \in \Omega(S_n),\ x \sim y \}. \end{aligned} \tag{5.18}$$

Note that both $P$ and $I$ are nonempty. By Lemma 5.2.8, the pair of subsets $P$, $I$ $(\subseteq G_n)$ satisfy C1–C3. Hence, by Theorem 3.1.5, there exist $q_1, \ldots, q_n \in \mathbb{R}[\varepsilon]_n$ satisfying (3.6) for all $k_1, \ldots, k_n \in \mathbb{Z}$. Thus the following holds for all $k_1, \ldots, k_n,\ l_1, \ldots, l_n \in \mathbb{N}$:

$$\begin{aligned} \sum_{i=1}^n k_i \alpha_i \prec \sum_{i=1}^n l_i \alpha_i \quad &\Longleftrightarrow \quad \sum_{i=1}^n (l_i - k_i)\alpha_i \in P && \text{(by Lemma 5.2.7)} \\ &\Longleftrightarrow \quad \sum_{i=1}^n (l_i - k_i) q_i > 0 && \text{(by (3.6))} \\ &\Longleftrightarrow \quad \sum_{i=1}^n k_i q_i < \sum_{i=1}^n l_i q_i . \end{aligned}$$

For the proof of the uniqueness part, suppose there exist $q'_1, \ldots, q'_n \in \mathbb{R}[\varepsilon]_n$ satisfying (5.13) in the nonnegative part $\Omega(S_n)$ of the $\mathbb{Z}$-free module $\langle G_n, +, 0 \rangle$. Let $P$, $I$ be the subsets of $G_n$ defined by (5.18). We shall show that $q'_1, \ldots, q'_n$ also satisfy (3.6) of

Theorem 3.1.5 with respect to $P$, $I$ in $G_n$: For all $k_1, \ldots, k_n \in \mathbb{Z}$,

$$
\begin{aligned}
& \sum_{i=1}^{n} k_i \alpha_i \in P \\
\iff \quad & \left\{ \begin{array}{l} \displaystyle\sum_{i=1}^{n} l_i \alpha_i \prec \sum_{i=1}^{n} m_i \alpha_i \\ \qquad \text{for some } l_i, m_i \in \mathbb{N} \text{ with } k_i = m_i - l_i \quad (i = 1, \ldots, n) \end{array} \right. \\
& \hfill \text{(by Lemma 5.2.5 and Lemma 5.2.7)} \\
\iff \quad & \left\{ \begin{array}{l} \displaystyle\sum_{i=1}^{n} l_i q'_i < \sum_{i=1}^{n} m_i q'_i \\ \qquad \text{for some } l_i, m_i \in \mathbb{N} \text{ with } k_i = m_i - l_i \quad (i = 1, \ldots, n) \end{array} \right. \\
& \hfill \text{(by (5.13))} \\
\iff \quad & \sum_{i=1}^{n} k_i q'_i > 0 .
\end{aligned}
$$

Therefore the uniqueness part of Theorem 5.2.1 follows from the corresponding one of Theorem 3.1.5. **Q.E.D.**

# Chapter 6

# Conclusion

In this chapter, we summarize our results and suggest future work.

The main theorem of this thesis is a lexicographical separation theorem stating that a convex cone and its convex complement in $\mathbb{F}^n$ can be separated by linear functions and the lexicographic order on $\mathbb{R}^n$, where $\mathbb{F}$ stands for an ordered field such that $\mathbb{Q} \subseteq \mathbb{F} \subseteq \mathbb{R}$. The main theorem is a modification of the lexicographical separation theorems due to Hausner and Wendel [18], Klee [23], Martínez-Legaz and Singer [30], by considering not only the real numbers $\mathbb{R}$ but also other ordered fields $\mathbb{F}$, and moreover by proving a kind of uniqueness result. We also provided applications of the main theorem to linear inequality systems, lexicographic expected utility, and extensive measurement.

It may be possible to generalize the main theorem for an arbitrary ordered field $\mathbb{F}$: the author conjectures that a convex cone and its convex complement in $\mathbb{F}^n$ can be separated by linear functions and the lexicographic order on $\tilde{\mathbb{F}}^n$, where $\mathbb{F}$ stands for an arbitrary ordered field and $\tilde{\mathbb{F}}$ denotes the completion of $\mathbb{F}$. If such a generalization of the main theorem is obtained, its applications will also be generalized in similar manners.

