4

PROGRAMMING IN LINEAR SPACES[1]

LEONID HURWICZ

I. Introduction

I.1. The present study has its origin in problems of optimal resource allocation, especially those related to the possibilities of a price mechanism. While for some purposes Pareto-optimality might be the more relevant concept, we have confined ourselves here to the case where by "optimal" is meant "efficient" resource allocation.[2]

The main result of the present chapter is an extension of the Kuhn-Tucker results [31] on "non-linear programming" to more general linear topological spaces.[3] (Numbers in square brackets indicate the references). The initial stimulus toward this type of generalization was the paper by Rosenbloom [39]. The need for it became apparent in the course of a larger study on problems of decentralization in resource allocation mechanisms.

The remainder of the Introduction is devoted to a brief statement of the nature of the problem. I.2 is primarily directed at the reader interested in the relevance of the study from the viewpoint of economics. I.3 provides a summary of some of the results being generalized and some of the mathematical issues arising.

II is devoted to introducing some of the basic concepts and notations. III and IV are devoted to the derivation of certain theorems on linear inequalities in linear topological spaces, among them the "Minkowski-Farkas Lemma," fundamental in the sequel, and another result of

[1] This study might not have been undertaken, and almost certainly would not have been completed, without persevering encouragement by Tjalling C. Koopmans of the Cowles Commission for Research in Economics. The author is greatly indebted to Professor Paul Rosenbloom, of the University of Minnesota, for having made him aware of the potentialities of the Banach space theory for the problems here treated. He also wishes to acknowledge valuable advice received from Professors Bernard R. Gelbaum and Gerhard K. Kalisch, also of the University of Minnesota, as well as from Kenneth J. Arrow and Hirofumi Uzawa of Stanford University. Thanks are due a number of readers, and to K. S. Kretschmer in particular, for pointing out errors in the first printing; see especially the note on p. 74.

[2] Cf. Koopmans [27].

[3] Some of the results, as for instance the generalized "Minkowski-Farkas Lemma," may be of independent interest.

DOI 10.1007/978-3-0348-0439-4_8, © Springer Basel 2014

importance in relating the theory of programming to that of games of strategy. An appendix to III relates the results of III to the theory of linear equations in Banach spaces, as formulated in a paper by Hausdorff [18].[4] (See the NOTE in brackets on page 74.) V.1 states conditions under which a Lagrangian saddle-point implies maximality (efficiency). V.2 deals with the problem of scalarization, i.e., of reducing a vectorial maximization problem to one of scalar maximization. V.3 contains the main results concerning the existence of saddle-points and "quasi-saddle-points." The third section, V.3.3 treats situations where the differential ("marginal") first-order conditions for the saddle-points are satisfied, while in the first, V.3.1, the differentiability is not assumed. (From the economist's viewpoint, the existence of a saddle-point corresponds to the existence of a price-vector equilibrating the market.) V.3.2 is devoted to the special ("linear") case which, in view of the interest in "linear programming" models, seemed worthy of separate direct treatment. The author has not completely avoided repetition where he feared that brevity might cause ambiguity. Also, many "obvious" and "trivial" proofs are spelled out in detail.

I.2. In problems of efficient resource allocation we deal with a model where commodities are classified into *resources*, typically available in limited amounts, and *desirables* in terms of which efficiency is defined. The amounts of resources used up and of the desirables produced are determined by the decision as to *activity levels*. Thus the model is of the type treated in activity analysis,[5] although not necessarily under the assumptions of additivity and linearity.

Some of the mathematical problems arising in models where linearity and additivity are not assumed have been explored by Kuhn and Tucker [31], and Slater [41]. The treatment in both papers is confined to the case where there is a finite number of commodities and a finite number of activities. This, of course, limits their domain of application. In economics there are many problems where, for instance, an infinity of commodities could be more naturally postulated, as in the case of problems involving time or location. But there is another reason why an economist may be interested in having a theory of resource allocation in which the commodity space or the activity space is of a more general nature: the logical structure of treatment of the more general situations often reveals the "deeper" or "intuitive" bases of important propositions and helps focus attention on the more fundamental features of the problem.

The economic interpretation of the Kuhn-Tucker and Slater results (discussed by Kuhn and Tucker) has to do with the possibility of reaching positions of efficient resource allocation through the price mechanism.

---

[4] On a number of occasions we explore the question of the necessity of the underlying hypotheses and several theorems are devoted to this.

[5] Cf. [27].

Roughly speaking,[6] when suitable conditions are satisfied [in the economist's language the main ones could be described as "perfect divisibility" (all positive multiples permissible), and absence of "external (dis-) economies of scale" and of "increasing returns"], (a) "competitive equilibrium" implies that an efficient point has been reached and (b) any efficient point can be one of "competitive equilibrium" (provided prices are properly selected). When the absence of "increasing returns" is not assumed (while the assumptions of "perfect divisibility" and absence of "external (dis-) economies of scale" are retained), it is still possible to obtain criteria[7] permitting classification of certain situations as non-efficient.

Results of the latter type are of considerable importance, for they serve as a basis for the development of a theory of resource allocation applicable to a class of situations not excluding "increasing returns."[8]

In attempting to generalize results of this type, the writer was guided by his interest in cases where "increasing returns" might prevail and hence "marginal" type phenomena would have to be considered. Mathematically, this meant working in a space with an operation of differentiation possessing most of its usual properties. The Banach spaces form the most general class of spaces with which the writer was familiar at the time this was written, although a more general theory of differentiation does exist.[9] However, in theorems where the differential operations were not used, an attempt was made to obtain proofs valid for a more general class of linear spaces. In V.3.1, the author has treated the case of Lagrangian saddle-points by methods of the type used by Slater, i.e., relying on convexity but not using differentiability.

If one is to treat phenomena of "indivisibilities," one must go beyond linear spaces. But since one knows that most of the important results valid in linear spaces cannot be expected to hold when "indivisibilities" appear, it becomes desirable to reappraise the objectives of the theory of resource-allocating mechanisms, especially in their "decentralization" aspects. This is done to some extent in another paper now being completed by the writer.

I.3. In this section we give a brief description of some of the results being generalized in the present chapter.

---

[6] For a more precise statement the reader is referred to Koopmans [27], Kuhn and Tucker [31], and Arrow and Hurwicz [2], and Chapter 3 of the present book.

[7] These criteria are of differential ("marginal") first-order nature; they involve prices, but not the full conditions of "competitive equilibrium." Cf. Theorem 1 in Kuhn and Tucker [31].

[8] Cf. the work of Hotelling [21], Lange and Taylor [33], and Lerner [35], especially as it involves "marginal cost" pricing. Cf. also Arrow and Hurwicz [2], and Chapter 6 of the present book.

[9] Cf. Hyers [22], and its Bibliography. See V. 3.3.8, where results involving differential operations are extended to a wider class of linear spaces.

Let $\mathscr{X}$, $\mathscr{Y}$, and $\mathscr{Z}$ be finite-dimensional Euclidean spaces,[10] their dimensionalities being respectively n_x, n_y, n_z.

In each space we define certain ordering relations. If $v' = (v'_1, \cdots, v'_{n_v})$, $v'' = (v''_1, \cdots, v''_{n_v})$ are two vectors in the space $\mathscr{V}$ (where $\mathscr{V}$ may be $\mathscr{X}$, $\mathscr{Y}$, or $\mathscr{Z}$), we write

$$v' \geqq v'' \text{ to mean } v'_i \geqq v''_i \text{ for } i = 1, 2, \cdots, n_v ,$$
$$v' \geq v'' \text{ to mean } v' \geqq v'' \text{ and } v'' \not\geqq v' \text{ (i.e., } v' \neq v'') ,$$
$$v' > v'' \text{ to mean } v'_i > v''_i \text{ for } i = 1, 2, \cdots, n_v .$$

The origins of the three spaces treated are denoted by $0_x, 0_y, 0_z$, but the subscripts are omitted where no danger of confusion seems to exist.

Given a set Y in the space $\mathscr{Y}$ of desirables (Y is the "attainable" set, cf. [27], p. 47), we define its *maximal* (= "efficient") subset $\hat{Y}$ by the condition

$$\hat{Y} = \{y \in Y : \; y' \in Y, y' \geqq y \text{ imply } y' \leqq y\} .$$

I.e., y is a maximal (= efficient) element of Y if, and only if, $y' \geq y$ does not hold for any element y' of Y.

The set Y, however, is given indirectly, since our decision variable is x, not y. Thus there are given two (single-valued) functional relations[11] f, g with a common domain in $\mathscr{X}$ and ranges in $\mathscr{Y}$ and $\mathscr{Z}$ respectively and

$$Y = f[P_x \cap g^{-1}(P_z)]$$

where P_x and P_z are the respective non-negative orthants of $\mathscr{X}$ and $\mathscr{Z}$. I.e., $y \in Y$ if, and only if, $y = f(x)$ with $x \geqq 0$ and $g(x) \geqq 0$.

Somewhat inaccurately we shall call an element x *maximal* when $f(x)$ is maximal and we write $\hat{X} = f^{-1}(\hat{Y})$.

The problem of finding necessary and sufficient conditions for the maximality of a point x in $\mathscr{X}$ is usually called the problem of *vectorial* maximization of $f(x)$ subject to the *constraints* $x \geqq 0$, $g(x) \geqq 0$. When $n_y = 1$, we are of course dealing with scalar maximization. Corresponding to a given maximization problem one may construct its *Lagrangian* (*expression*) given by

$$\Phi(x, z^* \,;\, y^* \,;\, f, g) = y^*[f(x)] + z^*[g(x)]$$

where

$$y^*(y) = y^{*\prime}y = \sum_{i=1}^{n_y} y_i^* y_i$$

[10] In the economic interpretation, $\mathscr{X}$ is the space of activity level vectors, $\mathscr{Y}$ the space of the vectors of desirables, $\mathscr{Z}$ that of resources.

[11] In line with the prevailing practice, we use f where Kuhn and Tucker use g and vice versa.

and

$$z^*(z) = z^{*\prime}z = \sum_{i=1}^{n_z} z_i^* z_i \, .$$

(In matrix and vector notation A' is the transpose of A. Depending on the content, we omit some or all of the detail following the symbol Φ.)

We say that Φ has a *non-negative saddle-point* at $(x_0, z_0^* ; y_0^*)$ if and only if

$$x_0 \geqq 0, \qquad z_0^* \geqq 0, \qquad y_0^* > 0,$$

and

$$\Phi(x, z_0^* ; y_0^*) \leqq \Phi(x_0, z_0^* ; y_0^*) \leqq \Phi(x_0, z^* ; y_0^*) \, ,$$

for all $x \geqq 0$ and all $z^* \geqq 0$.

Φ is said to have a *non-negative quasi-saddle-point* at $(x_0, z_0^* ; y_0^*)$ if and only if

$$x_0 \geqq 0, \qquad z_0^* \geqq 0, \qquad y_0^* > 0 \, ,$$

and

$$\begin{cases} \Phi_x^0 \leqq 0, & \Phi_x^{0\prime} x_0 = 0, \\ \Phi_{z^*}^0 \geqq 0, & \Phi_{z^*}^{0\prime} z_0^* = 0 \, . \end{cases}$$

[Here $\Phi_x^0 = < \partial\Phi/\partial x_1, \cdots, \partial\Phi/\partial x_{n_x} >$ (see fn. 1, p. 2, for this notation) with all derivatives evaluated at $(x_0, z_0^* ; y_0^*)$; $\Phi_z^{0*} = < \partial\Phi/\partial z_1^*, \cdots, \partial\Phi/\partial z_{n_z}^* >$ with all derivatives evaluated at $(x_0, z_0^* ; y_0^*)$.] It was shown by Kuhn and Tucker ([31], Lemma 1) that a non-negative saddle-point is always a non-negative quasi-saddle-point. The converse is false.

In order to state the main results of Kuhn and Tucker we need three concepts :

a) a function f, with convex domain D in $\mathscr{X}$ and range in $\mathscr{Y}$, is said to be *concave* if, and only if, for any x', x'' in D and any $0 < \theta < 1$ the inequality $(1-\theta)f(x') + \theta f(x'') \leqq f[(1-\theta)x' + \theta x'']$ holds ;

b) the function g, with domain in $\mathscr{X}$ and range in $\mathscr{Y}$, is said to be *regular* if and only if the "constraint qualification" (cf. Kuhn and Tucker [31], p. 483) is satisfied ;

c) x_0 is *properly maximal* if it is a proper solution of the vector maximum problem in the sense of Kuhn and Tucker ([31], p. 488).[12]

The following are results of interest :

1. (Kuhn and Tucker [31], Theorem 4.) Let x_0 be properly maximal, f and g differentiable, and g regular for $x \geqq 0$. Then there exist y_0^*, z_0^* such that $\Phi(x, z^* ; y^*)$ has a non-negative quasi-saddle-point at $(x_0, z_0^* ; y_0^*)$. (Note : For $n_y = 1$, the term "properly" may be omitted and Theorem 4 becomes Theorem 1 of Kuhn and Tucker.)

[12] Let $f(x_0)$ be properly maximal whenever x_0 is. Then the result of Arrow, Barankin, and Blackwell [1] seems to show that, at least when $Y = f[P_x \cap g^{-1}(P_z)]$ is closed and convex, the set of properly maximal y's is dense in the set of maximal points.

2. (Kuhn and Tucker [31], Theorem 6.) Let both f and g be differentiable and concave, and g regular for $x \geqq 0$. Then x_0 is properly maximal if and only if there exist y_0^*, z_0^* such that $\Phi(x, z^* ; y^*)$ has a non-negative saddle-point at $(x_0, z_0^* ; y_0^*)$. (Note: For $n_y = 1$, the term "properly" may be omitted and Theorem 6 becomes Theorem 3 of Kuhn and Tucker.)

The following comments may be found helpful in following the later sections of this paper.

1. The "if" part of Theorem 6 fails to hold when $y^* \geq 0$ instead of the stronger $y^* > 0$ which is postulated. This raises a difficulty in generalizing to linear (or even Banach) spaces, since in some of them a $y^* > 0$ may not exist.

2. The "if" part of Theorem 6 remains valid when the assumptions of differentiability and concavity of f and g and regularity of g are abandoned.[13]

3. The "only if" part of Theorem 6 depends on Theorem 4 and the concavity of f and g.

4. The proof of Theorem 4 consists in "scalarizing" the problem by means of an appropriate $y^* > 0$ (which will exist if x_0 is properly maximal) and then using Kuhn-Tucker Theorem 1 covering the case $n_y = 1$.

5. The crucial step in the proof of the Kuhn-Tucker Theorem 1 involves the use of the Minkowski-Farkas Lemma which states that if, A being an $m \times n$ matrix,

$$Ax \geqq 0 \quad \text{implies } b'x \geqq 0 \text{ for all } x ,$$

then there exists $t \geqq 0$ such that $b = A't$. (Cf. [31], p. 484.) Thus in attempting to generalize the results of Kuhn and Tucker the success hinges on finding the conditions under which the linear topological space counterpart of the Minkowski-Farkas proposition is valid.

6. The relationship of the present chapter to the results of Kuhn and Tucker is similar to that of Goldstine's paper [15] to, say, Bliss's discussion in [4], p. 210 ff. Goldstine treats the case of constraints in the form of equalities and imposes requirements strict enough to imply the existence of unique Lagrangian functionals ("multipliers"). Some of these results (in the "relaxed" form where uniqueness need not be present) are special cases of the theorems obtained in the present chapter.

7. Slater [41], assumes f and g to be continuous and postulates that they have a property (which we shall call "almost concavity")[14] implied by (but not implying) the concavity of both f and g; neither f nor g

[13] This suggests itself in reading Slater [41], p. 11

[14] Suppose that, for some $y^* \geqq 0$, $z^* \geqq 0$, $y^*[f(x^1)] + z^*[g(x^1)] = y^*[f(x^2)] + z^*[g(x^2)]$. Then "almost concavity" of (f, g) requires that $y^*[f(x)] + z^*[g(x)] \geqq y^*[f(x^i)] + z^*[g(x^i)]$ for all x on the segment joining x^1, x^2.

is assumed differentiable; instead of requiring that g be regular, it is required that, for some $x_1 \geqq 0$, $g(x_1) > 0$. (When this is so, we shall call g *Slater-regular*.) If by a *Slater-maximal* element of $Y = f[P_x \cap g^{-1}(P_z)]$ is meant a $y_0 \in Y$ such that $y' > y_0$ for no $y' \in Y$, and if x_0 is called Slater-maximal when $f(x_0)$ is Slater-maximal, then Slater's main result (Slater [41], Theorem 3) may be stated as follows:

Let f and g be continuous and almost concave and let g be Slater-regular. Then x_0 is Slater-maximal if, and only if, there exist

$$y_0^* \geq 0, z_0^* \geqq 0 \text{ such that}$$
$$\Phi(x, z_0^* ; y_0^*) \leq \Phi(x_0, z_0^* ; y_0^*) \leq \Phi(x_0, z^* ; y_0^*)$$

for all $x \geqq 0$ and all $z^* \geqq 0$.

It may be noted that the concept of Slater-maximality is weaker than that of maximality (as previously defined) and that it makes the "if" part of the theorem valid even though $y_0^* > 0$ is not required.

If one wanted to substitute "maximal" for "Slater-maximal" in Slater's Theorem 3, it is clear from known examples that one would have to require x_0 to be properly maximal and not merely maximal, as well as to specify that $y_0^* > 0$.

Slater's Theorem 3 is, of course, a counterpart of the Kuhn and Tucker Theorem 6. In the special case $n_y = 1$ the two concepts of maximality coincide; also $y_0^* \geq 0$ becomes equivalent to $y_0^* > 0$. Hence in this case the Slater result differs from that of Kuhn and Tucker only with regard to the hypotheses, since the assertion is precisely the same.

II. Notation, Terminology, and Some Fundamental Lemmas

II.1.0. This chapter deals with problems arising in spaces here called *linear topological spaces*. These spaces have both an algebraic structure (they are linear systems, i.e., sets of vectors, with vector addition and scalar multiplication) and a topological structure (they are topological spaces), and, furthermore, the two structures are related by the requirement that each of the algebraic operations be a (jointly) continuous function of its two arguments.[15]

The concept of a linear topological space, to be introduced more formally below, is a natural generalization of the properties of the finite-dimensional space (the real line being the simplest case) in its

[15] The definition of a linear topological space used in this paper is exactly the same as that used in Bourbaki [7], p. 1, for the term "espace vectoriel topologique." Our concept of a linear topological space is, therefore, broader than, e.g., that used by Bourgin [9], where the additional assumption is made that the space satisfies the Hausdorff separation axiom (i.e., that it is a T_2 space).

On the other hand, there are authors (e.g., Hille [20]) who use a concept broader than ours by relaxing slightly the nature of the continuity requirement for the algebraic operations, the continuity being required in each argument separately, but not necessarily jointly. Many of our results remain valid for this broader class of spaces.

customary Euclidean distance (metric) topology. Since any linear system must contain all scalar multiples of all its elements, the scalars used in a linear system being real numbers, the set of all integers (or even the set of all rational numbers) is not a linear system, hence not a linear topological space. From the economist's viewpoint this rules out applications involving indivisibilities.

II.1.1. *Linear topological spaces.* A linear topological space is both a linear system and a topological space. To avoid ambiguities, and for the sake of completeness, we supply some of the standard information concerning these concepts.

II.1.1.1. *Linear systems.* What we call a linear system is a purely algebraic concept. A fuller label would be "real linear system" since the scalars used are the reals. (Banach uses the term "linear space," Bourbaki "vector space"; our usage of the term "linear system" agrees with Hille's.) We shall find it convenient to refer to the elements of a linear system as *vectors.*

Since a linear system is an *additive group,* we start by defining the latter. A set $\mathscr{X}$ is called an additive group if it satisfies the following conditions:

1. With each pair (x', x'') of elements of $\mathscr{X}$ is associated a unique element x of $\mathscr{X}$; x is called the *sum* of x' with x'' and this is written as $x = x' + x''$.

2. Addition is associative; i.e., given any three elements x', x'', x''' of $\mathscr{X}$, $x' + (x'' + x''') = (x' + x'') + x'''$.

3. There is in $\mathscr{X}$ an element (the identity element of addition, later called the origin) denoted by 0_x (or, more simply, by 0) such that $x + 0_x = 0_x + x = x$ for every element x of $\mathscr{X}$.

4. To each element x of $\mathscr{X}$ corresponds uniquely an element $-x$ (called the negative of x) such that $x + (-x) = 0_x$. [Subtraction is defined by the relation $x' - x'' = x' + (-x'')$.] The foregoing conditions imply that the law of cancellation holds, i.e., that

$$x' + x = x'' + x \text{ implies } x' = x''$$

for any three elements x', x'', x of the group.

An additive group is called *commutative* (*Abelian*) if it satisfies the following additional condition:

5. $x' + x'' = x'' + x'$ for any two elements of the group.

A linear system is a commutative additive group in which there is further an operation of *scalar multiplication* (by reals, which we shall often call scalars). I.e.,

6. With each pair (α, x) where α is a scalar (real) and x a vector (an element of $\mathscr{X}$), there is associated a unique vector x', called their scalar product; this is written as $x' = \alpha \cdot x$. [Scalar multiplication is

commutative, i.e., $\alpha \cdot x = x \cdot \alpha$; the multiplication symbol $(\cdot)$ is often omitted.]

7. Scalar multiplication is distributive with regard to both scalars and vectors, i.e.,

$$(\alpha' + \alpha'')x = \alpha' x + \alpha'' x$$

and

$$\alpha(x' + x'') = \alpha x' + \alpha x''$$

for all selections of the scalars and vectors.

8. Scalar multiplication is associative, i.e., $\alpha'(\alpha'' x) = (\alpha' \alpha'')x$ for all selections of scalars and vectors.

9. The number one is the identity element of scalar multiplication, i.e., $1 \cdot x = x$ for all vectors x.

The preceding conditions imply that $(-1) \cdot x = -x$ and $0 \cdot x = 0_x$. [The last equation is an example of a situation where both the number 0 (zero) and the vector 0_x (origin) appear together. This is sometimes written simply as $0 \cdot x = 0$ and one must infer from the context that 0 denotes a scalar on the left and a vector on the right.]

Definition. A *linear system* is a set satisfying condition 1–9 above, i.e., an additive commutative group with scalar multiplication which is commutative, distributive, and associative, with reals as scalars and 1 as the identity element of scalar multiplication.

Algebraic set operations. Let X, X', X'' be subsets of a linear system and α a scalar (a real number). We write

$$\alpha X = \{\alpha x : x \in X\} ,$$
$$X' + X'' = \{x' + x'' : x' \in X', x'' \in X''\} ,$$
$$X' - X'' = \{x' - x'' : x' \in X', x'' \in X''\} .$$

Also,

$$-X = (-1)X = \{-x : x \in X\} .$$

These algebraic operations must be distinguished from the set-theoretic operations of union and difference. The union of two sets X' and X'' is written as $X' \cup X''$; the set-theoretic difference (i.e., the set of all elements that are in X' but not in X'') is written as $X' \sim X''$. The complement of X (with respect to X') is the difference $X' \sim X$.

We should also note that the algebraic operations do not have some of the properties suggested by the symbolism; e.g., it need not be true that $X + X = 2X$.

Some geometric terms. Given two vectors x', x'', the set $\{\lambda x' + (1 - \lambda)x'' : 0 \leq \lambda \leq 1\}$ is called the *segment joining* x' and x''. A set is called *convex* if with any two points x', x'' it also contains all points of the segment joining them. If $-X = X$, the set X is called *symmetric* (with respect to the origin). X is said to be *star-shaped from* the point x if, with any point x', it also contains the segment joining x and x'.

A subset X of the linear system $\mathscr{X}$ is called *absorbing* if, given any point x in the system $\mathscr{X}$, there is a point x' in the set X and a positive real number λ such that $x = \lambda x'$.[16]

II.1.1.2. *Topological spaces.* To define a topological space[17] it is convenient to start by introducing the concept of "a topology." A collection S of subsets of a given set A is called *a topology for A* if it satisfies the following conditions: (1) A is an element of S and so is the empty set ϕ; (2) the intersection of any two sets belonging to S belongs to S; (3) the union of the members of any (possibly infinite) sub-collection of S belongs to S. The subsets of A which belong to S are called open (relative to S, or in S). The union of all open sets contained in a given set is called its *interior*.

We *topologize* a set by selecting a topology for it. Any set can be topologized, for the two-element collection $\{\phi, A\}$ is a topology for A, i.e., it satisfies the above three conditions; such a two-element topology will be referred to as the *coarse* topology for A; it is sometimes called in the literature the indiscrete or trivial topology. On the other hand, the power set (sometimes written 2^A) of A, i.e., the set of all subsets of A, is also a topology for A, to be called the *fine* (often called discrete) topology for A. When A has two or more elements, the two topologies differ; for instance, one-element sets are open in the fine topology, but not in the coarse topology. Given two topologies for a set A, we call S' *finer than* S'' (and S'' *coarser than* S') if S'' is a proper subset of S', i.e., if every set open in S'' is also open in S' and there are some sets open in S' that are not open in S''. (Two topologies are non-comparable with respect to fineness when neither is a subset of the other.) Clearly, the fine topology is the finest topology possible, while the coarse topology is the coarsest topology possible. In most cases of applied interest, we deal with topologies that are somewhere between the fine and the coarse topologies.

Denote by $R^\sharp$ the linear system whose elements are all real numbers, i.e., the "real line." Its so-called "natural" topology is defined as consisting of all subsets B of $R^\sharp$ characterized by the following property: each element of B must belong to an "open interval" which is a subset of B. (An open interval is defined as the set of all numbers greater than some fixed number and less than another fixed number; an open interval is an open set in the natural topology, but there are open sets which are not open intervals, e.g., the set of all numbers other than zero.) A set is *closed* (in a specified topology) if its complement (with

[16] This usage of the term *absorbing*, as well as some of the subsequent formulation, is due to the author's exposure to lectures by Professor Hans Radstrom of the Royal Institute of Technology in Stockholm, to whom the author is also indebted for clarification on certain properties of linear spaces.

[17] See, for instance, Kelley [23].

respect to A) is open. A set may be both open and closed (e.g., the empty set and A), or it may be neither open nor closed (e.g., one-element sets in the coarse topology when A has two or more elements). What is ordinarily called a closed interval (i.e., one including its end-points) is a closed set in the natural topology of the real line. An interval including only one of its end-points is neither open nor closed in the natural topology. The *closure* of a set is the intersection of all closed sets containing it.

A *topological space* is defined formally as an ordered pair (A, S) where S is a topology for A. Often, when the topologization of A is understood, we refer to A itself as a topological space.

Let (A, S) be a topological space, B a subset of A, x an element of B. B is called a *neighborhood of* x (with respect to the topology S) if it contains a subset C which is open (with respect to the topology S). Obviously, any open set containing x is a neighborhood of x, but a neighborhood need not be open. (Some authors use a narrower concept of a neighborhood and require that it be an open set.) The collection of all neighborhoods of a given point x is called the *complete neighborhood system* for the point x. For instance, in the fine topology all sets of which x is an element constitute a complete neighborhood system for x; in particular, the one-element set consisting of x alone is a neighborhood of x. In the coarse topology, on the other hand, a point has only one neighborhood, namely, the set A. A topological space (A, S) is called a *Hausdorff* (topological) *space* if any two distinct points of the space have disjoint neighborhoods. Thus the fine topology is Hausdorff, but the coarse topology (when there are two or more elements in the space) is not. $R^{\#}$ in its natural topology is Hausdorff, for we can use as disjoint neighborhoods open intervals centered on the two points, the width of the intervals being less than half the distance of the two points. Most spaces of applied interest are Hausdorff.

Sometimes we are interested in certain subsets of the complete neighborhood system of a point. A subset F of the complete neighborhood system of a point is called a *fundamental system* of neighborhoods of the point if every neighborhood of the point contains a neighborhood belonging to the set F; if we call the neighborhoods belonging to F fundamental, we can say that every neighborhood of a point must contain a fundamental neighborhood of that point.

It is often convenient to define a topology indirectly, viz., by assigning to each point a of a set A a (non-empty) collection F_a and declaring it to be a fundamental neighborhood system of a. The complete neighborhood system of a is then defined as the collection G_a of subsets of A, each of which contains a fundamental set (i.e., a set belonging to F_a); finally a subset A' of A is declared as open if and only if it is a neighborhood of all of its points.

In order for such a procedure to result in a topology for A, the collection F_a must, of course, satisfy certain conditions. First, naturally, each set belonging to F_a must contain a, otherwise it would not qualify as a neighborhood of a; hence every set belonging to F_a is non-empty. Second, the collection F_a must satisfy the following finite intersection requirement: the intersection of any two sets belonging to F_a must contain a set belonging to F_a. (The intersection itself need not belong to F_a.) A non-empty collection F_a of sets each of which contains a and satisfying the preceding finite intersection requirement will be called a *neighborhood base* at a. We shall see that it is convenient to discuss the properties of linear topological spaces in terms of fundamental neighborhood systems and neighborhood bases.

It was mentioned earlier that a linear topological space is a set which is both a linear system and a topological space, with certain continuity conditions imposed on the algebraic operations of addition and scalar multiplication. To be able to state these conditions, we must introduce the concept of continuity.

Let A and B be two sets and let f denote a functional relation whose domain is A and whose range is B, i.e., which associates with each element a in A a unique element $b = f(a)$ in B. Given a subset A' of A, we define the *image* of A' by f as the set $\{f(a) \in B : a \in A'\}$. Given a subset B' in B, we define as the *inverse image* of B' by f the set $\{a \in A : f(a) \in B'\}$. The image of A' by f is denoted by $f(A')$; the inverse image of B' by f is denoted by $f^{-1}(B')$.

Now let us topologize A and B, with S denoting the topology for A, T the topology for B. The function f is said to be *continuous* if the inverse image $f^{-1}(B')$ of every set B' open in T is itself open (in S). It is important to realize that continuity depends not only on the nature of the function, but also on the manner in which the two spaces have been topologized. Thus if S is fine, any function on A is continuous. Similarly, the constant function (which has the same value for all elements of A) is continuous for any topology, since the inverse image of the one-element set (consisting of the constant f value) is the whole space A. Now suppose $A = B$ and f is the identity function, i.e., $f(a) = a$ for all a in A. (When $A = B = R^{\sharp}$, the identity function is represented by the positively inclined 45° straight line through the origin.) Whether f is continuous depends on the topologization of A and B. If A and B are given the same topologies (i.e., $S = T$), then f is continuous, since $f^{-1}(B') = B'$ for all B'. But, even though the sets A and B are the same, their topologies may differ. For instance, let $A = B = R^{\sharp}$, with f still the identity function, and let B have the natural topology while to A we give the coarse topology. Let B' be a finite open interval on the B-axis which is an open set in the natural topology. The inverse image of B' is the same interval, taken on the

A-axis; but, in the coarse topology of the real line, a non-empty proper subset of the line is not open; hence $f^{-1}(B')$ is not open (in S); hence with this topologization the identity function is not continuous.

One more topological concept is essential in discussing the properties of linear topological spaces. As was indicated earlier, the continuity of the operations in such a space is joint continuity in the two arguments. To clarify the point, consider the operation of scalar multiplication. We may write, for a scalar (real) α and a vector x, $\alpha x = g(\alpha, x)$, so that the scalar product may be viewed as a function of two variables α and x. In order to explain what is meant by the *joint* continuity of g in the two variables, we restate the situation as follows: First, we write $\alpha x = f((\alpha, x))$, i.e., we now view the scalar product as a function whose domain is the set of ordered pairs (α, x), i.e., the Cartesian product $R^\# \times \mathscr{X}$, while the range, of course, is the set $\mathscr{X}$. To apply the above definition of continuity, we must topologize the product set $R^\# \times \mathscr{X}$. Similarly, addition may be viewed as a function on the Cartesian product $\mathscr{X} \times \mathscr{X}$ with the range in $\mathscr{X}$. Here, again, the product set must be topologized.

In both cases, the appropriate topologization (i.e., the one implicit in the definition of a linear topological space) is the so-called product topology which we shall now define.

Let there be two topological spaces (A, S) and (B, T) and let $C = A \times B$. The *product topology*, about to be defined, will be denoted by $P[S, T]$; hence the topological product space is written as $(C, P[S, T])$. To define the product topology, it is enough to characterize the open sets of C. A set C' is open in the $P[S, T]$ topology if and only if every point c' of C' is a member of a set of the form $A' \times B'$ where A' is open (in S), B' is open (in T), and $A' \times B' \subseteq C'$. As an illustration, if $A = B = R^\#$ (the real line), so that C is the Cartesian plane (the set of all ordered pairs of real numbers), and $S = T =$ the natural topology for the reals, then the product topology $P[S, T]$ is the usual Euclidean topology of the plane where a set is open if every one of its points can be enclosed in a disk of positive radius wholly belonging to the set.

Another example is obtained if C is again $R^\# \times R^\#$ but $S = T =$ coarse topology; here the product topology is the coarse topology of the plane. Similarly, the topological product of fine topological spaces is fine. An interesting case is obtained if we take $A = B = R^\#$ but with different topologies on the two spaces, viz., A coarse and B natural. In the product topology a one-element set (a point) is not closed and the topology is not Hausdorff: every open set must contain an infinite "strip" (of positive width) parallel to the A-axis, and any neighborhood containing $(a', 0)$ must contain all other points of the form $(a, 0)$.

II.1.1.3. *Linear topological spaces*. We now have a sufficient vocabulary to provide a precise definition of a linear topological space.

Definition. Let X be a linear system and S a topology for X. (X, S) is said to be a *linear topological space* if (1) addition is continuous in the product topology $P[S, S]$, and (2) scalar multiplication is continuous in the product topology $P[N, S]$, where N denotes the natural topology of the real line.

To illustrate, let X be the real line $R^{\#}$. As might have been expected, $(R^{\#}, N)$, i.e., the real line in its natural topology, is a linear topological space; the real line in its coarse topology also turns out to be a linear topological space; the real line in its fine (discrete) topology is not a linear topological space, although it is a linear system and a topological space. Hence the continuity conditions in the above definition are not automatically satisfied for every linear system which is also a topological space.

The verification of the continuity properties of the algebraic operations directly from the topology of the space can be quite awkward; the situation becomes much more transparent when the properties of linear topological spaces are stated in terms of neighborhood systems. Furthermore, we may confine ourselves to the discussion of the neighborhood system of the origin 0_x; if G_{0_x} is a collection of neighborhoods of the origin, the corresponding collection of neighborhoods of any point x is given by $\{x\} + G_{0_x} = G_x$. Thus the topology of a linear topological space may be defined in terms of a fundamental neighborhood system of the origin, the corresponding complete system of neighborhoods then being defined as the collection of sets containing a fundamental set, and finally, an open set being defined as a set which is a neighborhood of each of its elements. But to follow such procedures we must know what types of neighborhoods one may encounter in linear topological spaces.

The answer to this question is contained in a theorem we shall state in a moment. To simplify this statement, we shall coin an *ad hoc* term; we shall call a non-empty family G of sets *acceptable* if it satisfies the following conditions: (1) if V is in the family G, then the family must also contain a set W such that $W + W \subseteq V$; (2) every set in the family is symmetric, i.e., $V = -V$ for all V in G; (3) every set of the family contains the origin, i.e., $0_x \in V$ for each V in G; (4) every set of the family is star-shaped from the origin, i.e., if a point x is in V, then so is the whole segment joining x to the origin; (5) every set in the family is absorbing, i.e., if x is any point of the space X and V is a set in the family G, then V has an element x' such that $x = \lambda x'$ for some positive number λ; (6) the family G is invariant under homotheties (from the origin), i.e., if V is a set in the family and α a real number different from zero, then the set αV is also in the family.

THEOREM (Bourbaki [7], Prop. 5, p. 7).

A. *If (X, S) is a linear topological space, then there exists an accepta-*

ble (i.e., satisfying conditions 1–6 above) fundamental neighborhood system of the origin.

B. *In a linear system X, let F be a neighborhood base at 0_x (i.e., a non-empty collection of sets each containing the origin and such that an intersection of any two members of the collection contains a member of the collection) and suppose that F is acceptable (i.e., satisfies conditions 1–6 above). Then there exists a topology (and only one topology) such that F is the fundamental neighborhood system of the origin in that topology. In this topology, X is a linear topological space.*

[The above six conditions are somewhat redundant, since 3 follows from the others. We have chosen this form, however, partly in order to show that a linear topological space is a linear topological group, i.e., an additive Abelian group with a topology in which addition and subtraction are both continuous (jointly in the two arguments). Conditions 1–3 above are precisely those characterizing a fundamental neighborhood system of the origin (identity element of addition) of a topological group. Cf. Bourbaki [6], p. 6.]

We can now verify our statements about the various topologizations of the real line. Thus for its coarse topology the neighborhood base at the origin consists of the single one-element set $\{R^{\sharp}\}$. It may be seen that this base is acceptable (i.e., satisfies conditions 1–6). For the natural topology of the real line we use the family of all open intervals centered on 0; again, the family is acceptable. Hence $R^{\sharp}$ is indeed a linear topological space in both the coarse and the natural topology. But the situation is different when $R^{\sharp}$ is given its fine (discrete) topology. Since the one-element set consisting of the origin is open in this topology, it is a neighborhood and hence any fundamental neighborhood system of the origin must contain $\{0\}$. However, $\{0\}$ is not an absorbing set (i.e., condition 6 of acceptability is violated) and hence $R^{\sharp}$ in its fine topology does not have an acceptable fundamental system; hence it is not a linear topological space.

A linear topological space satisfying the Hausdorff separation axiom (distinct points have disjoint neighborhoods) is called a *Hausdorff linear* (topological) *space*. It will be noted that the real line, depending on its topologization, may fail to be a linear topological space (in the fine topology), it may be a Hausdorff linear space (in the natural topology), or it may be a non-Hausdorff linear topological space (in the coarse topology). Euclidean finite-dimensional spaces are all Hausdorff linear.

A linear topological space may or may not possess a fundamental neighborhood system of the origin consisting of convex neighborhoods. If, in a linear topological space, there exists such a fundamental system consisting of convex neighborhoods (i.e., every fundamental neighborhood is convex), the space is called a *locally convex* (linear topological) *space*. We may note that the real line forms a locally convex space in both its natural and its coarse topology. This is not accidental: according to a theorem due to Tychonoff (cf. [43], p. 769) every finite-dimensional linear topological space is locally convex. [A linear system is finite-dimensional, say n-dimensional, if there exists a finite set of elements $x_1, x_2, \cdots, x_n$ such that every element x of X can be written in the form $x = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$ where the α_i are scalars (reals).] Furthermore, if a space is finite-dimensional and Hausdorff linear, then its topology is Euclidean.