In this thesis, we presented the main theorem and its applications not only in terms of the lexicographic order but also in terms of the polynomial ring $\mathbb{R}[\varepsilon]$, where $\mathbb{R}[\varepsilon]$ denotes the smallest ring containing both $\mathbb{R}$ and an infinitesimal $\varepsilon$. We used the fact that the lexicographic order on $\mathbb{R}^n$ can be described by the polynomial ring $\mathbb{R}[\varepsilon]$, in the sense that

$$(a_0, a_1, \ldots, a_n) <_L (b_0, b_1, \ldots, b_n) \quad \text{iff} \quad a_0 + a_1\varepsilon + \cdots + a_n\varepsilon^n < b_0 + b_1\varepsilon + \cdots + b_n\varepsilon^n .$$

Such a description is well-known in the literature, but we gave a new role to an infinitesimal in this thesis: (i) We used an infinitesimal as a useful tool of proving the main theorem. The use of an infinitesimal makes the proof easier because it allows both addition and multiplication. (ii) We adopted an infinitesimal as a solution to linear inequality systems. The use of an infinitesimal enable us to give a necessary and sufficient condition for the existence of solutions to infinite systems of linear inequalities.

It should be noted that the lexicographic order can also be described by an *infinitely large number* $\omega$, i.e. $k < \omega$ for all positive integer $k$, together with an infinitesimal $\varepsilon$. More precisely, one can show that, as the direct analogy of decimal number system,

$$(a_n, \ldots, a_1, a_0, a_{-1}, \ldots, a_{-m}) <_L (b_n, \ldots, b_1, b_0, b_{-1}, \ldots, b_{-m}) \quad \text{iff} \quad a_n\omega^n + \cdots + a_1\omega + a_0 + a_{-1}\varepsilon + \cdots + a_{-m}\varepsilon^m < b_n\omega^n + \cdots + b_1\omega + b_0 + b_{-1}\varepsilon + \cdots + b_{-m}\varepsilon^m .$$

Therefore the main theorem and its applications in this thesis can also be presented in terms of an infinitely large number $\omega$, as well as an infinitesimal $\varepsilon$. As will be mentioned below, the use of an infinitely large number $\omega$ together with an infinitesimal $\varepsilon$ enables us to obtain a generalized solution to linear inequality systems.

As an application of the main theorem to linear inequality systems, we gave a necessary and sufficient condition for the existence of solutions to infinite system of linear inequalities: we showed that the origin $\mathbf{0}$ is not contained in the convex hull of a subset $P$ of $\mathbb{R}^n$ if and only if the inequality system

$$\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n \ > \ 0 \quad \text{for all} \ \ (\lambda_1, \lambda_2, \ldots, \lambda_n) \in P$$

has solutions $x_1, x_2, \ldots, x_n$ in $\mathbb{R}[\varepsilon]$. The result is a generalization of the well-known theorem of the alternatives for finite linear inequality systems. Also, we gave a generalization of Farkas' lemma for lexicographical inequality systems.

The author obtains a further generalization of the above result, using an infinitely large number $\omega$ together with an infinitesimal $\varepsilon$: Let

$$\mathbb{R}[\omega, \varepsilon] \ = \\ \{ r_n \omega^n + \cdots + r_1 \omega + r_0 + r_{-1}\varepsilon + \cdots + r_{-n}\varepsilon^n \mid n \in \mathbb{N} \ \ r_n, \ldots, r_1, r_0, r_{-1}, \ldots, r_{-n} \in \mathbb{R} \},$$

and let $P$ be a nonempty subset of $\mathbb{R}^{n+1}$. Then, the origin $\mathbf{0}$ is not contained in the convex hull of $P \cup (0, \ldots, 0, -1)$ if and only if the inequality system

$$\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n \ > \ \lambda_{n+1} \quad \text{for all} \ \ (\lambda_1, \lambda_2, \ldots, \lambda_n, \lambda_{n+1}) \in P$$

has solutions $x_1, x_2, \ldots, x_n$ in $\mathbb{R}[\omega, \varepsilon]$. This result will be presented in a future paper.

As other application of the main theorem, we presented two kinds of lexicographic utility representations: one is about lexicographic expected utility, and the other is about lexicographic extensive utility.