Most linear topological spaces occurring in applications are locally convex, but there do exist linear topological spaces that are not locally convex. Tychonoff's example (*loc. cit.*, p. 768) is the space denoted by $l_{1/2}$ consisting of all the infinite sequences $x = (x_1, x_2, \cdots)$ of numbers x_i such that

$$\sum_{i=1}^{\infty} |x_i|^{1/2} < \infty .$$

The space $l_{1/2}$ is topologized in the following manner. We construct a fundamental neighborhood system of the origin consisting of sets of the form

$$\{x : (\sum_{i=1}^{\infty} |x_i|^{1/2})^2 < \rho\}$$

with ρ varying over positive reals. It was shown by Tychonoff that (a) this space is a linear topological space, but (b) there does not exist a fundamental neighborhood system of the origin consisting of convex sets. (The fact that the fundamental system as given does not consist of convex sets is by itself inconclusive, since there might exist another fundamental system consisting of convex sets and yielding the same topology.) Hence $l_{1/2}$ is a non-locally convex linear topological space. On the other hand, the space is Hausdorff linear; this can be shown by utilizing the fact that the function

$$q(x) = (\sum_{i=1}^{\infty} |x_i|^{1/2})^2$$

satisfies the inequality $q(x' + x'') \leqq 2[q(x') + q(x'')]$. Let x, y be two distinct elements of the space such that $q(x - y) = a$, where a is necessarily positive. Now select a neighborhood of x to be the set of points such that $q(x - z')$ is less than $a/6$; similarly, select a neighborhood of y consisting of points such that $q(y - z'')$ is less than $a/6$. Suppose that there is a point z belonging to both neighborhoods. Then, in virtue of the inequality, $q(x - y) \leqq 2[a/6 + a/6]$ which contradicts the assumption made. Hence the space does satisfy the Hausdroff separation axiom.

We have had examples of spaces that are both Hausdorff and locally convex (real line in its natural topology), Hausdorff but not locally convex (the space $l_{1/2}$), locally convex but not Hausdorff (real line in its coarse topology). To complete the picture let us point out that the topological product of $l_{1/2}$ with the real line in its coarse topology is neither Hausdorff nor locally convex, although it is a linear topological space.

A fundamental neighborhood system of the origin in a locally convex space can always be defined (Bourbaki [7], pp. 95–96) by means of a set of functions called *semi-norms*. A semi-norm is defined as a finite real-valued function p on a linear system satisfying the following two requirements: (1) for any scalar α and any vector x, $p(\alpha x) = |\alpha| \cdot p(x)$; (2) for any two vectors, $p(x' + x'') \leqq p(x') + p(x'')$. It follows that $p(0_x) = 0$ and $p(x)$ is always non-negative. A semi-norm is called a *norm* if it has the further property (3) $p(x) = 0$ only if $x = 0_x$. Hence, for a norm, $p(x) = 0$ if, and only if, $x = 0_x$. The norm of x is usually written $\|x\|$. For instance, let X be the linear system consisting of all ordered pairs (x_1, x_2) of real numbers (the Cartesian plane). Then $p^*(x) = |x_1| + |x_2|$ is a norm, while $p^{**}(x) = |x_1|$ is a semi-norm, but not a norm.

If p is a semi-norm, the set $\{x : p(x) < \lambda\} = [p ; \lambda]$ is called an *open strip* (of width 2λ). Denote by $[p]$ the set of all open strips $[p ; \lambda]$, obtained by keeping p fixed while λ varies over the positive reals. Given a set P of semi-norms, we shall denote by $[P]$ the set of all open strips, obtained by taking all the sets $[p ; \lambda]$ with λ varying over the positive reals and p over P. Finally, let F(P) denote the set of all finite intersections of members of $[P]$. The set $F(P)$ is a fundamental neighborhood system of the origin, as can be verified from the "acceptability" conditions 1–6 above; also, the elements of $F(P)$ are convex sets, since all strips are convex and so are their intersections. Hence $F(P)$ defines a locally convex topology for the linear system on which the semi-norms are defined. Conversely (Bourbaki [7], p. 96, Prop. 4), every locally convex topology can be defined by a fundamental set $F(P)$ for a suitably chosen set of semi-norms P.

A *normed space* is a linear system where the fundamental neighborhood system of the origin consists of sets (open spheres) $S(\rho) = \{x : \|x\| < \rho\}$ where ρ is the *radius* of the *sphere*. (The spheres are centered at the origin.) The system consists of spheres with the radius varying over the positive reals, although a smaller system (e.g., with rational radii) would be sufficient. That a normed space is a locally convex linear space follows from the fact that the spheres constitute an "acceptable" family (i.e., satisfy conditions 1–6 above) and are convex sets. Also, a normed space is Hausdorff. The proof proceeds exactly as in the case of the space $l_{1/2}$ above, except that the relevant inequality does not have the factor 2 on the right-hand side.

A linear topological space is called *normable* if its topology can be defined by a norm as just indicated. From what has just been said it follows that a normable space must be locally convex Hausdorff. However, not every locally convex Hausdorff linear space is normable. Because of the convenience in dealing with normed spaces, it is of interest to know under what conditions a space is normable. In order to do so, we must introduce a new concept, that of a *bounded* subset of a linear topological space. A subset B is said to be bounded if, given *any* neighborhood V of the origin, there is a positive scalar λ such that $B \subseteq \lambda V$; this is expressed by saying that a bounded set is *absorbed* by every neighborhood.

We may now state Kolmogoroff's theorem on normability of linear topological spaces: a linear topological space is normable if and only if it is locally convex Hausdorff and there exists a bounded neighborhood of the origin.

The following is an example of a non-normable locally convex Hausdorff space. Its elements are all the infinite numerical sequences $x = (x_1, x_2, \cdots)$. The space is topologized by the set $P = \{p^1, p^2, \cdots\}$ of semi-norms where $p^d(x) = \max(|x_1|, |x_2|, \cdots, |x_d|)$. Its topology, being based on the family $F(P)$ as the fundamental neighborhood system of the origin, is necessarily a locally convex linear space. It is also Hausdorff because for each element x other than 0_x of the space there exists a norm p in P such that $p(x) \neq 0$. (Cf. Bourbaki [7], p. 97, Prop. 5.) Now if this space were normable, there would exist a bounded neighborhood of the origin; hence, by definition of a fundamental system, there would exist a bounded set of the family $F(P)$, since a subset of a bounded set is bounded. Hence to establish the non-normability, it is enough to show that no member of the family $F(P)$ is bounded. Now the members of the family $F(P)$ are formed by finite intersections of the open strips defined by the norms p^d. Hence it is true for each member of $F(P)$ that, starting with, say, the k-th component, the values of the components with subscripts $\geqq k$ are completely unrestricted. Now let V_k be a member of $F(P)$, with the components whose subscripts $\geqq k$ are unrestricted, while the components 1 through $k-1$ cannot exceed M $(0 < M < \infty)$ in absolute value. We show that V_k is not absorbed by a neighborhood V_{k+1}. This follows from the fact that in V_{k+1} the $(k+1)$th component is restricted, while in V_k it is not. Hence, no matter what $\lambda > 0$ we choose, there will be elements in V_k that are not in λV_{k+1}. Hence V_k is not bounded; but since V_k is a typical member of $F(P)$, no set in $F(P)$ is bounded. By the previous argument it follows that there is no bounded neighborhood of the origin, and hence the space is not normable.

Many spaces we deal with are normed; in particular, the finite-dimensional Euclidean spaces are normed. The norm of a point x in a Euclidean n-dimensional space can be defined in various ways. The *Euclidean* norm of x is defined as

$$\left(\sum_{i=1}^{n} x_i^2\right)^{1/2};$$

another norm (which results in the same topology) can be defined as $\max(|x_1|, |x_2|, \cdots, |x_n|)$.

The space of all infinite sequences $x = (x_1, x_2, \cdots)$ of numbers x_i with only finitely many components different from zero can be normed by defining

$$||x|| = \left(\sum_{i=1}^{\infty} x_i^2\right)^{1/2}.$$

In a normed space it is possible to define a *distance* function $d(x', x'') = ||x' - x''||$. In any space on which a distance function has been defined one can introduce a "metric" topology, by using as the fundamental neighborhood system for x_0 the (metric) spheres, i.e., the sets $\{x : d(x, x^0) < \rho\}$ with the radius ρ varying over positive reals. Because of the triangle inequality satisfied by the distance function a metric space is always a Hausdorff topological space, the proof being analogous to that sketched above for the normed and $l_{1/2}$ spaces.

Let $(x^1, x^2, \cdots)$ be an infinite sequence of points x^i in a metric space with a distance function d. The sequence is said to be a *Cauchy sequence* if, given $\varepsilon > 0$, there exists a positive integer N such that $d(x^m, x^n) < \varepsilon$ provided both m and n are greater than N. A sequence $(x^1, x^2, \cdots)$ is said to converge to x^0 if, for any $\varepsilon > 0$, there exists a positive integer N such that $d(x^n, x^0) < \varepsilon$ provided n is greater than N. A sequence is said to be *convergent* if it converges to some element x^0 of the space. It is known that every convergent sequence is a Cauchy sequence. On the other hand, there are spaces with non-convergent Cauchy sequences. One example of such a space is that of infinite sequences with only finitely many components different from zero. A space where every Cauchy sequence is convergent is called *complete*. The reals are complete in their natural topology, while the rationals with the same topology (i.e., defined by the Euclidean distance or norm) form a space that is not complete because there are sequences of reals converging to an irrational number. A normed space which is also complete is called a *Banach space*. Thus the reals (as well as all finite-dimensional Euclidean spaces) are Banach, but the above space of infinite sequences with only finitely many non-zero components is not Banach, though normed. The classic example of an infinite-dimensional Banach space is the space l_2 of all infinite numerical sequences $x = (x_1, x_2, \cdots)$ such that

$$\sum_{i=1}^{\infty} x_i^2 < \infty,$$

the norm being defined as the square root of the preceding infinite sum;

l_2 belongs to a sub-class of Banach spaces known as *Hilbert spaces*, where with each pair of elements it is possible to associate a number called their *inner product* $x' \cdot x''$, with $x' \cdot x''$ linear in each of its arguments, $x' \cdot x'' = x'' \cdot x'$, $x \cdot x$ always non-negative, and $x \cdot x = 0$ if and only if $x = 0_x$. In a space with such an inner product it is possible to define the norm of a vector as $||x|| = (x \cdot x)^{1/2}$, and the resulting normed space is called Hilbert if it is Banach, i.e., if it is complete. (Some authors use somewhat different definitions of a Hilbert space.) According to this definition the Euclidean spaces are Hilbert, with the inner product defined as

$$\sum_{i=1}^{n} x_i' x_i'' .$$

The resulting norm is, of course, that corresponding to the Euclidean distance.

As an example of a Banach space which is not a Hilbert space we may take the space of all infinite bounded numerical sequences $x = (x_1, x_2, \cdots)$. The norm of this space is defined by $||x|| = \sup(|x_1|, |x_2|, \cdots)$.

II.1.2. *Linear transformations.* A function T whose domain is a linear system $\mathscr{X}$ and the range a subset of a linear system $\mathscr{Y}$ is called a *linear transformation* on $\mathscr{X}$ into $\mathscr{Y}$ if it is additive and homogeneous, i.e., if

$$T(x' + x'') = T(x') + T(x'') \qquad \text{for all } x', x'' \text{ in } \mathscr{X} ,$$

and

$$T(\alpha x) = \alpha T(x) \qquad \text{for all real } \alpha \text{ and all } x \text{ in } \mathscr{X} .$$

If both spaces are linear topological, an additive continuous function is homogeneous (Hille [20], Theorem 2.6.1, p. 16). The converse, however, is not true; i.e., there are (in infinite dimensional spaces) linear transformations which are not continuous at any point (see Bourbaki [7], p. 93).

If both spaces are Banach, an additive function is continuous if and only if it is bounded (i.e., carries bounded sets into bounded sets). Hence, in Banach spaces, "linear bounded" as applied to transformations is synonymous with "linear continuous."

A linear transformation on $\mathscr{X}$ whose range is a subset of the reals (i.e., a homogeneous additive real-valued function on $\mathscr{X}$) is called a *linear functional* on $\mathscr{X}$. There are linear functionals on locally convex spaces that are not continuous at any point (Bourbaki [7], p. 93). Since a linear functional is both convex and concave, it follows that, even in locally convex spaces, a convex or concave function need not be continuous. This is of interest in connection with results of this chapter where only concavity, but not continuity, of a function is assumed, since it proves that the concavity assumption is less restrictive; the same remark applies to results where only linearity, but not continuity, of transformations is assumed.

II.1.3. *The conjugate space.* Let $\mathscr{X}$ be a linear topological space. The set $\mathscr{X}^*$ of all linear *continuous* (with respect to the natural topology of the reals) functionals on $\mathscr{X}$ is called the *conjugate* (adjoint, dual) space of $\mathscr{X}$. $\mathscr{X}^*$ is a linear system whose typical element will be written as x^*; the null element (origin) of $\mathscr{X}^*$ (i.e., the real-valued function on $\mathscr{X}$ whose value is zero for each element of $\mathscr{X}$) is denoted by 0_x^* or 0, as the occasion demands.

Two ways of topologizing the conjugate space are of particular interest. They are respectively labeled "strong" and "weak star," the latter usually being written "weak*."

In each case the topology is defined through a fundamental neighborhood system.

In the *strong* topology a fundamental neighborhood of 0_x^* is of the form

$$U(\varepsilon, B) = \{x^* \in \mathscr{X}^* : |x^*(x)| < \varepsilon \quad \text{for all } x \in B\}$$

where B is a *bounded* set, taking all neighborhoods $U(\varepsilon, B)$ with ε varying over the positive reals and B over the class of *all bounded* sets in $\mathscr{X}$.

In the *weak* topology* a fundamental neighborhood of 0_x^* is also of the form

$$U(\varepsilon, B) = \{x^* \in \mathscr{X}^* : |x^*(x)| < \varepsilon \quad \text{for all } x \in B\} ,$$

but B here is required to be a *finite* set; the fundamental system is again obtained by letting ε vary over positive reals and B over the class of *all finite* sets in $\mathscr{X}$.

Since every finite set is bounded, it follows that every weak* neighborhood is also a strong neighborhood, but there may be strong neighborhoods that are not weak* neighborhoods. It follows that the strong topology is at least as fine as, and possibly finer than, the weak* topology. I.e., every set open (resp. closed) in the weak* topology is also open (resp. closed) in the strong topology, but the converse need not be true.

In both topologies the conjugate space is a Hausdorff locally convex linear topological space (Bourbaki [8], pp. 16–19). Moreover, when the space $\mathscr{X}$ is normed, the conjugate space is *normable* in its *strong* topology, the norm of an element x^* of the conjugate space being defined by

$$||x^*|| = \sup_{||x|| \leq 1} |x^*(x)| .$$

In its strong (norm) topology, the conjugate of any normed space is complete, hence it is a Banach space.

In finite-dimensional Euclidean spaces, the strong and weak* topologies coincide. But in infinite-dimensional spaces the strong topology is, in most cases likely to be considered, actually finer than the weak* topology.

In particular, if $\mathscr{X}$ is an infinite-dimensional normed space, the strong topology is finer than the weak* topology. (Cf. Bourbaki [8], p. 111, where it is shown that the set of elements of norm one in the conjugate space is not closed in the weak* topology, although it is closed in the strong topology.)

II.1.4. *Separation by hyperplanes in linear topological spaces.*

II.1.4.1. Let $\mathscr{X}$ be a linear system. A subset of $\mathscr{X}$ is called *linear* if it is closed under the operations of addition and scalar multiplication. A translate of a linear set M, i.e., a set of the form $\{x_0\} + M$ where M is a linear set, is called a (linear) *variety*.[18] If M is a linear set such that M is a proper subset of $\mathscr{X}$ and there is no linear proper subset of $\mathscr{X}$ in which M is contained, M is called a *maximal* linear set. A translate of a maximal linear set is called a *maximal variety.*

With each maximal variety V one may associate a non-null (i.e., $\neq 0_x^*$) linear functional x^* on $\mathscr{X}$ and a real number α such that $V = \{x \in \mathscr{X} : x^*(x) = \alpha\}$. On the other hand, every pair (x^*, α) where x^* is a linear functional and α a real number defines a maximal variety.

If $\mathscr{X}$ is a linear topological space, a maximal variety may or may not be a closed set. We shall call a closed maximal variety a *hyperplane.* (Terminologies of various writers differ. In Bourbaki, hyperplane is synonymous with a maximal variety.) In a linear topological space $\mathscr{X}$, a maximal variety $V = \{x \in \mathscr{X} : x^*(x) = \alpha\}$, where x^* is a linear functional and α a real number, is closed if and only if x^* is continuous. I.e., a maximal variety is a hyperplane if and only if the functional defining the variety is continuous.

We may now state a theorem underlying a great many results concerning convex sets in linear topological spaces. The theorem is variously called the Hahn-Banach Theorem (geometric form) (cf. Bourbaki [7], p. 69) and the Bounding Plane Theorem.

THEOREM II.1. *Let $\mathscr{X}$ be a linear topological space, A an open convex (non-empty) subset of $\mathscr{X}$, and M a linear variety disjoint from A (i.e., $A \cap M = \phi$). Then there exists a hyperplane H containing M and disjoint from A (i.e., $M \subseteq H$ and $H \cap A = \phi$).*

Hence, under the hypotheses of the Theorem there exists a *continuous* linear functional x^* and a real α such that $x^*(x) = \alpha$ for all x in M and $x^*(x) < \alpha$ for all x in A.

In what follows we shall need the following

COROLLARY II.1. *Let $\mathscr{X}$ be a linear topological space and A a convex subset with non-empty interior. Then, for any point x_0 of $\mathscr{X}$ which is not in the interior of A, there exists a continuous linear functional x_0^* such that $x_0^*(x) \leqq x_0^*(x_0)$ for all x in A.*

[18] In particular, every point of the space, viewed as a one-element set, is a linear variety.

The geometric interpretation of the preceding Corollary is that through every point not in the interior of A there is a hyperplane "bounding" the set A, provided A is convex and has a non-empty interior.

In certain contexts, however, we want a somewhat stronger separation property. Given a set A and a point x_0 outside the set, we are interested in the existence of a hyperplane such that A is wholly on one "side" of it (possibly touching H) while x_0 is on the other "side" (not touching H). I.e., we are looking for a continuous functional x_0^* such that

$$\sup_{x \in A} x_0^*(x) < x_0^*(x_0) .$$

It is intuitively clear that we shall have to require that A be a closed convex set. But it turns out that restrictions must also be imposed on the nature of the linear topological space. The desired result follows from Prop. 4 in Bourbaki [7], p. 73. It was established by Mazur for Banach spaces and by Bourgin for Hausdorff locally convex spaces; we shall refer to it as the Mazur-Bourgin Theorem.

THEOREM II.2. (Mazur-Bourgin.) *Let $\mathscr{X}$ be a locally convex linear topological space, A a (non-empty) convex closed subset of $\mathscr{X}$, and x_0 a point outside A, i.e., $x_0 \notin A$. Then there exists a hyperplane "strictly separating" x_0 from A, i.e., there exists a continuous linear functional x_0^* such that the inequality*

$$\sup_{x \in A} x_0^*(x) < x_0^*(x_0) \tag{1}$$

holds.

Following Bourgin,[19] we shall refer to a set that can be "strictly separated" from points not in it as *regularly° convex*. Hence the preceding theorem states that in a locally convex space closed convex sets are regularly° convex. (Also, it is the case that a regularly° convex set is closed and convex.) It may be noted, however, that the class of spaces in which a closed convex set is regularly° convex is wider than that of locally convex spaces, as shown by Klee ([25], (10.1), p. 459). This is of interest since the regular° convexity of certain sets is a crucial property in several results of this chapter. If spaces in which closed convex sets are regularly° convex are called c-regular (as suggested by E. Michael, see Klee [26], p. 106), we may note here that many of the results of this chapter which presuppose local convexity of the space are valid for all c-regular spaces. However, this additional generality does not seem of serious applied interest in our problems.

On the other hand, we may in some cases wish to ensure the regular° convexity of certain sets without restricting ourselves to locally convex spaces. This can be accomplished by imposing an additional requirement

[19] In a slightly modified fashion: what we call regularly° convex (regularly circle-convex) he calls regularly $\mathscr{X}$ convex (where $\mathscr{X}$ is the underlying space).

on the nature of the set A, viz., that it possess a non-empty interior (see, for instance, Klee [25], Theorem 9.7, p. 456). However, the assumption of a non-empty interior rules out certain worth-while applications. Specifically, the sets in whose regular° convexity we are interested are those consisting of the vectors with non-negative coordinates (the non-negative cones); in a Euclidean space of finite dimension such a set (the non-negative orthant) does have an interior, but in infinite-dimensional spaces this is not always the case. In particular, for the l_p spaces ($p \geq 1$), the non-negative cone has no interior points (cf. Klee [24], p. 771); in other spaces, such as the space (m) of infinite sequences, the non-negative cone does have interior points.

Let $\mathscr{X}$ be a linear topological space, $\mathscr{X}^*$ its conjugate space. Given any element x_0 of the space $\mathscr{X}$, we can define a functional f_{x_0} on the conjugate space $\mathscr{X}^*$ by the relation

$$f_{x_0}(x^*) = x^*(x_0) \qquad \text{for all } x^* \in \mathscr{X}^* .$$

It may be verified that f_{x_0} is additive and homogeneous, hence linear. Now it may be noted that f_{x_0} is a continuous functional on $\mathscr{X}^*$ if $\mathscr{X}^*$ is given its weak* topology; in fact, the weak* topology is the coarsest topology for which all functionals f_x are continuous. Since the strong topology of the conjugate space is finer than (or at least as fine as) the weak* topology, it follows that the functionals f_x are also continuous when $\mathscr{X}^*$ is given its strong topology. Hence the set of all functionals f_x obtained by letting x vary over the whole space $\mathscr{X}$ is a subset of the conjugate of $\mathscr{X}^*$, whether the latter has the weak* or the strong topology. When the set of all f_x (as x varies over $\mathscr{X}$) equals the conjugate of $\mathscr{X}^*$, we call $\mathscr{X}$ *reflexive.* (For instance, the Euclidean spaces are reflexive and so is l_2.) Let $\mathscr{X}$ be a linear topological space and $\mathscr{X}^*$ its conjugate. A subset X^* of $\mathscr{X}^*$ is said to be *regularly convex* (this is not to be confused with the notion of regular° convexity defined earlier) if, given an element x_0^* not in X^*, there exists an element x_0 of the underlying space $\mathscr{X}$ such that

$$\sup_{x^* \in X^*} x^*(x_0) < x_0^*(x_0) . \tag{2}$$

The relation (2) can be understood more easily if we rewrite it as

$$\sup_{x^* \in X^*} f_{x_0}(x^*) < f_{x_0}(x_0^*) \tag{2'}$$

where f_{x_0} is defined as above. Now f_{x_0} is a continuous functional on $\mathscr{X}^*$, as just shown, whether the topology of the conjugate space is weak* or strong. Hence (2′) demands that it be possible to strictly separate X^* from a point x_0^* outside of it by a hyperplane (in either topology) and, furthermore, that the hyperplane be of the type defined by an f_x functional. Now we know that in either topology the conjugate space is locally convex; hence, provided X^* is convex and closed, there always exists *some* hyperplane strictly separating the point and the set (by the Mazur-Bourgin Theorem). However, it does not follow that the separating hyperplane will be *of the f_x type*, i.e., determined by an element of the underlying space. It is therefore noteworthy that, as shown by Bourgin ([9], Theorem 18, p. 655), if $\mathscr{X}$ is a Hausdorff[20] linear space, X^* is regularly convex, if and only if it is convex and closed in the weak* topology. Of course, X^* is closed in the strong topology if it is closed in the weak* topology.

II.1.4.2. *Regularly convex envelope.*

LEMMA II.1. *Let $\mathscr{A}$ be a collection of regularly convex subsets A of $\mathscr{W}^*$ and assume the intersection*

$$I_{\mathscr{A}} = \bigcap_{A \in \mathscr{A}} A$$

of these sets to be non-empty. Then $I_{\mathscr{A}}$ is also regularly convex.[21]

PROOF. Let $w_0^* \notin I_{\mathscr{A}}$. Then $w_0^* \notin A_0$ for some $A_0 \in \mathscr{A}$. Since A_0 is regularly convex, there exists $w_0 \in \mathscr{W}$ such that

$$\sup_{w^* \in A_0} w^*(w_0) < w_0^*(w_0) .$$

Now $I_{\mathscr{A}} \subseteq A_0$, so that

$$\sup_{w^* \in I_{\mathscr{A}}} w^*(w_0) \leqq \sup_{w^* \in A_0} w^*(w_0) ,$$

hence,

$$\sup_{w^* \in I_{\mathscr{A}}} w^*(w_0) < w_0^*(w_0)$$

and the conclusion of the Lemma follows.

If $B \subseteq \mathscr{W}^*$, we denote by $\tilde{B}$ the intersection of all regularly convex sets in $\mathscr{W}^*$ containing B. By the preceding Lemma, $\tilde{B}$ is regularly convex; it is called the *regularly convex envelope* of B.

Clearly, $B = \tilde{B}$ if and only if B is regularly convex.

II.1.5. If T is a linear continuous transformation on $\mathscr{X}$ into $\mathscr{Y}$ (where both $\mathscr{X}$ and $\mathscr{Y}$ are topological linear spaces), we may define a functional φ on $\mathscr{X}$ by the relation

$$\varphi(x) = y_0^*[T(x)] \qquad \text{for all } x \in \mathscr{X} ,$$

where y_0^* is a fixed element of $\mathscr{Y}^*$, i.e., a linear continuous functional on $\mathscr{Y}$. We have

$$\varphi(\alpha x) = y_0^*[T(\alpha x)] = y_0^*[\alpha T(x)] = \alpha y_0^*[T(x)] = \alpha\varphi(x)$$

and

$$\begin{aligned}\varphi(x' + x'') &= y_0^*[T(x' + x'')] = y_0^*[T(x') + T(x'')] \\ &= y_0^*[T(x')] + y_0^*[T(x'')] = \varphi(x') + \varphi(x'') .\end{aligned}$$

[20] It may be shown that the restriction to Hausdorff spaces may be removed.

[21] This is stated for Banach spaces in Krein and Šmulian [30], p. 556.

Hence φ is linear; φ is also continuous,[22] hence it is an element of $\mathscr{X}^*$ and may be denoted by $x^*_{y^*_0}$.

Consider now the functional relation associating with each $y^* \in \mathscr{Y}^*$ the corresponding $x^*_{y^*} \in \mathscr{X}^*$, as just defined. This relation is denoted by T^* and is called the *adjoint* of T. We write

$$x^*_{y^*} = T^*(y^*) \qquad (x^*_{y^*} \in \mathscr{X}^*, y^* \in \mathscr{Y}^*),$$

where

$$x^*_{y^*}(x) = y^*[T(x)] \qquad \text{for all } x \in \mathscr{X}.$$

We note that, for all $x \in \mathscr{X}$,

$$x^*_{\alpha y^*}(x) = \alpha y^*[T(x)] = \alpha x^*_{y^*}(x)$$

and

$$x^*_{y^*_1+y^*_2}(x) = (y^*_1 + y^*_2)[T(x)] = y^*_1[T(x)] + y^*_2[T(x)]$$
$$= x^*_{y^*_1}(x) + x^*_{y^*_2}(x).$$

I.e.,

$$T^*(\alpha y^*) = \alpha T^*(y^*)$$

and

$$T^*(y^*_1 + y^*_2) = T^*(y^*_1) + T^*(y^*_2),$$

so that T^* is a linear transformation.

When $\mathscr{X}$ and $\mathscr{Y}$ are Banach spaces, T^* is also continuous. (Hille [20], Def. 2.13.1 and Theorem 2.13.3, p. 27. Note that here continuity is equivalent to boundedness.)

When $\mathscr{X}$ and $\mathscr{Y}$ are finite-dimensional Euclidean spaces, let A denote the matrix such that

$$T(x) = Ax.$$

Here linear functionals belong to their respective spaces ($\mathscr{X} = \mathscr{X}^*$, $\mathscr{Y} = \mathscr{Y}^*$) and $x^*(x) = x^{*\prime}x$, etc., where the prime denotes transposition. Hence the relation $x^*(x) = y^*[T(x)]$ may be written as $x^{*\prime}x = y^{*\prime}Ax$, i.e., $x^{*\prime} = y^{*\prime}A$, so that $x^* = A'y^*$. I.e., the adjoint transformation T^* corresponds to a premultiplication by the transpose A' of the matrix A representing the given transformation T.

II.2.1. Let A be a set and ρ a transitive binary relation in A. When the relation holds for the ordered pair $a', a'' \in A$, we write $a' \rho a''$. When it does not, we write $a' \bar{\rho} a''$. An element $a_0 \in A_1$, $A_1 \subseteq A$ is said to be *ρ-maximal* in A_1 (or, more briefly, maximal) if, for any $a' \in A$, the relations $a' \in A_1$, $a' \rho a_0$ imply $a_0 \rho a'$.

Let ψ be a real-valued function on A. Then ψ is said to be *isotone* (with respect to ρ) if

$$a' \rho a'' \text{ implies } \psi(a') \geq \psi(a'');$$

[22] Cf. Kuratowski [32], p. 74, (6).

ψ is said to be *strictly isotone* (with respect to ρ) if, in addition,

$$a' \rho a'' \text{ and } a'' \bar{\rho} a' \text{ imply } \psi(a') > \psi(a'') .$$

In what follows we usually deal with transitive reflexive relations denoted by $\geqq$ or similar symbols. (The denial of $\geqq$ is written $\ngeqq$.) We then write $a' \geq a''$ to mean $a' \geqq a''$ and $a'' \ngeqq a'$.

II.2.2. If $\mathscr{W}$ is a linear system and $K \subseteq \mathscr{W}$, K is said to be a *cone*[23] if

$$w \in K, \lambda \geqq 0 \text{ imply } \lambda w \in K.$$

K is said to be a *convex cone* if K is a cone and a convex set.

A set $K \subseteq \mathscr{W}$ is a *convex cone* if and only if it satisfies

$$w \in K, \lambda \geqq 0 \text{ imply } \lambda w \in K ,$$

and

$$w' \in K, w'' \in K \text{ imply } w' + w'' \in K .$$

It may be noted that both the space $\mathscr{W}$ and the one-element set $\{0_w\}$ are convex cones.

II.2.3. Given a convex cone $K \subseteq \mathscr{W}$, a transitive reflexive relation to be denoted by $\geqq$ (or $\geqq_K$ if we wish to be more explicit) may be defined as follows: for any $w', w'' \in \mathscr{W}$, $w' \geqq w''$ if and only if $w' - w'' \in K$. (In particular, $w \geqq 0_w$ if and only if $w \in K$.)

Example. Let $\mathscr{W}$ be the Euclidean two-space of elements $w = (w^{(1)}, w^{(2)})$ where $w^{(1)}, w^{(2)}$ are real numbers. Then the following convex cones are of interest in defining ordering relations:

$$K_1 = \{w : w^{(1)} \geqq 0, w^{(2)} \geqq 0\} ,$$
$$K_2 = \{w : w^{(1)} \geqq 0, w^{(2)} = 0\} ,$$
$$K_3 = \{0_w\} ,$$
$$K_4 = \mathscr{W} .$$

We see that

$$w \geqq_{K_1} 0_w \text{ means } w^{(1)} \geqq 0, w^{(2)} \geqq 0 ,$$
$$w \geqq_{K_2} 0_w \text{ means } w^{(1)} \geqq 0, w^{(2)} = 0 ,$$
$$w \geqq_{K_3} 0_w \text{ means } w^{(1)} = 0, w^{(2)} = 0 ,$$

and

$w \geqq_{K_4} 0_w$ holds for all $w \in \mathscr{W}$, i.e., it is a vacuous constraint.[24] Other relations could be obtained by replacing $\geqq$ by $>$ in the definitions of K_1 and K_2. Thus we have a great range of possibilities

[23] It would be more precise to speak of a cone with the vertex at origin, but we omit the qualifying phrase since no other cones will be considered. (Our use of the term "cone" may seem unnatural, but it permits us to define the "convex cone" as a cone which is convex.)

[24] This makes it possible to cover simultaneously the cases of unconstrained and (non-vacuously) constrained maximization by orderings based on convex cones.

covering equalities, inequalities ($\geqq$ or $>$), and their various combinations. This makes it possible to obtain results which can be specialized in a variety of ways.

Let $\mathscr{W}$ be a linear topological space and K a convex cone in $\mathscr{W}$. In the applications, we are interested in the topological, as well as the algebraic properties of the cone K. In some theorems, we assume that the cone K is *closed.* This is obviously true of the cones K_1, K_2, K_3, K_4 in the natural (Euclidean) topology of the plane. On the other hand, the cone

$$K_5 = \{(w^{(1)}, w^{(2)}) : w^{(1)} > 0, w^{(2)} > 0\}$$

is not closed in the natural topology of the plane. We see that lack of closedness may result from using cones corresponding to strict, rather than weak, inequalities. In the economic applications the inequalities are usually of the weak type, hence the closedness of the corresponding cones is not a serious restriction.

Another topological property assumed for certain convex cones is that they have *non-empty interiors.* Of the preceding examples, using again the Euclidean topology of the plane, K_1, K_4, and K_5 have interior points, while K_2 and K_3 do not. The requirement of a non-empty interior can be troublesome in infinite-dimensional spaces. Thus consider a space $l_p (p \geqq 1)$ whose elements are infinite sequences $x = (x_1, x_2, \cdots)$ such that

$$\sum_{i=1}^{\infty} |x_i|^p < \infty \ .$$

This space is normable, the norm being of x defined as

$$\left(\sum_{i=1}^{\infty} |x_i|^p\right)^{1/p} .$$

Now consider the convex cone K consisting of all the elements x of l_p whose every coordinate is non-negative, this being the natural counterpart of the non-negative orthant in a finite-dimensional space. It may be seen that K has no interior, i.e., every element of K is a boundary point. To see this, take an arbitrary element x' of K. Given a positive number ε, however small, one can find an element x'' of l_p whose distance from the element x' is less than ε and such that x'' has at least one negative coordinate ; this can be accomplished by taking x'' such that all but one of the coordinates of x'' are the same as the corresponding coordinates of x', while one coordinate of x'' (with a sufficiently high subscript) is the negative of the corresponding coordinate of x'.

On the other hand, let (m) denote the space of infinite bounded sequences $x = (x_1, x_2, \cdots)$, normed by

$$||x|| = \sup_{1 \leq i} (|x_i|) \ ,$$

and define K, as in the preceding example, as the set of all x with non-negative coordinates. Here any point x whose coordinates are all positive is an interior point of the cone.

II.2.4. Let $K \subseteq \mathscr{W}$ be a convex cone. Then the conjugate $K^{\oplus}$ of K is defined by

$$K^{\oplus} = \{w^* \in \mathscr{W}^* : w^*(w) \geqq 0 \text{ for all } w \in K\} .$$

Since $K^{\oplus}$ is a cone, it is called the *conjugate cone* of K. Clearly, $K^{\oplus}$ is the set of linear continuous functionals isotone with respect to $\geqq_K$. We note that $K^{\oplus}$ is never empty, since $0^* \in K^{\oplus}$.

In accord with the notational principles of II.2.3, we write $w^* \geqq_{K^{\oplus}} 0^*_w$ (or, more simply, $w^* \geqq 0$), and call w^* *non-negative on* K if $w^* \in K^{\oplus}$. Furthermore, we write $w^* >_{K^{\oplus}} 0$ (or : $w^* > 0^*_w$) and call w^* *strictly positive on*[25] K if $w^* \geqq_{K^{\oplus}} 0$, and $w \geq_K 0$ implies $w^*(w) > 0$. It is seen that $w^* > 0$ if and only if w^* is a linear continuous functional strictly isotone with respect to $\geqq_K$.

II.2.5. LEMMA II.2.[26] *Let K be a closed convex cone in a locally convex linear space $\mathscr{W}$ and let $w_0 \in \mathscr{W}$ be such that*

$$w^*(w_0) \geqq 0 \textit{ for all } w^* \geqq 0 .$$

Then $w_0 \in K$.

PROOF. Suppose $w_0 \notin K$. By virtue of the Mazur-Bourgin Theorem[27] K is regularly° convex, since it is closed and convex and $\mathscr{W}$ is locally convex, so that there exists a $w_0^* \in \mathscr{W}^*$ such that

$$\sup_{w \in K} w_0^*(w) < w_0^*(w_0) .$$

Now, since $w_0^*(w_0)$ is a fixed number and K a cone, we must have

$$\sup_{w \in K} w_0^*(w) = 0 .$$

Let $w_1^* = - w_0^*$. Then

$$w_1^*(w) \geqq 0 \text{ for all } w \in K; \text{ i.e., } w_1^* \in K^{\oplus} ,$$

and

$$w_1^*(w_0) < 0 ,$$

which contradict the hypothesis of the Lemma, hence the proof is completed.

II.2.6.1. LEMMA II.3. *If $K \subseteq \mathscr{W}$ is a convex cone, then the conjugate cone $K^{\oplus}$ is regularly convex.*

PROOF. If $K^{\oplus} = \mathscr{W}^*$, no $w_0^* \notin K^{\oplus}$ exists and the condition of regularity is (vacuously) satisfied. Now let $K^{\oplus} \neq \mathscr{W}^*$ and take $w_0^* \notin K^{\oplus}$. Then there exists a $w_1 \in K$ such that $w_0^*(w_1) < 0$. Write $w_0 = - w_1$. Since

---

[25] The reader should be warned that this term has a somewhat unusual meaning. In particular, if K is the origin, *every* linear functional is strictly positive on K.

[26] For the case of linear normed spaces, cf. Krein and Rutman [29], p. 16.

[27] Cf. Theorem II.2. in II.1.4.1.

$w^*(w_1) \geqq 0$ for all $w^* \in K^\oplus$ (because $w_1 \in K$ and by definition of $K^\oplus$), we have

$$w^*(w_0) = - w^*(w_1) \leqq 0 \qquad \text{for all } w^* \in K^\oplus ,$$

while

$$w_0^*(w_0) = - w_0^*(w_1) > 0 ,$$

which shows that $K^\oplus$ is regularly convex. This completes the proof.

II.2.6.2. For the special case of $\mathscr{W}$ linear normed, the preceding result follows from Krein and Rutman [29], p. 38, where it is proved that $K^\oplus$ is weak* closed; since the convexity of $K^\oplus$ is evident, this implies the regular convexity of $K^\oplus$; cf. II.1.4.1.