The lexicographic expected utility representation given in this thesis is a modification of Hausner's lexicographic expected utility theory, by omitting the existence of irrational-valued probabilities: in this thesis we restricted our attention to $\mathbb{F}$-valued probabilities, where $\mathbb{F}$ stands for an ordered field such that $\mathbb{Q} \subseteq \mathbb{F} \subseteq \mathbb{R}$, and showed that lexicographic expected utility theory can be founded on the domain of $\mathbb{F}$-valued lotteries. The author plans to apply the result to game theory: it will be interesting to determine in each game what kind of field $\mathbb{F}$ is needed, as the domain of probabilities, to obtain a Nash equilibrium.

For lexicographical extensive utility representation, we established a scheme of conditions which is necessary and sufficient for the existence of extensive utilities on indivisible items. The scheme of conditions throws a new light on the distinction between possibility and impossibility of "addition of utility." The scheme of conditions may also be applicable to physical sciences, whenever one wants to obtain extensive measurement.

# Bibliography


[1] Birkhoff, G., *Lattice Theory*, 3rd edition, American Mathematical Society, New York, 1967.

[2] Blackwell, D. and Girshick, G. A., *Theory of Games and Statistical Decisions*, Wiley, New York, 1954.

[3] Brams, S. J. and Fishburn, P. C., Fair Division of Indivisible Items between Two People with Identical Preferences: Envy-freeness, Pareto-optimality, and Equity, *Social Choice and Welfare*, **17** (2000), 247–267.

[4] Chipman, J. S., The Foundations of Utility, *Econometrica*, **28** (1960), 193–224.

[5] Debreu, G., Representation of a Preference Order by a Numerical Function, in Thrall, R. M., Coombs, C. H., and Davis, R. L., editors, *Decision Processes*, Wiley, New York, 1954, pp. 159–166.

[6] Fishburn, P. C., *Utility Theory for Decision Making*, Wiley, New York, 1970.

[7] Fishburn, P. C., A Study of Lexicographic Expected Utility, *Management Science*, **17** (1971), 672–678.

[8] Fishburn, P. C., Lexicographic Orders, Utilities and Decision Rules: A Survey, *Management Science*, **20** (1974), 1442–1471.

[9] Fishburn, P. C., On the Nature of Expected Utility, in Allais, M. and Hagen, O., editors, *Expected Utility Hypothesis and the Allais Paradox*, Reidel, Dordrecht, 1979, pp. 243–257.

[10] Fishburn, P. C., *The Foundations of Expected Utility*, Reidel, Dordrecht, 1982.

[11] Fuchs, L., *Partially Ordered Algebraic Systems*, Pergamon Press, 1963.

[12] Gale, D., *The Theory of Linear Economic Models*, McGraw-Hill, New York, 1960.

[13] Goldblatt, R., *Lectures on the Hyperreals: An Introduction to Nonstandard Analysis*, Springer, New York, 1998.

[14] Hahn, H., Über die Nichtarchimedischen Grossensysteme, *Sitzungsberichte der Matematisch-Naturwissenschaftlichen Klasse der Kaiserlichen Wiener Akademie der Wissenschaftlichen*, Abteilung IIa, **116** (1907), 601–655.

[15] Hammond, P. J., Elementary Non–Archimedean Representations of Probability for Decision Theory and Games, in Humphreys, P., editor, *Patrick Suppes: Scientific Philosopher, Volume 1*, Kluwer Academic Publishers, 1994, pp. 25–59.

[16] Hammond, P. J., Objective Expected Utility: A Consequentialist Perspective, in Barcera, S., Hammond, P. J., and Seidl, C., editors, *Handbook of Utility Theory, Volume 1*, Kluwer Academic Publishers, 1998, pp. 143–211.

[17] Hausner, M., Multidimensional Uilities, in Thrall, R. M., Coombs, C. H., and Davis, R. L., editors, *Decision Processes*, Wiley, New York, 1954, pp. 167–180.

[18] Hausner, M. and Wendel, J. G., Ordered Vector Spaces, *Proceedings of the American Mathematical Society*, **3** (1952), 977–982.

[19] Herstein, I. N. and Milnor, J., An Axiomatic Approach to Measurable Utility, *Econometrica*, **21** (1953), 291–297.

[20] Hicks, J. R., *Value and Capital*, 2nd edition, Oxford University Press, London, 1946.

[21] Jensen, N. E., An Introduction to Bernoullian Utility Theory I: Utility Functions, *Swedish Journal of Economics*, **69** (1967), 163–183.