II.3.1. A very abstract version of a (partial ordering) maximization problem of the type considered in the present chapter in connection with the Lagrangian saddle-points can be formulated as follows.

Let $\mathscr{X}$ be an arbitrary space, X a subset of $\mathscr{X}$, $\mathscr{Y}$ an arbitrary space with the transitive relation ρ, and $\mathscr{Z}$ a linear system with $z' \geqq z''$ defined to mean $z' - z'' \in P_z$ where P_z is a convex cone.

Furthermore let f be a function on $\mathscr{X}$ into $\mathscr{Y}$ and g a function on $\mathscr{X}$ into $\mathscr{Z}$.

Let the *constraints* be $x \in X$ and $g(x) \geqq 0_z$.

Let X_0 denote the *permissible x-set*, i.e.,

$$X_0 = X \cap g^{-1}(P_z) = \{x \in \mathscr{X} : x \in X, g(x) \geqq 0_z\} ,$$

while

$$Y_0 = f(X_0) = \{y : y = f(x), x \in X, g(x) \geqq 0_z\}$$

is the *permissible y-set*.

Denote by $\hat{Y}_0$ the ρ-maximal subset of Y_0; i.e.,

$$\hat{Y}_0 = \{y_0 \in y_0 : y' \in Y_0, y' \rho y_0 \text{ imply } y_0 \rho y'\}$$

and call $\hat{Y}_0$ the *maximal y-set*, while $\hat{X}_0 = f^{-1}(\hat{Y}_0)$ is called the *maximal x-set*. An element of a (y- or x-) maximal set is called *maximal*.

The objective is typically to characterize $\hat{X}_0$. Hence a maximization problem is uniquely determined by the selection of

$$\pi \equiv (\mathscr{X}, X ; \mathscr{Y}, \rho ; \mathscr{Z}, P_z ; f, g)$$

and we may refer to π as the (*partial ordering*) *maximization problem*.

In some contexts we only need $\mathscr{Y}$, ρ, and Y_0, without reference to how Y_0 is defined. In others specializing assumptions are made with regard to the entities defining π.

II.3.2. For a given maximization problem π, as defined in the preceding section, we define a *generalized Lagrangian expression* Φ_π or (where safe) Φ by

$$\Phi = \Phi_\pi = \Phi_\pi(x, \zeta^* ; \eta^*) = \eta^*[f(x)] + \zeta^*[g(x)] , \quad x \in \mathscr{X} ,$$

where ζ^* and η^* are real-valued functions on $\mathscr{Z}$ and $\mathscr{Y}$ respectively.

That is, Φ_π is a real-valued function in the Cartesian product space $\mathscr{X} \times [\zeta^*] \times [\eta^*]$ where $[\zeta^*]$ and $[\eta^*]$ are the spaces of real-valued functions on $\mathfrak{Z}$ and $\mathscr{Y}$ respectively.

II.3.3. Let $\mathfrak{Z}$ be a topological linear space, $\mathfrak{Z}^*$ its conjugate space. Symbols such as z^*, z_0^* denote elements of $\mathfrak{Z}^*$. We say that Φ_{π_1} (where π_1 differs from π in that π_1 requires $\mathfrak{Z}$ to be a linear topological space) has an *isotone saddle-point* at $(x_0, z_0^* ; \eta_0^*)$ if

(1) $x_0 \in X, z_0^* \geqq 0, \eta_0^* \in [\eta_0^*]$ and η_0^* is strictly isotone with respect to ρ,

(2) $\Phi_{\pi_1}(x, z_0^* ; \eta_0^*) \leqq \Phi_{\pi_1}(x_0, z_0^* ; \eta_0^*) \leqq \Phi_{\pi_1}(x_0, z^* ; \eta_0^*)$
for all $x \in X$ and all $z^* \geqq 0$.

Now specialize the partial ordering maximization problem π_1 to the *vectorial* (ordering) *maximization problem* π_2 as follows. Let $\mathscr{X}$ be a linear system, P_x a convex cone in $\mathscr{X}$, $x \geqq 0$ be defined as $x \in P_x$, and $X = P_x$. Furthermore, let $\mathscr{Y}$ be a linear topological space, P_y be a convex cone in $\mathscr{Y}$, and let ρ be $\geqq_{P_y}$. (Hence $y_0^* \in \mathscr{Y}^*$, and $y_0^* > 0$ means y_0^* is strictly positive on P_y.)

We then say that the Lagrangian expression Φ_{π_2} has a *non-negative saddle-point* at $(x_0, z_0^* ; y_0^*)$ if

(1′) $$x_0 \geqq 0 , \qquad z_0^* \geqq 0 , \qquad y_0^* > 0$$

and

(2′) $$\Phi_{\pi_2}(x, z_0^* ; y_0^*) \leqq \Phi_{\pi_2}(x_0, z_0^* ; y_0^*) \leqq \Phi_{\pi_2}(x_0, z^* ; y_0^*)$$
for all $x \geqq 0$ and all $z^* \geqq 0$.

II.4. Let $\mathscr{X}$ and $\mathscr{Y}$ be linear systems and let f be a (single-valued) function with a convex domain $\mathscr{D} \subseteq \mathscr{X}$ and range $\mathscr{R} \subseteq \mathscr{Y}$. Then the *function* f is said to be *concave* if, given any $x', x'' \in \mathscr{D}$ and any real number $0 < \theta < 1$, we have

$$(1 - \theta)f(x') + \theta f(x'') \leqq f[(1 - \theta)x' + \theta x''] ,$$

where $y' \geqq y''$ means $y' - y'' \in K$ for a given convex cone K in $\mathscr{Y}$.

II.5.1. Let $\mathscr{W}$ be a Banach space and h a (single-valued) function whose domain is a set A of reals and the range a subset of $\mathscr{W}$, i.e.,

$$w = h(\alpha) , \qquad \alpha \in A, w \in \mathscr{W} .$$

Following Graves[28] we define the *first derivative* $h'(\alpha_0) = \left.\frac{d}{d\alpha}h(\alpha)\right|_{\alpha=\alpha_0}$ of h with regard to α at α_0 as the element of $\mathscr{W}$ such that

$$\lim_{\alpha \to \alpha_0} \left\| \frac{h(\alpha) - h(\alpha_0)}{\alpha - \alpha_0} - h'(\alpha_0) \right\| = 0 .$$

[28] Reference [17], p. 164.

Similarly,

$$\frac{d^2}{d\alpha^2}h(\alpha)\Big|_{\alpha=\alpha_0} = \frac{d}{d\alpha}\frac{d}{d\alpha}h(\alpha)\Big|_{\alpha=\alpha_0}, \text{ etc.}$$

Now let $\mathscr{X}$ and $\mathscr{Y}$ be Banach spaces and f a function on $\mathscr{X}$ into $\mathscr{Y}$. Then $f(x_0 + \alpha x')$, α real, $x_0, x' \in \mathscr{X}$, may be regarded, for fixed x_0 and x', as a function of the real variable α with values in $\mathscr{Y}$. We define the *first, second, etc., variation of f at x_0 with increment x'* by

$$\delta f(x_0\,;\,x') = \frac{d}{d\alpha}f(x_0 + \alpha x')\Big|_{\alpha=0},$$

$$\delta^2 f(x_0\,;\,x') = \frac{d^2}{d\alpha^2}f(x_0 + \alpha x')\Big|_{\alpha=0}, \text{ etc.}$$ [29]

When the domain of f is open and $\delta f(x_0\,;\,x')$ exists and is continuous in x', $\delta f(x_0\,;\,x')$ is called the *Fréchet differential* of f at x_0 with increment x'. It has been shown[30] that the Fréchet differential $\delta f(x_0\,;\,x')$ so defined is linear (i.e., homogeneous and additive) as well as continuous in x'; also that

$$\lim_{||x'||\to 0} \frac{1}{||x'||}\left\| f(x_0 + x') - f(x_0) - \delta f(x_0\,;\,x')\right\| = 0$$

for all x in the domain of f.

II.5.2. The "function of a function rule" is valid for Fréchet differentials[31] and may be stated as follows.

Let $\mathscr{Y}, \mathscr{X}, \mathscr{Z}$ be Banach spaces; f a function on $\mathscr{X}$ into $\mathscr{Y}$, g on $\mathscr{Z}$ into $\mathscr{X}$.

$$y = f(x)\,, \qquad y_0 = f(x_0)\,,$$
$$x = g(z)\,, \qquad x_0 = g(z_0)\,,$$

and assume that f and g possess Fréchet differentials at x_0 and z_0 respectively. Write

$$f(g(z)) = h(z)$$

so that h is a function in $\mathscr{Z}$ into $\mathscr{Y}$. Then, for $\zeta \in \mathscr{Z}$,

$$\delta h(z_0\,;\,\zeta) = \delta f(x_0\,;\,\delta g(z_0\,;\,\zeta))\,.$$

The reader is referred to Fréchet [11], [13], Hildebrandt and Graves

---

29 An equivalent definition of $\delta f(x_0\,;\,x')$ is

$$\delta f(x_0\,;\,x') = \lim_{\alpha\to 0}\frac{f(x_0+\alpha x') - f(x_0)}{\alpha}$$

where, for a function $w=h(\alpha)$ of real variable α with values in $\mathscr{W}$, we write

$$\lim_{\alpha\to\alpha_0} h(\alpha) = w_0,\ w_0 \in \mathscr{W} \quad \text{if and only if} \lim_{\alpha\to\alpha_0} ||\,h(\alpha)-w_0\,|| = 0\,.$$

Cf. Hildebrandt and Graves [19], p. 136, and Hille [20], pp. 71–72.

30 Hille [20], p. 73 and p. 72, Def. 4.3.4.

31 Cf. Hildebrandt and Graves [19], pp. 141–44; Graves [17], p. 649.

[19], Graves [16], [17], and Hille [20] for an account of the properties of Fréchet differentials.[32]

II.6.1. Let $\mathscr{X}$ and $\mathscr{Y}$ be two Banach spaces. Consider the linear system whose elements are the ordered pairs $(x, y), x \in \mathscr{X}, y \in \mathscr{Y}$, with addition and scalar multiplication defined by

$$(1') \qquad \begin{cases}(x', y') + (x'', y'') = (x' + x'', y' + y'') \\ \alpha(x, y) = (\alpha x, \alpha y), \ \alpha \text{ real.}\end{cases}$$

Then the linear system of the ordered pairs (x, y) will become a Banach space if it is normed in such a way that[33]

$$(1'') \qquad \lim_{n\to\infty} x_n = x_0 \text{ and } \lim_{n\to\infty} y_n = y_0 \quad \text{if and only if } \lim_{n\to\infty} \left\|(x_n, y_n) - (x_0, y_0)\right\| = 0 \,.$$

Such a Banach space of the ordered pairs (x, y) is denoted by $\mathscr{X} \times \mathscr{Y}$ and is called the *(Banach) product* of $\mathscr{X}$ and $\mathscr{Y}$. Writing, for $A \subseteq \mathscr{X}$, $B \subseteq \mathscr{Y}$, $A \times B = \{(x, y) : x \in A, y \in B\}$ we have[34] $A \times B$ closed if and only if both A and B are closed.

More generally, let $\mathscr{X}$ and $\mathscr{Y}$ be linear topological spaces and consider the linear system of the ordered pairs (x, y) with the operations defined by (1′) above.

Then the space of pairs (x, y), again to be denoted by $\mathscr{X} \times \mathscr{Y}$ (and called *linear topological product*), may be topologized by choosing as a base[35] the sets

$$\{(x', y') : x' \in U_x, y' \in \mathscr{Y}\} \,,$$
$$\{(x'', y'') : x'' \in \mathscr{Y}, y'' \in U_y\} \,,$$
$$\{(x''', y''') : x''' \in U_x, y''' \in U_y\} \,,$$

where U_x is any open set in $\mathscr{X}$, U_y any open set in $\mathscr{Y}$. It may be noted for later reference that[36] with this topology $A \times B$ is closed if A and B both are.

It is known [37] that if $\mathscr{X}$ and $\mathscr{Y}$ are linear topological spaces, then so is $\mathscr{X} \times \mathscr{Y}$; if $\mathscr{X}$ and $\mathscr{Y}$ are locally convex, then so is $\mathscr{X} \times \mathscr{Y}$.

II.6.2. Let[38] $\mathscr{X} = \mathscr{X}' \times \mathscr{X}''$ be the (Banach) product of the two Banach spaces $\mathscr{X}', \mathscr{X}''$. The symbols x' and ξ' denote elements of $\mathscr{X}'$, x'' and ξ'' those of $\mathscr{X}''$, x and ξ those of $\mathscr{X}$. If f is a function on

[32] See also V.3.3.8 for a discussion of differentials in a class of spaces wider than Banach.

[33] Banach [3], pp. 181–82, especially eq. (33), where examples of norms satisfying (1) are given. Cf. also Hyers [22], pp. 3, 5, and Tychonoff [43], p. 772.

[34] Cf. Kuratowski [32], 24.II.1, p. 219.

[35] Cf. Lefschetz [34], p. 6 (6.1); p. 10, Section 12.

[36] Lefschetz [34], p. 11 (12.6).

[37] Tychonoff [43], p. 772; Bourgin [9], p. 639; Hyers [22], pp. 3, 5. In these sources it is shown how a linear topological product of an arbitrary family of spaces is formed.

[38] We confine ourselves to the product of two spaces. The treatment of $\mathscr{X}^{(1)} \times \mathscr{X}^{(2)} \times \cdots \times \mathscr{X}^{(n)}$ is quite analogous.

$\mathscr{X}$ into the Banach space $\mathscr{Y}$, $\delta f(x ; \xi)$ will denote the Fréchet differential of f at x with increment ξ.

Then the *partial Fréchet differential* of f *with respect to* x' *at* x_0 *with increment* ξ' is written as $\delta_{x'} f(x_0 ; \xi')$ and is defined by

(1) $$\delta_{x'} f(x_0 ; \xi') = \delta f(x_0 ; (\xi', 0_{x''})) , \qquad x_0 \in \mathscr{X}, \qquad x', \xi' \in \mathscr{X}' .$$

We have[39] the additivity law

(2) $$\delta f(x_0 ; (\xi', \xi'')) = \delta_{x'} f(x_0 ; \xi') + \delta_{x''} f(x_0 ; \xi'') .$$

II.6.3. We shall now state the "function of a function" rule for the case of a function of several variables.

Let

$$\mathscr{X} = \mathscr{X}^{(1)} \times \mathscr{X}^{(2)} \times \cdots \times \mathscr{X}^{(n)}$$
$$\mathscr{Z} = \mathscr{Z}^{(1)} \times \mathscr{Z}^{(2)} \times \cdots \times \mathscr{Z}^{(m)}$$

where all spaces are Banach and so are the products. Also, let f be a function in $\mathscr{X}$ into the Banach space $\mathscr{Y}$, $g^{(i)}$ in $\mathscr{Z}$ into $\mathscr{X}^{(i)}$.

$$y = f(x) , \qquad y_0 = f(x_0) ,$$
$$x^{(i)} = g^{(i)}(z) , \qquad x_0^{(i)} = g^{(i)}(z_0) \qquad (i = 1, 2, \cdots, n)$$

and assume that f and each of the $g^{(i)}$ possess Fréchet differentials at x_0 and z_0 respectively. Write

$$h(z) = f((g^{(1)}(z), g^{(2)}(z), \cdots, g^{(n)}(z)))$$

so that h is a function in $\mathscr{Z}$ into $\mathscr{Y}$. Then, for $\zeta \in \mathscr{Z}$,

$$\delta h(z_0 ; \zeta) = \sum_{i=1}^{n} \delta f x^{(i)}(x_0 ; \sum_{j=1}^{m} \delta z^{(j)} g^{(i)}(z_0 ; \zeta^{(i)})$$

where $\zeta^{(i)} \in \mathscr{Z}^{(i)}$.[40]

II.6.4. We shall find it convenient to define a "quasi-saddle-point" for Lagrangian expressions. We say that

$$\Phi(x, z^* ; y_0^*) \equiv y_0^*[f(x)] + z^*[g(x)]$$

has a non-negative *quasi-saddle-point* at $(x_0, z_0^* ; y_0^*)$ if and only if $y_0^* > 0$, $x_0 \geqq 0$, $z_0^* \geqq 0$, and the following relations hold:

$$\delta_x \Phi((x_0, z_0^*) ; \xi) \leqq 0 \qquad \text{for all } x \geqq 0, x = x_0 + \xi ,$$
$$\delta_x \Phi((x_0, z_0^*) ; x_0) = 0 ,$$
$$\delta_{z^*} \Phi((x_0, z_0^*) ; \zeta^*) = \zeta^*[g(x_0)] \geqq 0 \qquad \text{for all } z^* \geqq 0, \zeta^* = z^* - z_0^* ,$$
$$\delta_{z^*} \Phi((x_0, z_0^*) ; z_0^*) = z_0^*[g(x_0)] = 0 .$$

It is seen that if Φ has a non-negative saddle-point at $(x_0, z_0^* ; y_0^*)$, then it necessarily has a non-negative quasi-saddle-point there, but the converse is not true.

---

39 Cf. Hildebrandt and Graves [19], p. 138.

40 Cf. Fréchet [11], pp. 318–21. (The reprinted version in [13] is free of the misprints in [11].)

III. The "Minkowski-Farkas Lemma"

III.1. Throughout III, $\mathscr{X}$ is a linear topological space, $\mathscr{Y}$ a locally convex[41] linear space, $y' \geqq y''$ means $y' - y'' \in P_y$, where P_y is a closed convex cone. T is a linear continuous transformation on $\mathscr{X}$ into $\mathscr{Y}$. $\mathscr{X}, \mathscr{Y}, P_y, T$ are fixed throughout. $\mathscr{X}^*, \mathscr{Y}^*$ are the conjugate spaces of $\mathscr{X}, \mathscr{Y}$, and T^* is the adjoint of T.

III.2. If $x^* \in \mathscr{X}^*$ is such that

$$(1) \qquad x^* = T^*(y^*) \text{ for some } y^* \in \mathscr{Y}^* ,$$

we say that eq. (1) is *solvable.* If x^* is such that eq. (1) holds for some $y^* \geqq 0$, we say that eq. (1) is *positively*[42] *solvable.* We then also say that x^* *makes eq.* (1) *positively solvable.*

The set of all $x^* \in \mathscr{X}^*$ which make eq. (1) positively solvable will be denoted by Z_T, i.e.,

$$(2) \qquad Z_T = \{x^* \in \mathscr{X}^* : x^* = T^*(y^*), y^* \geqq 0\} .$$

Note that, since $\{y^* : y^* \geqq 0\} = P_y^{\oplus}$, we have

$$(3) \qquad Z_T = T^*(P_y^{\oplus}) .$$

The point x^* is said to be *positively normal with regard to* T if

$$(4) \qquad \text{for all } x \in \mathscr{X}, T(x) \geqq 0 \text{ implies } x^*(x) \geqq 0 .$$

We shall denote by V_T the set of all x^* positively normal with regard to a given T, i.e.,

$$(5) \qquad V_T = \{x^* \in \mathscr{X}^* : \text{for all } x \in \mathscr{X}, T(x) \geqq 0 \text{ implies } x^*(x) \geqq 0\} .$$

III.3. THEOREM III.1. *If x^* makes eq.* (1) *positively solvable, then x^* is positively normal with regard to T. In set language,*

$$(6) \qquad Z_T \subseteqq V_T .$$

PROOF. Let $x^* = T^*(y^*)$ for some $y^* \geqq 0$. Then we have, by the definition of T^*,

$$(7) \qquad x^*(x) = (T^*y^*)(x) = y^*(Tx), \qquad \text{for all } x \in \mathscr{X} .$$

Therefore, since $y^* \geqq 0$, $T(x) \geqq 0$ implies $x^*(x) \geqq 0$.

III.4. THEOREM III.2. *If every x^* positively normal with regard to T makes eq.* (1) *positively solvable, then the set of all x^* which make eq.* (1) *positively solvable is regularly convex. In set language, if $V_T \subseteqq Z_T$, then Z_T is regularly convex.* (*We may note that, in view of* (6), *Theorem* III.2 *may be equivalently restated as follows: if $V_T = Z_T$, then Z_T is regularly convex.*)

PROOF. We note that the set

$$(8) \qquad X_T = \{x \in \mathscr{X} : T(x) \geqq 0\}$$

is a convex cone and $V_T = X_T^{\oplus}$, so that, by Lemma II.3 in II.2.6.1, V_T is regularly convex and hence so is $Z_T = V_T$.

[41] Cf. II.1.1.3.

[42] "Non-negatively" would be more accurate but awkward.

III.5. THEOREM III.3. *V_T coincides with the regular convex envelope of Z_T*:

(9) $$V_T = \tilde{Z}_T .$$

PROOF. In view of Theorem III.1, it will suffice to establish

(10) $$V_T \subseteq \tilde{Z}_T ,$$

i.e., that $x^* \notin \tilde{Z}_T$ implies $x^* \notin V_T$.

Consider some $x_0^* \notin \tilde{Z}_T$. We shall find x_1 such that $T(x_1) \geqq 0$ while $x_0^*(x_1) < 0$.

Since $\tilde{Z}_T$ is regular convex (cf. II.1.4.2), there must exist $x_0 \in \mathscr{X}$ such that

(11) $$\sup_{x^* \in \tilde{Z}_T} x^*(x_0) < x_0^*(x_0) .$$

Since a cone Z_T is contained in $\tilde{Z}_T$ and

$$\sup_{x^* \in \tilde{Z}_T} x^*(x_0)$$

is finite, (11) implies

(12) $$x^*(x_0) \leqq 0 < x_0^*(x_0) \text{ for any } x^* \in Z_T .$$

Now write

(13) $$x_1 = - x_0 .$$

Then (12) may be written

(14) $$x_0^*(x_1) = - x_0^*(x_0) < 0 ,$$

and

(15) $x^*(x_1) = y^*(Tx_1) = (T^*y^*)(x_1) = - x^*(x_0) \geqq 0$, for any $y^* \geqq 0$,

since $x^* = T^*(y^*) \in Z_T$.

Now since $P_y = \{y^* : y^* \geqq 0\}$ is assumed closed and $\mathscr{Y}$ locally convex, the Lemma II.2 in II.2.5 applies. It follows that

(16) $$T(x_1) \geqq 0 .$$

But (14) and (16) together imply $x_0^* \notin V_T$.

III.6. THEOREM III.4. *The positive normality of x^* with regard to T is equivalent to x^* making eq.* (1) *positively solvable if and only if the set of all x^* making eq.* (1) *solvable is regularly convex. In set language,*

(17) $Z_T = V_T$ *if and only if* Z_T *is regularly convex.*

PROOF. If $Z_T = V_T$, the regular convexity of Z_T follows from Theorem III.2. On the other hand, if Z_T is regularly convex, we have $Z_T = \tilde{Z}_T$ (cf. Lemma II.1 in II.1.4.2). The equality $Z_T = V_T$ then follows from Theorem III.3.

III.7. *The finite-dimensional Euclidean case.* In a reflexive Banach space, a set is regularly convex if and only if it is convex and (strongly) closed (cf. II.1.4). Since Z_T is always convex, for reflexive Banach spaces one may substitute "(strongly) closed" for "regularly convex" in Theorem III. 1, 2, 3, and 4.

In particular, if $\mathscr{X}$ and $\mathscr{Y}$ are finite-dimensional Euclidean spaces (hence Banach and reflexive in the Euclidean distance topology) and T is represented by a matrix, Z_T is a polyhedral convex cone (cf. Gale in [14], p. 290, Def. 1′) which is closed in the Euclidean distance topology. Hence for this case Z_T is necessarily regularly convex and $Z_T = V_T$ for all T. The Minkowski-Farkas Lemma as usually stated asserts that $V_T \subseteq Z_T$ in the finite-dimensional Euclidean case. This follows from Theorem III.3, since Z_T is known to be regularly convex.

[Let $y = (y_1, y_2, \cdots, y_n)$ and write $I = \{1, 2, \cdots, n\}$.

Partition I into I' and I'' where $I' \cup I'' = I, I' \cap I'' = \phi$, and either I' or I'' may be empty. The relation $y \geqq 0$ is interpreted as meaning

$$y_i \geqq 0 \qquad \text{if } i \in I' ,$$
$$y_i = 0 \qquad \text{if } i \in I'' .$$

The Minkowski-Farkas Lemma is usually stated for $I = I'$, but it is clear that $P_y = \{y \in \mathscr{Y} : y \geqq 0\}$, where the meaning of $y \geqq 0$ is that just stated, is necessarily closed.]

III. *Appendix*: *Relationship with Hausdorff's results.* [NOTE: This appendix is incorrect in its present form and should be ignored. For technical reasons, however, it was impossible to eliminate it from the present printing.]

IIIa.1. Suppose that x^* is positively normal with regard to T (cf. III.2) and let, for some $x' \in \mathscr{X}$,

(1) $$T(x') = 0 .$$

Then

(2′) $$T(x') \geqq 0$$

and

(2″) $$T(- x') \geqq 0 .$$

Since x^* is positively normal, the preceding inequalities yield, respectively,

(3′) $$x^*(x') \geqq 0$$

and

(3″) $$x^*(- x') \geqq 0 ,$$

i.e.,

(4) $$x^*(x') = 0 .$$

Hence, if x^* is positively normal with regard to T, we have

(5) for all $x \in \mathscr{X}$, $T(x) = 0$ implies $x^*(x) = 0$.

Call x^* satisfying (5) *normal with regard to* T. I.e., we have shown that *if* x^* *is positively normal with regard to* T, *then it is also normal with regard to* T.

IIIa.2. We shall now show that

(6) *if* x^* *is normal with regard to* T, *then either* x^* *or* $-x^*$ *is positively normal with regard to* T.

For suppose it could happen that (5) holds and neither x^* nor $-x^*$ is positively normal with regard to T. Then there must exist $x_1, x_2 \in \mathscr{X}$ such that

(7.1) $$T(x_1) > 0 \ ,$$

(7.2) $$x^*(x_1) > 0 \ ,$$

(8.1) $$T(x_2) > 0 \ ,$$

(8.2) $$x^*(x_2) < 0 \ .$$

[Suppose no such pair x_1, x_2 exists. Then it must be that either $T(x) > 0$ implies $x^*(x) \geqq 0$ or $T(x) > 0$ implies $x^*(x) \leqq 0$. This, in conjunction with (5), would then yield (6).]

Let

(9) $$\lambda = \frac{T(x_1)}{T(x_2)} \ .$$

Then

(10) $$T(x_1 - \lambda x_2) = T(x_1) - \lambda T(x_2) = 0 \ .$$

On the other hand, by (7), (8), and (9) (which imply $\lambda > 0$),

(11) $$x^*(x_1 - \lambda x_2) = x^*(x_1) - \lambda x^*(x_2) > 0 \ .$$

Hence (5) fails to hold for $x_1 - \lambda x_2 \in \mathscr{X}$ which establishes the validity of (6).

IIIa.3. Write

(12) $$F_T = \{x^* : \text{for all } x \in \mathscr{X}, T(x) = 0 \text{ implies } x^*(x) = 0\}$$

(the set of x^* normal with regard to T) and recall that the set of all x^* positively normal with regard to T is denoted by V_T. Hence the results in IIIa.1 and IIIa.2 may be written as

(13) $$F_T = V_T \cup (-V_T) \ .$$

IIIa.4. We shall now show that[43]

(14) $$F_T = V_T - V_T \ .$$

First,

(15) $$F_T = V_T \cup (-V_T) \subseteq V_T - V_T \ ,$$

for any element in $V_T \cup (-V_T)$ is either of the form $x_1^* - 0_x^*$ where $x_1^* \in V_T$ or of the form $0_x^* + x_2^*$ where $x_2^* \in (-V_T)$. Note that $0_x^* \in V_T \cap (-V_T)$.

On the other hand, let x^* be an element of $V_T - V_T$, i.e.,

(16) $$x^* = x_1^* + x_2^* \ , \qquad x_1^* \in V_T \ , \qquad x_2^* \in (-V_T) \ .$$

By (13), $x_i^* \in F_T$ $(i = 1, 2)$. But then $x^* \in F_T$, since F_T is a linear set.

[43] $A - B$ is the set of all elements of the form $a - b$, $a \in A$, $b \in B$. $A - A$ is neither empty nor the null element!

For let $x_i^* \in F_T$ $(i = 1, 2)$. Then $T(x) = 0$ implies $x_i^*(x) = 0$. Consider $x^* = \alpha_1 x_1^* + \alpha_2 x_2^*$ and suppose $T(x) = 0$; then $x^*(x) = \alpha_1 x_1^*(x) + \alpha_2 x_2^*(x) = 0$, hence $x^* \in F_T$.

IIIa.5.0. From now on we shall assume that

(17′) all the spaces considered are Banach (which implies that both T and T^* are bounded, since T was assumed continuous);

(17″) for every $y^* \in \mathscr{Y}^*$, there exist $y_1^* \in P_y^{\oplus}$, $y_2^* \in P_y^{\oplus}$ such that $y^* = y_1^* - y_2^*$.

(17″) is equivalent to the condition that P_y is a *normal cone*; cf. Krein and Rutman [29], Def. 2.2, p. 22, and p. 24.

IIIa.5.1. *Example.* Let $\mathscr{Y}$ be the space of all continuous real-valued functions $y(t)$ defined on the closed interval $[0, 1]$. (This space is usually denoted by $C[0, 1]$.) Then[44] every bounded linear functional y^* can be defined by

$$y^*(y) = \int_0^1 y(t)dg \qquad (y \in \mathscr{Y}) \tag{a}$$

where g is a function of bounded variation. Now define $y \in P_y$ (i.e., $y \geqq 0_y$) to mean

$$y(t) \geqq 0 \qquad \text{for all } 0 \leqq t \leqq 1\,. \tag{b}$$

Then $y^* \geqq 0_y^*$ (i.e., $y^* \in P^{\oplus}$) means that the function g in (a) is monotone non-decreasing. But it is well known (e.g., Titchmarsh [42], p. 355, Sec. 11.4) that if g is a function of bounded variation, it can be expressed as

$$g = g_1 - g_2 \tag{c}$$

where g_1, g_2 are monotone non-decreasing. I.e., the cone P_y is normal.

IIIa.5.2. The condition (17″) may be written as

$$\mathscr{Y}^* = P_y^{\oplus} - P_y^{\oplus}\,. \tag{18}$$

Now suppose

$$x^* \in T^*(\mathscr{Y}^*)\,, \tag{19}$$

i.e.,

$$x^* = T^*(y^*) \text{ for some } y^* \in \mathscr{Y}^*\,. \tag{20}$$

Then, by using (17″), we have

$$x^* = T^*(y_1^* - y_2^*) \qquad (y_i^* \in P_y^{\oplus},\ i = 1, 2)\,, \tag{21}$$

i.e.,

$$x^* = x_1^* - x_2^* \tag{22}$$

where

$$x_i^* = T^*(y_i^*) \qquad (y_i^* \in P_y^{\oplus},\ i = 1, 2)\,, \tag{23}$$

so that, by definition of Z_T (cf. III.2 (2)),

$$x_i^* \in Z_T \qquad (i = 1, 2)\,, \tag{24}$$

[44] Banach [3], Section 4.1, pp. 59–61.

i.e.,

$$x^* \in Z_T - Z_T \,;$$

hence

(25) $$T^*(\mathscr{Y}^*) \subseteq Z_T - Z_T \,.$$

On the other hand, let

(26) $$x^* \in Z_T - Z_T \,.$$

Then the relations (22), (23) hold for some $y_i^* \in P_y^\oplus$ $(i = 1, 2)$, and hence (20) holds for $y^* = y_1^* - y_2^*$, so that (19) follows and

(27) $$T^*(\mathscr{Y}^*) \supseteq Z_T - Z_T \,.$$

(Note that (27) holds even if P_y is not assumed normal.) Equations (25) and (27) together yield

(28) $$T^*(\mathscr{Y}^*) = Z_T - Z_T \,.$$

IIIa.5.3. Consider now the case when Z_T is regularly convex. We know (Theorem III.4) that in this case

(29) $$Z_T = V_T \,.$$

But then, from (14) and (28) we have

(30) $$T^*(\mathscr{Y}^*) = F_T \,.$$

When (30) holds, Hausdorff ([18], p. 307) says that the equation $x^* = T^*(y^*)$ is *normally solvable*; he calls the equation $y = T(x)$ *normally solvable* if and only if

(31) $$T(\mathscr{X}) = F_{T^*} \,,$$

where

(32) $$F_{T^*} = \{y : \text{for all } y^* \in \mathscr{Y}^*,\, T^*(y^*) = 0 \text{ implies } y^*(y) = 0\} \,.$$

Hausdorff shows (*ibid.*, Theorem *X*, pp. 308, 310) that in Banach spaces the following four properties are equivalent: the normal solvability of $x^* = T^*(y^*)$, the normal solvability of $y = T(x)$, the closedness of $T(\mathscr{X})$, and the closedness of $T^*(\mathscr{Y}^*)$, i.e.,

(33) $$(30) \Leftrightarrow (31) \Leftrightarrow T(\mathscr{X}) \text{ closed} \Leftrightarrow T^*(\mathscr{Y}^*) \text{ closed.}$$

IIIa.5.4. Now, under the assumption that P_y is normal and Z_T regularly convex, we have obtained (30). It follows from (33) that both $T(\mathscr{X})$ and $T^*(\mathscr{Y}^*)$ are closed.

The example below[45] shows that Z_T need not be regularly convex when P_y is normal. This is of importance, since it shows that the assumption of regular convexity in the theorems in IV is not automatically satisfied.

Let $\mathscr{X} = C[0, 1]$ and

$$y = T(x)$$

where

[45] Closely related to one suggested by Professor B. Gelbaum.

$$y(t) = \int_0^t x(t)dt \ .$$

Then y is absolutely continuous, hence continuous, and we may take $\mathscr{Y} = C[0, 1]$ also, As noted earlier, we may define $y \geqq 0_y$ to mean $y(t) \geqq 0, 0 \leq t \leq 1$ in which case P_y is normal. Now take any function $y_0 \in C[0, 1]$ which is not absolutely continuous (e.g., the one given by Titchmarsh [42], Sec. 11. 72, p. 366). Then y_0 is not[46] in the range $T(\mathscr{X})$. But y_0 is a strong (uniform) limit of a sequence of polynomials,[47] hence y_0 is an element of the closure of $T(\mathscr{X})$. Hence $T(\mathscr{X})$ is not closed, hence (by (33)), eq. (30) fails, so that Z_T cannot be regularly convex.

IIIa.5.5. Consider now the special case when

(34) $$P_y = \{0_y\} \ .$$

(P_y is (vacuously) normal, but this fact is of no relevance in what follows.) Then

(35) $$P_y^{\oplus} = \mathscr{Y}^* \ .$$

In this case we have (cf. III, eq. (3))

(36) $$Z_T = T^*(\mathscr{Y}^*) \ .$$

Also, using (13), we get

(37) $$V_T = F_T$$

since

(38) $$V_T = - V_T \ .$$

[Let $x^* \in V_T$. Then $T(x) \geqq 0_y$ implies $x^*(x) \geqq 0$. But, for $P_y = \{0_y\}$, $y \geqq 0_y$ is equivalent to $- y \geqq 0_y$; hence $T(x) \geqq 0_y$ implies $T(- x) \geqq 0_y$ which in turn yields $x^*(- x) \geqq 0$ or $- x^*(x) \geqq 0$. The latter relation means that $- x^* \in V_T$. Hence $V_T \subseteqq (- V_T)$. That $(- V_T) \subseteqq V_T$ is shown in the same fashion.]

Now suppose that

(39) $$Z_T = V_T \ .$$

This is equivalent to

(40) $$F_T = T^*(\mathscr{Y}^*) \ ,$$

i.e., Hausdorff's normal solvability of the equation $x^* = T^*(y^*)$.

By Theorem III.4, (39) implies that $Z_T = F_T = T^*(\mathscr{Y}^*)$ is regularly convex; hence (cf. II.1.4.1) $T^*(\mathscr{Y}^*)$ is closed in the weak* topology, hence it is (strongly) closed. Thus we have obtained Hausdorff's result (part of his Theorem X), viz., that the normal solvability implies the closure of $T^*(\mathscr{Y}^*)$, as a special case of our Theorem III.4. On the other hand, suppose the space $\mathscr{X}$ to be reflexive[48] and let $T^*(\mathscr{Y}^*)$ be

[46] Cf. Titchmarsh [42], Section 11.71, p. 364.

[47] The "Weierstrass Theorem," cf. Rudin [40], Section 7.24, p. 131.

[48] Cf. II. 1.4.

closed. In this case (cf. II.1.4) regular convexity is equivalent to regular° convexity and the latter is always equivalent to closure with convexity. Hence, since $T^*(\mathscr{Y}^*)$ is closed and convex, it is regularly convex and this implies, by Theorem III.4, the equalities (39) and (40).

I.e., we have shown, as a special case of our results in III, when $\mathscr{X}$ is reflexive, the (strong) closure of $T^*(\mathscr{Y}^*)$ is a sufficient condition for the normal solvability of the equation $x^* = T^*(y^*)$ which is also a part of Hausdorff's Theorem X.

IV. Further Theorems on Linear Inequalities

IV.1. In IV all spaces are assumed locally convex linear. Products of topological spaces are understood to be linear topological products, hence the product spaces are also locally convex linear.

IV.2. Let U denote a linear continuous transformation on $\mathscr{X}$ into $\mathscr{Z}$. We introduce the transformation T (which is easily seen to be linear and will also be shown to be continuous) on $\mathscr{X}$ into the product space $\mathscr{Y} = \mathscr{Z} \times \mathscr{X}$ defined by

$$(1') \qquad T(x) = (U(x), x) \qquad \text{for all } x \in \mathscr{X} .$$

In the notation of the type used in matrix calculus we may write

$$(1'') \qquad T = \begin{pmatrix} U \\ \\ I \end{pmatrix},$$

where $I(x) = x$, for all $x \in \mathscr{X}$. (I.e., I is the identity transformation in $\mathscr{X}$.)

If P_x, P_z are convex cones in $\mathscr{X}$ and $\mathscr{Z}$ respectively, and $x' \geqq x''$, $z' \geqq z''$ mean $x' - x'' \in P_x$, $z' - z'' \in P_z$ respectively, then for $y = (z, x)$, we write $y' \geqq y''$ if and only if $y' - y'' \in P_y$, where

$$(2) \qquad P_y = P_z \times P_x = \{(z, x) : z \geqq 0, x \geqq 0\} .$$

It may be noted that if P_z and P_x are closed, then so is P_y (cf. II.6.1).