[22] Kaneko, M., *Game Theory and Konnyaku Mondô* (in Japanese), Nihon-Hyôron-Sha, to appear.

[23] Klee, V., Separation and Support Properties of Convex Sets—A Survey, in Balakrishnan, A. V., editor, *Control Theory and the Calculus of Variations*, Academic Press, New York, 1969, pp. 235–303.

[24] Kuhn, H. W. and Tucker, A. W., *Linear Inequalities and Related Systems*, Princeton University Press, Princeton, 1956.

[25] Lang, S., *Algebra*, 3rd edition, Addison-Wesley, 1993.

[26] Machina, M. J., Choice under Uncertainty: Problems Solved and Unsolved, *Journal of Economic Perspectives*, **1** (1987), 121–154.

[27] Martínez-Legaz, J. E., Exact Quasiconvex Conjugation, *Zeitschrift für Operations Research*, **27** (1983), 257–266.

[28] Martínez-Legaz, J. E., Lexicographical Order, Inequality Systems and Optimization, in Thoft-Christensen, P., editor, *System Modelling and Optimization*, Springer, Berlin, 1984, pp. 203–212.

[29] Martínez-Legaz, J. E., Lexicographic Utility and Orderings, in Barcera, S., Hammond, P. J., and Seidl, C., editors, *Handbook of Utility Theory, Volume 1*, Kluwer Academic Publishers, 1998, pp. 345–369.

[30] Martínez-Legaz, J. E. and Singer, I., Lexicographical Separation in $\mathbb{R}^n$, *Linear Algebra and Its Applications*, **90** (1987), 147–163.

[31] Martínez-Legaz, J. E. and Singer, I., Lexicographical Order, Lexicographical Index, and Linear Operators, *Linear Algebra and Its Applications*, **128** (1990), 65–95.

[32] Martínez-Legaz, J. E. and Singer, I., Compatible Preorders and Linear Operators on $\mathbb{R}^n$, *Linear Algebra and Its Applications*, **153** (1991), 53–66.

[33] Nakamura, Y., Finite-Dimensional Utilities, *Economic Theory*, **16** (2000), 209–218.

[34] Nakamura, Y., Lexicographic Quasilinear Utility, *Journal of Mathematical Economics*, **37** (2002), 157–178.

[35] Narens, L., Measurement without Archimedean Axioms, *Philosophy of Science*, **41** (1974), 374–393.

[36] Narens, L., *Abstract Measurement Theory*, M.I.T. Press, Cambridge, MA, 1985.

[37] Owen, G., *Game Theory*, 3rd edition, Academic Press, 1995.

[38] Roberts, F. S., *Measurement Theory*, Addison-Wesley, 1979.

[39] Rockafellar, R. T., *Convex Analysis*, Princeton University Press, Princeton, 1970.

[40] Scott, D., Measurement Structures and Linear Inequalities, *Journal of Mathematical Psychology*, **1** (1964), 233–247.

[41] Sion, M., On General Minimax Theorem, *Pacific Journal of Mathematics*, **8** (1958), 171–176.

[42] Skala, H. J., *Non–Archimedean Utility Theory*, Reidel, Dordrecht, 1975.

[43] Teh, H. -H., Construction of Orders in Abelian Groups, *Proceedings of the Cambridge Philosophical Society*, **57** (1961), 476–482.

[44] von Neumann, J., Zur Theorie der Gesellschaftspiele, *Mathematische Annalen*, **100** (1928), 295–320.

[45] von Neumann, J. and Morgenstern, O., *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, 1944.

[46] Wald, A., Generalization of a Theorem by von Neumann Concerning Zero-Sum Two-Person Games, *Annals of Mathematics*, **46** (1945), 281–286.

# Publications

K. Ikeda, "Non-Archimedean Solutions to Linear Inequality Systems," the 7th International Symposium on Generalized Convexity/Monotoniciy, Hanoi, August 27–31, 2002.