IV.3. Theorem IV.1.

A. *Let $\mathscr{X}$ be a linear topological space, $\mathscr{Z}$ locally convex, U a linear continuous transformation on $\mathscr{X}$ to $\mathscr{Z}$, P_x and P_z closed convex cones in $\mathscr{X}$ and $\mathscr{Z}$ respectively, and assume that the set*

$$(3) \qquad X_T^* = \{x^* \in \mathscr{X}^* : x^* = T^*(y^*), y^* \geqq 0\}$$

is regularly convex.

B. *It follows that, for any $x^* \in \mathscr{X}^*$, if*

$$(4) \qquad U(x) \geqq 0, x \geqq 0 \qquad \textit{imply } x^*(x) \geqq 0 \qquad \textit{for all } x \in \mathscr{X},$$

then there exists a $z_0^* \geqq 0$ such that

(5) $$z_0^*[U(x)] \leqq x^*(x) \qquad \textit{for } x \geqq 0 ,$$

and

(6) $$x^*(x) = 0,\ U(x) \geqq 0,\ x \geqq 0 \qquad \textit{imply } z_0^*[U(x)] = 0 .$$

PROOF. (5) may be rewritten as

(4′) $$T(x) \geqq 0 \qquad \text{implies } x^*(x) \geqq 0 \qquad \text{for all } x \in \mathscr{X} .$$

Furthermore, T is continuous in x.[49]

Since X_T^* is assumed regularly convex, Theorem III.4 yields a functional $y_0^* \geqq 0$ such that

(7) $$x^*(x) = y_0^*[T(x)] \qquad \text{for all } x \in \mathscr{X} .$$

Now, since

(8) $$y = (z, x) = (z, 0) + (0, x) ,$$

we have

(9) $$y_0^*(y) = y_0^*((z, 0)) + y_0^*((0, x)) .$$

We shall write

(10.1) $$y_0^*((z, 0)) = z_0^*(z) \qquad \text{for all } z \in \mathscr{Z} ,$$

(10.2) $$y_0^*((0, x)) = x_0^*(x) \qquad \text{for all } x \in \mathscr{X} ,$$

where $y_0^*((z, 0))$ is continuous in z and $y_0^*((0, x))$ is continuous in x. Since $z \geqq 0$ implies $(z, 0) \geqq 0$ and $x \geqq 0$ implies $(0, x) \geqq 0$, it follows that, for z_0^*, x_0^* defined by (10), $y_0^* \geqq 0$ yields

(11.1) $$z_0^* \geqq 0 ,$$

(11.2) $$x_0^* \geqq 0 .$$

Thus

(12) $$x^*(x) = y_0^*[T(x)] = y_0^*[(U(x), x)]$$
$$= z_0^*[U(x)] + x_0^*(x) \qquad \text{for all } x \in \mathscr{X} .$$

Since $x_0^* \geqq 0$, (5) follows.

Now let x_1 satisfy the hypotheses of (6), i.e.,

(13.1) $$x^*(x_1) = 0$$

and

(13.2) $$T(x_1) \geqq 0 .$$

Equations (13.1) and (5) yield

(14) $$z_0^*[U(x_1)] \leqq 0 .$$

[49] We have $T(x) = (U(x), I(x))$ where $I(x) = x$ for all $x \in \mathscr{X}$. Then (cf. Lefschetz [34], p. 7 (8.2)), T is continuous if every inverse image of a member of a sub-base in $\mathscr{Y} = \mathscr{Z} \times \mathscr{X}$ is open. Such a sub-base is given (cf. Lefschetz [34], p. 10) by the collection of sets

$$Y' = \{y' = (z_1', x') \colon z' \in N_z,\ x' \in \mathscr{X}\} , \qquad Y'' = \{y'' = (z'', x'') \colon z'' \in \mathscr{Z},\ x'' \in N_x\}$$

where N_z, N_x are open sets in $\mathscr{Z}$ and $\mathscr{X}$ respectively. Now the inverse image $T^{-1}(Y') = \{x \colon T(x) \in Y'\} = \{x \colon U(x) \in N_z,\ I(x) \in \mathscr{X}\} = \{x \colon U(x) \in N_z\} = U^{-1}(N_z)$ which is open since N_z is open and U continuous. Similarly $T^{-1}(Y'') = \{x \colon U(x) \in \mathscr{Z},\ I(x) \in N_x\} = N_x$ which is open.

On the other hand, since $U(x_1) \geqq 0$, and $z_0^* \geqq 0$,

(15) $$z_0^*[U(x_1)] \geqq 0 .$$

Equations (14) and (15) yield the conclusion of (6).

IV.4. Let all hypotheses under A in Theorem IV.1 hold, except that X_T^* is not assumed regularly convex while $\mathscr{X}$ and $\mathscr{Z}$ are taken to be normed spaces. Suppose there exists a $z_0^* \geqq 0$ such that (5) holds. Then define

(16) $$\varphi(x) = x^*(x) - z_0^*[U(x)] \qquad \text{for all } x \in \mathscr{X} .$$

Clearly φ is linear, and, because of (5),

(17) $$x \geqq 0 \qquad \text{implies } \varphi(x) \geqq 0 .$$

Also, φ is bounded, since, for any $x \in \mathscr{X}$,

(18) $$\begin{aligned} |\varphi(x)| = |x^*(x) - z_0^*[U(x)] &\leqq |x^*(x)| + |z_0^*[U(x)]| \\ &\leqq \|x^*\| \cdot \|x\| + \|z_0^*\| \cdot \|U\| \cdot \|x\| \\ &= (\|x^*\| + \|z_0^*\| \cdot \|U\|)\|x\| . \end{aligned}$$

Thus

(19) $$\varphi \in \mathscr{X}^*, \qquad \varphi \geqq 0_x^* .$$

Now define

(20) $$\psi(y) = \psi((z, x)) = z_0^*(z) + \varphi(x) \qquad \text{for all } z \in \mathscr{Z} \text{ and all } x \in \mathscr{X} ,$$

which is linear in y, since

(21.1) $$\begin{aligned} \psi(\alpha y) = \psi(\alpha(z, x)) = \psi((\alpha z, \alpha x)) &= z_0^*(\alpha z) + \varphi(\alpha x) \\ &= \alpha z_0^*(z) + \alpha\varphi(x) = \alpha\psi(y) \end{aligned}$$

and

(21.2) $$\begin{aligned} \psi(y' + y'') = \psi((z' + z'', x' + x'')) &= z_0^*(z' + z'') + \varphi(x' + x'') \\ &= z_0^*(z') + z_0^*(z'') + \varphi(x') + \varphi(x'') \\ &= \psi(y') + \psi(y'') . \end{aligned}$$

Also,

(22) $$y \geqq 0 \text{ implies } \psi(y) \geqq 0$$

since if $(z, x) \geqq 0$ then $z \geqq 0$ and $x \geqq 0$ and both z_0^* and φ are non-negative functionals.

Finally, ψ is continuous. For let $(z_n, x_n) = y_n \to y_0 = (z_0, x_0)$, $n = 1, 2, \cdots$. Then, by II.6.1, eq. (1), $z_n \to z_0$ and $x_n \to x_0$. Hence, since z_0^* and φ are continuous, $z_0^*(z_n) \to z_0^*(z_0)$ and $\varphi(x_n) \to \varphi(x_0)$, and therefore, $\varphi(y_n) \to \varphi(y_0)$.

Hence

(23) $$\psi \in \mathscr{Y}^* , \qquad \psi \geqq 0_y^* .$$

Because of (16), we have

(24) $$x^*(x) = z_0^*[U(x)] + \varphi(x) \qquad \text{for all } x \in \mathscr{X} ,$$

i.e., by (20),

(25) $$x^*(x) = \psi[(U(x)), x)] \\ = \psi[T(x)] \qquad \text{for all } x \in \mathscr{X} ,$$

or

(26) $$x^* = T^*(\psi) , \qquad \psi \geqq 0_y^* .$$

Now (26) holds for all $x^* \in \mathscr{X}^*$; it follows from Theorem III.4 that the set X_T^* is regularly convex. Thus we have shown that, at least in normed spaces, given the other hypotheses under A in Theorem IV.1, *the assumption of regular convexity of* X_T^* *is necessary* (as well as sufficient) for the validity of the conclusions. We may state this as

THEOREM IV.2.

A. *Let* $\mathscr{X}$ *and* $\mathscr{Z}$ *be normed spaces,* U *a linear bounded transformation on* $\mathscr{X}$ *to* $\mathscr{Z}$, P_x *and* P_z *closed convex cones in* $\mathscr{X}$ *and* $\mathscr{Z}$ *respectively. Then the condition that the set*

(27) $$X_T^* = \{x^* \in \mathscr{X}^* : x^* = T^*(y^*), y^* \geqq 0\}$$

be regularly convex is equivalent to the following: for any $x^* \in \mathscr{X}^*$, *if*

(28) $$U(x) \geqq 0, x \geqq 0 \quad \textit{imply } x^*(x) \geqq 0 \quad \textit{for all } x \in \mathscr{X} ,$$

then there exists a $z_0^* \geqq 0$ *such that*

(29) $$z_0^*[U(x)] \leqq x^*(x) \qquad \textit{for } x \geqq 0$$

and

(30) $$x^*(x) = 0, U(x) \geqq 0, x \geqq 0 \quad \textit{imply } z_0^*[U(x)] = 0 .$$

IV.5. The following result generalizes Theorem IV.1 to situations where non-homogeneous inequalities appear.

THEOREM IV.3.

A. *Let all the hypotheses under* A *in Theorem* IV.1 *hold, the transformation* T *being defined in* (37), (38) *below.*

B. *It follows that if, for some* $\bar{x} \in \mathscr{X}$,

(31) $$\bar{x} \geqq 0 \textit{ and } U(\bar{x}) - a \geqq 0 ,$$

and if, for some $x^* \in \mathscr{X}^*$,

(32) $$x \geqq 0 \textit{ and } U(x) - a \geqq 0 \quad \textit{imply } x^*(x) - \beta \geqq 0 ,$$

then there exists a $z_0^* \geqq 0$ *such that*

(33) $$z_0^*[U(x) - a] \leqq x^*(x) - \beta \qquad \textit{for } x \geqq 0$$

and

(34) $$x^*(x) = \beta, U(x) - a \geqq 0, x \geqq 0 \quad \textit{imply } z_0^*[U(x) - a] = 0$$

PROOF. Consider the product space

(35) $$\mathscr{W} = \{w : w = (\rho, x), \rho \text{ real}, x \in \mathscr{X}\}$$

and the linear transformation

(36) $$P(w) = P((\rho, x)) = -a\rho + U(x) .$$

on $\mathscr{W}$ into $\mathscr{Z}$.

Then define

(37) $$T = \begin{pmatrix} P \\ I \end{pmatrix}, \qquad I(w) = w \qquad \text{for all } w \in \mathscr{W},$$

i.e.,

(38) $$T(w) = (P(w), w) \qquad \text{or} \qquad T((\rho, x)) = (-a\rho + U(x), (\rho, x))$$

and T is a linear transformation on $\mathscr{W}$ into $\mathscr{X} \times \mathscr{W}$.

Now suppose we have shown that

(39) $$T(w) \geqq 0 \quad \text{implies } w^*(w) \geqq 0 \quad \text{for all } w \in \mathscr{W}$$

where we define w^* by

(40) $$w^*(w) = w^*((\rho, x)) = -\beta\rho + x^*(x) .$$

One can ascertain easily that Theorem IV.1 applies, with w replacing x, w^* replacing x^*, and P replacing U.

Hence there exists a $z_0^* \geqq 0$ such that

(41) $$z_0^*[P(w)] \leqq w^*(w) \qquad \text{for } w \geqq 0$$

and

(42) $$w^*(w) = 0, T(w) \geqq 0 \qquad \text{imply} \qquad z_0^*[P(w)] = 0 .$$

Equation (41), written out explicitly, yields, by (36) and (40),

(43) $$z_0^*[-a\rho + U(x)] \leqq -\beta\rho + x^*(x) \qquad \text{for } \rho \geqq 0, x \geqq 0 .$$

Letting $\rho = 1$ we obtain (33).

Similarly, using (36), (40), and (38) in (42), and putting $\rho = 1$, we obtain (34).

Therefore, it remains to establish (39) which, written out explicitly, states that

(44) $$\left.\begin{aligned} -a\rho + U(x) &\geqq 0 \\ \rho &\geqq 0 \\ x &\geqq 0 \end{aligned}\right\} \text{ imply } \; -\beta\rho + x^*(x) \geqq 0 .$$

Suppose (44) is false. Then the hypotheses of (44) must hold and the conclusion fail for some $\rho_0 \geqq 0$, $x_0 \geqq 0$. We shall first consider the case $\rho_0 > 0$. I.e., we have

(45.1) $$\begin{aligned} -a\rho_0 + U(x_0) &\geqq 0 \\ \rho_0 &> 0 \\ x_0 &\geqq 0 \end{aligned}$$

and

(45.2) $$-\beta\rho_0 + x^*(x_0) < 0 ,$$

so that

(46.1) $$\begin{aligned} -a + U\left(\frac{x_0}{\rho_0}\right) &\geqq 0 \\ \frac{x_0}{\rho_0} &\geqq 0 \end{aligned}$$

and

(46.2) $$-\beta + x^*\left(\frac{x_0}{\rho_0}\right) < 0 \ .$$

This, however, violates (32). Hence the implication in (44) has been established for $\rho > 0$. We shall now take up the case $\rho = 0$. I.e., we must show that

(47) $$\left.\begin{array}{r} U(x) \geqq 0 \\ x \geqq 0 \end{array}\right\} \text{ imply } x^*(x) \geqq 0 \ .$$

Let x_1 satisfy the hypotheses of (47) and take a real $\lambda > 0$. Then, by (31),

(48) $$-[a - U(\bar{x})]\lambda + U(x_1) \geqq 0$$

and hence

(49) $$-a\lambda + U(x_1 + \lambda\bar{x}) \geqq 0 \ .$$

Note also that

(50) $$x_1 + \lambda\bar{x} \geqq 0 \ .$$

Hence, for $\rho_0 = \lambda$, $x_0 = x_1 + \lambda\bar{x}$, the hypotheses of (45.1) are satisfied, so that

(51) $$-\beta\lambda + x^*(x_1 + \lambda\bar{x}) \geqq 0 \ .$$

We therefore have

(52) $$x^*(x_1) + \lambda[x^*(\bar{x}) - \beta] \geqq 0 \qquad \text{for all } \lambda > 0 \ .$$

Suppose now that

(53) $$x^*(x_1) = -\varepsilon < 0 \ .$$

Then (52) is false for any $\lambda > 0$ if $x^*(\bar{x}) - \beta \leqq 0$.
Hence suppose

(54) $$x^*(\bar{x}) - \beta = \eta > 0 \ .$$

and take $\lambda = \varepsilon/(2\eta)$. Then (52) becomes

(55) $$-\varepsilon + \frac{\varepsilon}{2\eta}\eta > 0 \ ,$$

i.e.,

(56) $$-\frac{\varepsilon}{2} > 0$$

which contradicts (53). Hence

(57) $$x^*(x_1) \geqq 0$$

which establishes the validity of (47).

IV.6. Consider now the special case of Theorem IV.3, where $P_x = \mathscr{X}$, so that the restriction $x \geqq 0$ is necessarily satisfied for all x. In this case 0_x^* is the only non-negative element of $\mathscr{X}^*$. [For otherwise there would be some $x_0^* \in \mathscr{X}^*$ with $x_0^*(x_0) > 0$ for some $x_0 \in \mathscr{X}$, hence $x_0^*(-x_0) < 0$ even though $-x_0 \in P_x$, which contradicts $x_0^* \geqq 0$.] Now the counterpart of (12) for Theorem IV.3 is

(58) $$-\beta\rho + x^*(x) = z_0^*[-a\rho + U(x)] + \tau_0^*\rho + x_0^*(x) \qquad \text{for all } \rho \text{ and } x \in \mathscr{X}.$$

When $P_x = \mathscr{X}$, it follows that $x_0^*(x) = 0$ and (58) becomes

(59) $\quad -\beta\rho + x^*(x) = z_0^*[-a\rho + U(x)] + \tau_0^*\rho \qquad$ for all ρ and all $x \in \mathscr{X}$.

Putting $\rho = 0$, (59) reduces to

(60) $$x^*(x) = z_0^*[U(x)] \qquad \text{for all } x \in \mathscr{X}.$$

Furthermore, if $U(\hat{x}) - a \geqq 0$ and $x^*(\hat{x}) \leqq x^*(x)$ for all $U(x) - a \geqq 0$, then (34) in Theorem IV.3 yields

(61) $$z_0^*[U(\hat{x}) - a] = 0 .$$

which, by (60), implies

(62) $$x^*(\hat{x}) = z_0^*(a) .$$

Hence, with $\beta \leqq x^*(\hat{x})$ by hypothesis, we have

(63) $$x^*(\hat{x}) = z_0^*(a) \geqq \beta ,$$

as in Dantzig's Corollary ([10], p. 334).

We may state these results as

COROLLARY IV.3.

A. *Let* $\mathscr{X}$ *and* $\mathscr{Z}$ *be locally convex linear spaces,* U *a linear transformation on* $\mathscr{X}$ *to* $\mathscr{Z}$, P_z *a closed convex cone in* $\mathscr{Z}$, *and assume that the set*

(64) $$X_U^* = \{x^* \in \mathscr{X}^* : x^* = U^*(z^*), z^* \geqq 0\}$$

is regularly convex.

B. *It follows that if, for some* $\bar{x} \in \mathscr{X}$,

(65) $$U(\bar{x}) - a \geqq 0 ,$$

and if, for some $x^* \in \mathscr{X}^*$,

(66) $$U(x) - a \geqq 0 \qquad \textit{implies } x^*(x) - \beta \geqq 0 ,$$

then there exists a $z_0^* \geqq 0$ *such that*

(67) $$z_0^*[U(x)] = x^*(x) \qquad \textit{for all } x \in \mathscr{X} ;$$

furthermore

(68) $$\min_{U(x) \geqq a} x^*(x) = z_0^*(a) \geqq \beta .$$

It will be noted that, in Banach spaces at least, the assumption of regular convexity of X_U^* is necessary as well as sufficient if U is bounded; this follows from Theorem IV.2.

In finite-dimensional Euclidean spaces the requirement of regular convexity of X_U^* is necessarily satisfied (cf. III.9) and if $z \geqq 0$ means, as usual, that each of its coordinates is non-negative, then P_z is closed. Hence the hypotheses of the regular convexity of X_U^* and the closure of P_z may be omitted, and we obtain, as a special case of Corollary IV.3, the Lemma (and its Corollary) stated by Dantzig in [10], p. 334. This, of course, suggests the possibility of generalizing the Dantzig result on the equivalence of linear programming and game problems, since the Lemma plays a crucial role in Dantzig's proof and Corollary IV.3 above provides its generalization.

V. The Lagrangian Saddle-Point Theorem

V.1. *Isotone Lagrangian saddle-point implies maximality.*

V.1.0. Contrary to the customary sequence, we find it more convenient to start with the theorem indicated in the title, rather than with one in which the implication goes in the opposite direction. This is done in order that the reason for requiring that the functional on $\mathscr{Y}$ be strictly isotone and that on $\mathscr{Z}$ isotone may become more readily apparent.