K. Ikeda, "Lexicographic Extensive Utility on Indivisible Items," to be submitted to the *Mathematical Social Sciences*.

K. Ikeda, "Non–Archimedean Additive Utilities for Discrete Items," Research Report IS-RR-2001-029, JAIST, 2001.

# Appendix

**Proof of Proposition 1.1.1 :** Let $z$ be the point in $B$ whose distance from $x$ is minimum. (Such a point exists because $B$ is closed.) Let

$$\begin{aligned} p_i &= z_i - x_i \qquad \text{for } i = 1, \cdots, n\,, \\ p_{n+1} &= \sum_{i=1}^{n} z_i\, x_i - \sum_{i=1}^{n} x_i^2\,. \end{aligned}$$

Clearly (1.1) holds. We shall show (1.2) also holds. By the above conditions,

$$\sum_{i=1}^{n} p_i\, z_i = \sum_{i=1}^{n} z_i^2 - \sum_{i=1}^{n} z_i\, x_i$$

and hence

$$\begin{aligned} \sum_{i=1}^{n} p_i\, z_i - p_{n+1} &= \sum_{i=1}^{n} z_i^2 - 2\sum_{i=1}^{n} z_i\, x_i + \sum_{i=1}^{n} x_i^2 \\ &= \sum_{i=1}^{n} (z_i - x_i) > 0\,. \end{aligned}$$

Therefore

$$\sum_{i=1}^{n} p_i\, z_i > p_{n+1}\,.$$

Suppose now that there exists $y \in B$ such that

$$\sum_{i=1}^{n} p_i\, y_i \leq p_{n+1}\,.$$

Because $B$ is convex, the line joining $y$ to $z$ must be entirely contained in $B$, i.e. for all $0 \leq r \leq 1$ we have $w_r = ry + (1-r)z \in B$. The square of the distance from $x$ to $w_r$ is given by

$$\rho^2(x, w_r) = \sum_{i=1}^{n} (x_i - ry_i - (1-r)z_i)^2\,.$$

Therefore

$$\frac{\partial \rho^2}{\partial r} = 2\sum_{i=1}^{n} (z_i - y_i)(x_i - ry_i - (1-r)z_i)$$

$$= \; 2 \sum_{i=1}^{n} (z_i \,-\, x_i)\, y_i \;-\; 2 \sum_{i=1}^{n} (z_i \,-\, x_i)\, z_i \;+\; 2 \sum_{i=1}^{n} r\,(z_i \,-\, y_i)^2$$
$$= \; 2 \sum_{i=1}^{n} p_i\, y_i \;-\; 2 \sum_{i=1}^{n} p_i\, z_i \;+\; 2r \sum_{i=1}^{n} (z_i \,-\, y_i)^2 \,.$$

Let us evaluate this at $r = 0$ (i.e. $w_r = z$):

$$\left.\frac{\partial \rho^2}{\partial r}\right|_{r=0} \;=\; 2 \sum_{i=1}^{n} p_i\, y_i \;-\; 2 \sum_{i=1}^{n} p_i\, z_i \,.$$

But the first term on the right-hand side of this equation was assumed to be less than or equal to $2p_{n+1}$, while the second is greater than $2p_{n+1}$. Therefore

$$\left.\frac{\partial \rho^2}{\partial r}\right|_{r=0} \;<\; 0\,.$$

It follows that, for $r$ close enough to zero,

$$\rho(x, w_r) \;<\; \rho(x, z)\,.$$

But this contradicts the way in which $z$ is chosen. Thus (1.2) must hold. □

**Proof of Proposition 1.1.3:** Suppose both (i) and (ii) hold. Then

$$0 \;=\; \langle x,\, 0\rangle \;=\; \langle\, x,\, (\sum_{i=1}^{m} p_i\, a_i\,)\,\rangle \;=\; \sum_{i=1}^{m} p_i\, \langle a_i,\, x\rangle \;>\; 0\,,$$

a contradiction. It remains to show that if (i) does not hold then (ii) holds.

Suppose (i) does not hold. Let $B$ be the convex hull[1] of the $m$ points $a_1, \ldots, a_m$ in $\mathbb{R}^n$. Applying Proposition 1.1.1, there exist numbers $p_1, \ldots, p_n,\, p_{n+1}$ such that

$$\sum_{j=1}^{n} 0 \cdot p_j \;=\; p_{n+1}$$

(this means $p_{n+1} = 0$) and

$$\sum_{j=1}^{n} y_j\, p_j \;>\; 0$$

for any $y = (y_1, \cdots, y_n)$ in $B$. In particular, this holds when $y$ is one of the vectors $a_1, \cdots, a_m$. Hence, $p = (p_1, \ldots, p_n)$ satisfies

$$\langle a_i,\, p\rangle \;>\; 0 \qquad \text{for } i = 1, \ldots, m\,.$$

□

[1]See Section 2.2 for the definition of the convex hull of $a_1, \ldots, a_m$.