V.1.1. Let π_{11} denote the partial ordering maximization problem obtained if in π_1 of II.3.3 the following two requirements are added:

(1) $\mathscr{Z}$ is a locally convex linear topological space;

(2) P_z is closed.

V.1.2. THEOREM V.0. *Let the generalized Lagrangian expression $\Phi_{\pi_{11}}(x, z^* ; \eta_0^*)$ have an isotone saddle-point at $(x_0, z_0^* ; \eta_0^*)$. Then x_0 is maximal.*

For the sake of convenience, we give a more explicit statement of the preceding theorem.

THEOREM V.1. *Let $x \in \mathscr{X}$, while the values of $f(x)$ and $g(x)$ are in $\mathscr{Y}$ and $\mathscr{Z}$, respectively, where $\mathscr{Y}$ is ordered by a relation written as $\geqq$ and $\mathscr{Z}$ is a locally convex, linear topological space such that $z \geqq 0$ means $z \in P_z$, P_z being a closed convex cone.*

Write

$$(1) \qquad \Phi(x, z^*) = \eta_0^*[f(x)] + z^*[g(x)]$$

where η_0^* is a strictly isotone functional on $\mathscr{Y}$ and z^* is linear continuous functional on $\mathscr{Z}$. Here $z^* \geqq 0$ means that $z^*(z) \geqq 0$ for all $z \geqq 0$.

Suppose that, for some $x_0 \in X$, $z_0^* \geqq 0$, $(X \subseteq \mathscr{X})$, we have

$$(2) \qquad \Phi(x, z_0^*) \leqq \Phi(x_0, z_0^*) \leqq \Phi(x_0, z^*), \qquad \textit{for all } x \in X \textit{ and all } z^* \geqq 0 .$$

Then

$$(3.1) \qquad g(x_0) \geqq 0$$

and, for all $x \in X$,

$$(3.2) \qquad g(x) \geqq 0, f(x) \geqq f(x_0) \qquad \textit{imply } f(x_0) \geqq f(x).$$

V.1.3. PROOF. The right-hand inequality in (2) implies that

$$(4) \qquad z_0^*[g(x_0)] \leqq z^*[g(x_0)] \qquad \text{for all } z^* \geqq 0 ;$$

hence, in particular,

$$(5) \qquad z_0^*[g(x_0)] \leqq (z_0^* + z_1^*)[g(x_0)] \qquad \text{for all } z_1^* \geqq 0 ,$$

since $z_0^* + z_1^* \geqq 0$ if $z_1^* \geqq 0$. But (5) gives

$$(6) \qquad 0 \leqq z_1^*[g(x_0)] \qquad \text{for all } z_1^* \geqq 0 ,$$

and (3.1) follows from Lemma II.2 in II.2.5 above, based on the Mazur-Bourgin Theorem.

We shall now show the validity of (3.2). Equation (2) yields

(7) $\Phi(x, z_0^*) \leqq \Phi(x_0, z^*)$ for all $x \in X$ and all $z^* \geqq 0$.

Using $z^* = 0_z^*$, this gives

(8) $\eta_0^*[f(x)] + z_0^*[g(x)] \leqq \eta_0^*[f(x_0)]$ for all $x \in X$.

Now let $x' \in X$ be such that the hypotheses of (3.2) are satisfied, i.e.,

(9.1) $$g(x') \geqq 0$$

and

(9.2) $$f(x') \geqq f(x_0) .$$

Suppose that the conclusion of (3.2) is false, i.e.,

(9.3) $$f(x_0) \ngeqq f(x') .$$

Since η_0^* is strictly isotone, (9.2) and (9.3) together imply (cf. II.2.1)

(10) $$\eta_0^*[f(x')] > \eta_0^*[f(x_0)] .$$

Also, since $z_0^* \geqq 0$, (9.1) gives

(11) $$z_0^*[g(x')] \geqq 0 .$$

From (10) and (11) it follows that

(12) $$\eta_0^*[f(x')] + z_0^*[g(x')] > \eta_0^*[f(x_0)] ,$$

which contradicts (8). Hence (9.3) is false and (3.2) follows.

It is important to note, that, in this proof, it would not have been sufficient to assume η_0^* isotone rather than strictly isotone (cf. V.2.6).

V.2. *Scalarization.*

V.2.1. Let $\mathscr{W}$ be a topological linear space and K a convex cone in $\mathscr{W}$. Then there always exists a linear continuous functional non-negative on K, since the null functional [$\psi(w) = 0$ for all $w \in \mathscr{W}$] has this property. However, even with additional assumptions on K (viz., that it is closed and pointed[50]), there may not exist any continuous linear functional strictly positive on K, as shown by example in Krein and Rutman ([29], pp. 21—22). On the other hand, it has been shown (*ibid.*, Theorem 2.1, p. 21) that if K is a closed pointed convex cone and $\mathscr{W}$ a separable Banach space, a linear continuous (equals bounded, in this case) functional strictly positive on K does exist. It is shown below that the requirement of pointedness can be removed (Lemma V.2.2 in V.2.5 below). ["Strictly positive" is defined in II.2.4.]

When a strictly positive functional exists, it may be used to "scalarize" the Lagrangian problem. It has been pointed out in V.1.2 that a non-negative functional is not adequate for our purposes (cf. also V.2.6 below).

V.2.2. Let $P_y \subseteq \mathscr{Y}$ be a convex cone in the linear topological space $\mathscr{Y}$. We write $y' \geqq y''$ if and only if $y' - y'' \in P_y$. $\hat{Y}$ denotes the $\geqq$ —maximal subset of the given permissible set Y.

[50] K is said to be *pointed* if $0_w \neq w \in K$ implies $-w \notin K$.

An element[51] $y_0 \in Y$ may have the property that there exists a $y^*_{y_0} = y^*_0 > 0$ such that

(1) $$y \in Y \text{ implies } y^*_0(y) \leqq y^*_0(y_0) .$$

In the light of the remarks in V.2.1, such a y^*_0 will not always exist in infinite-dimensional spaces. But even in the two-dimensional Euclidean space, where every closed convex cone does possess a strictly positive linear continuous functional, and with P_y chosen as the non-negative quadrant, the required y^*_0 may not exist for some y_0. This, of course, is not surprising since so far nothing has been assumed about the set Y. But even if (as is natural in certain problems) one were to assume Y to be convex, closed, and even bounded, $y^*_0 > 0$ may not exist for certain elements of $\hat{Y}$.[52]

Let $\hat{\hat{Y}}$ denote the subset of $\hat{Y}$ such that $y_0 \in \hat{\hat{Y}}$ if and only if there exists a y^*_0 such that (1) holds.

We shall now formulate a necessary condition for membership in $\hat{\hat{Y}}$.

Suppose that there exists a $y^*_0 > 0$ such that (1) holds for a given $y_0 \in \hat{Y}$. Then

(2) $$y^*_0(y - y_0) \leqq 0 \qquad \text{for all } y \in Y ,$$

i. e.,

(3) $$y^*_0(y) \leqq 0 \qquad \text{for all } y \in Y - y_0 .$$

Furthermore, by definition of strict positiveness,

(4) $$y^*_0(y) \geqq 0 \qquad \text{for all } y \geqq 0$$

and

(5) $$y^*_0(y) > 0 \qquad \text{for all } y \geq 0 .$$

Putting

(6) $$y^*_1 = -y^*_0 ,$$

we may rewrite (3), (4), and (5) as

(7.1) $$y^*_1(y) \geqq 0 \qquad \text{for all } y \in Y - y_0 ,$$

(7.2) $$y^*_1(y) \geqq 0 \qquad \text{for all } y \leqq 0,$$

(7.3) $$y^*_1(y) > 0 \qquad \text{for all } y \leq 0 ,$$

respectively.

Now consider the intersection K_0 of all convex cones containing the set

(8) $$(Y - y_0) \cup (-P_y) .$$

Clearly, K_0 is a convex cone, and furthermore

(9) $$y^*_1(y) \geqq 0 \qquad \text{for all } y \in K_0 .$$

[This follows from the fact that the set $\{y : y^*_1(y) \geqq 0\}$ is a convex cone

[51] The element y_0 in this context need not be $\geqq$ −maximal: cf. Theorem V.2.3.

[52] Arrow's example: $Y = \{y: y = (y_1, y_2), y_i \geqq 0, y_1^2 + y_2^2 \leqq 1\}, y_0 = (0, 1)$. Cf. also Kuhn and Tucker [31], p. 488, example.

which, by (7.1) and (7.2), contains $(Y - y_0) \cup (-P_y)$, and hence, by definition of K_0, it includes K_0.]

Writing $\bar{K}_0$ to denote the closure of K_0, we also have

(10) $$y_1^*(y) \geqq 0 \quad \text{for all } y \in \bar{K}_0 ,$$

by continuity of y_1^*.

[For linear normed (hence Banach) spaces, this has been noted in Krein and Rutman ([29], pp. 16–17). When $\mathscr{Y}$ is a linear topological space, (10) is proved as follows: Let $y' \in \bar{K}_0$ and suppose $y_1^*(y) < 0$, say $y_1^*(y) = -\alpha$. The inverse image by y_1^* of the open interval $(-3\alpha/2, -\alpha/2)$ is an open set containing y', hence containing at least one point, say y'' of K_0. Thus $y_1^*(y'') < -\alpha/2 < 0$, which contradicts (9).]

Now suppose that there exists an element y' with

(11.1) $$y' \in \bar{K}_0$$

and

(11.2) $$y' \geq 0 .$$

Define

(12) $$y'' = -y' ,$$

so that

(13) $$y'' \leq 0 .$$

Then, by (7.3),

(14) $$y_1^*(y'') > 0 .$$

But because of (11.1) and (10),

(15) $$y_1^*(y') \geqq 0 ,$$

hence

(16) $$y_1^*(y'') \leqq 0 ,$$

which contradicts (14). Hence we have

THEOREM V.2.1. *Let $\mathscr{Y}$ be a linear topological space and P_y a convex cone. If, for*[53] $y_0 \in Y$, *there exists* $y_0^* > 0$ *such that* (1) *holds, then the set* $\bar{K}_0$ [*the closure of the intersection of all the convex cones containing the set* $(Y - y_0) \cup (-P_y)$] *does not contain any* y' *such that* $y' \geq 0$.

V.2.3. *Definition.* If $\bar{K}_0$ contains no $y' \geq 0$ and y_0 is $\geqq$ −maximal, y_0 is said to be *properly maximal.*

V.2.4. THEOREM V.2.2. *Let $\mathscr{Y}$ be a linear topological space with the property that for every closed convex cone $K \subseteq \mathscr{Y}$, there is a linear continuous functional $y^* \in \mathscr{Y}^*$ strictly positive on K.*

Then, for every y_0 properly maximal, there exists a $y_0^* > 0$ such that (1) *is satisfied.*

PROOF. By hypothesis, there exists y_1^* strictly positive on $\bar{K}_0$. Then

[53] In this Theorem, y_0 need not be $\geqq$ −maximal. But Theorem V.2.3 asserts that y_0 must be $\geqq$ −maximal.

(7.1) and (7.2) are satisfied because $\bar{K}_0$ contains the sets $Y - y_0$ and $-P_y$ and because $y_1^* \in (\bar{K}_0)^{\oplus}$. Now take $y' \leq 0$. Then $y' \in \bar{K}_0$. Suppose $-y' \in \bar{K}_0$ also. Then, since y_0 is properly maximal, $-y' \not\geq 0$, which contradicts $y' \leq 0$. Hence, $-y' \notin \bar{K}_0$. But then, because of $y_1^* > 0$, $y_1^*(y') > 0$; i. e., (7.3) also holds. It is seen that

(17) $$y_0^* = -y_1^*$$

has the required property.

V.2.5. COROLLARY V.2.2. *Let $\mathscr{Y}$ be a separable linear normed space and y_0 properly maximal. Then there exists a $y_0^* > 0$ such that* (1) *holds.*

PROOF. In view of Theorem V. 2. 2, it will suffice to prove the following:

LEMMA V.2.2. *For every closed convex cone K in a linear normed separable space $\mathscr{Y}$, there is a linear bounded functional strictly positive on K.*

To prove the Lemma, we first note that Theorem 2.1, p. 21, in Krein and Rutman [29] is precisely equivalent to our Lemma for the case where K is pointed, i.e., where $0 \neq w \in K$ implies $-w \notin K$. Hence, it is sufficient to show that the Krein-Rutman proof can be extended to cover the case of K not assumed pointed.

Now the pointedness of K is not used in the Krein-Rutman proof in reaching the conclusion that there exists a y_0^* (in their notation f_0) such that

(18) $$y_0^* \in K^{\oplus}$$

and

(19) $$y_0 \in K,\ y_0^*(y_0) = 0 \quad \text{imply } y^*(y_0) = 0 \quad \text{for all } y^* \in K^{\oplus}\ .$$

We may now use Theorem 1.4, p. 17, in Krein and Rutman [29] which asserts[54] that for every $w_0 \in C$ where C is a closed convex cone in $\mathscr{W}$ and $-w_0 \notin C$, there exists $w_0^* \in C^{\oplus}$ such that $w_0^*(w_0) > 0$. This Theorem, together with (19), yields the conclusion that

(20) $$y_0 \in K,\ y_0^*(y_0) = 0 \quad \text{imply } -y_0 \in K$$

which, with (18), makes y_0^* strictly positive on K.

V.2.6. THEOREM V.2.3. *Let $\mathscr{Y}$ be a topological linear space and y_0^* strictly positive on the convex cone P_y, and let $y_0 \in Y$ be such that* (1) *holds, i.e., that $y \in Y$ implies $y_0^*(y) \leqq y_0^*(y_0)$. Then y_0 is maximal in Y.*

PROOF. Suppose not. Then, for some $y' \in Y$ we have

(21) $$y' \geq y_0\ .$$

Also, by (1), since $y' \in Y$,

[54] This Theorem follows from the Mazur-Bourgin Theorem (II.1.4.1) whenever $\mathscr{W}$ is locally convex.

(22) $$y_0^*(y' - y_0) \leqq 0 \, ,$$
while (21) together with $y_0^* > 0$ yields

(23) $$y_0^*(y' - y_0) > 0 \, .$$

The contradiction between (22) and (23) completes the proof.

It would not have been enough to assume $y_0^* \in P_y^{\oplus}$ (or even $y_0^* \geq 0_y^*$) instead of $y_0^* > 0$. For in that case (21) would only have yielded

(23′) $$y_0^*(y' - y_0) \geqq 0 \, ,$$

which does not contradict (22) since equality could hold in both. (E.g., in the case of P_y closed, $y' - y_0$ could be a boundary point of P_y with $y_0^*(y' - y_0) = 0$.)

V.2.7. Noting that the hypotheses of Theorems V.2.1 and V.2.3 are identical, we summarize the results of V.2 in Theorem V.2.4.

THEOREM V.2.4. *Let $\mathscr{Y}$ be a linear topological space ordered by the relation $\geqq$ (where $y' \geqq y'$ means $y' - y'' \in P_y$, P_y being a convex cone).*

A. *If there exists $y_0^* > 0$ such that* (1) *holds for some $y_0 \in Y$, then y_0 is properly $\geqq$ −maximal in Y.*[55]

B. *If for every closed convex cone $K \subseteq \mathscr{Y}$ there is a linear continuous functional $y^* \in \mathscr{Y}^*$ strictly positive on K (as is, for instance, the case in a separable linear normed space), then for every y_0 properly $\geqq$ −maximal there exists a $y_0^* > 0$ such that* (1) *is satisfied.*

V.3. *Maximality implies existence of a saddle-point.*

V.3.1. *Lagrangian saddle-points without differentiability.*

V.3.1.1. The basic idea of the Theorem presented in this section goes back to Slater's paper entitled "Lagrange Multipliers Revisited" [41]. The chief accomplishment of Slater's paper was to establish the existence of a saddle-point for the Lagrangian expression without using differentiability properties in any way whatever, the reliance being placed on the concavity properties of the relevant functions. (A more detailed comparison is given at the end of Part I of the present chapter.) Since the differentiability approach also used the concavity properties, Slater's result was a significant improvement. The present writer extended Slater's result (except for a slight strengthening of Slater's concavity requirements to conform with the usual ones) in a Cowles Commission Discussion Paper (Economics No. 2110) of September 1954. The present version differs significantly from the 1954 version. A suggestion, due to Hirofumi Uzawa, has made it possible not only to simplify the proof tremendously, but also to weaken the assumptions on the functions used (which are merely concave, but not necessarily continuous) and on the underlying spaces.

V.3.1.2. THEOREM V.3.1. *Let $\mathscr{X}$ be a linear system, $\mathscr{Y}$ and $\mathscr{Z}$ linear topological spaces. P_y, P_z are convex cones in $\mathscr{Y}$ and $\mathscr{Z}$ with non-empty*

[55] The point y_0 is $\geqq$ −maximal by V.2.3; proper maximality then follows from V.2.1.

interiors, $P_y \neq \mathscr{Y}$, *X a (fixed) convex subset of* $\mathscr{X}$, *f a concave function on X to* $\mathscr{Y}$, *g a concave function on X to* $\mathscr{Z}$. *Let there be a point* x_* *in X such that*

(1) $g(x_*) > 0$ (*i.e.,* $g(x_*)$ *is an element of the interior of* P_z).

If x_0 *maximizes* $f(x)$ *subject to* $g(x) \geqq 0$ *and* $x \in X$, *then there exist linear continuous functionals*

(2) $$y_0^* \geq 0 \quad \text{and} \quad z_0^* \geqq 0$$

such that, for the Lagrangian expression

(3) $$\Phi(x, z^*) = y^*[f(x)] + z^*[g(x)] ,$$

the saddle-point inequalities

(4) $$\Phi(x, z_0^*) \leqq \Phi(x_0, z_0^*) \leqq \Phi(x_0, z^*)$$

hold for all $x \in X$ *and all* $z^* \geqq 0$.

(We may note that in applications X is usually a convex cone—e.g., the non-negative orthant of the system $\mathscr{X}$.)

PROOF. Let $\mathscr{W}$ be the topological product space $\mathscr{Y} \times \mathscr{Z}$ and consider the subset of $\mathscr{W}$ defined by

(5) $$A = \{(y, z) : y \in \mathscr{Y}, y \leqq f(x), z \in \mathscr{Z}, z \leqq g(x) \quad \text{for some } x \in X\}.$$

The set A is convex because of the concavity of the functions f and g and the convexity of the set X. Also, A has interior points because P_z and P_y have non-empty interiors.

Consider the point $(f(x_0), 0_z) = w_0$ of the space $\mathscr{W}$. The point w_0 is an element of A, since, by hypothesis, $0_z \leqq g(x_0)$. On the other hand, w_0 does not belong to the interior of A; for if w_0 were interior to A, there would exist an element x in X such that $f(x_0) < f(x)$ and $0_z \leqq g(x)$, which cannot happen because of the assumed maximality of x_0.

Hence, we may apply the Corollary of the Hahn-Banach (Bounding Plane) Theorem (see Corollary II. 1 of II.1.4 above) and obtain a non-null functional w_0^* such that

$$w_0^*(w) \leqq w_0^*(w_0) \quad \text{for all } w \in A .$$

Writing $w_0^* = (y_0^*, z_0^*)$, this implies

(6) $$y_0^*(y) + z_0^*(z) \leqq y_0^*[f(x)] \quad \text{for all } (y, z) \text{ in } A .$$

Since $(f(x), g(x))$, with x in X, belongs to A, we have in particular

(7) $$y_0^*[f(x)] + z_0^*[g(x)] \leqq y_0^*[f(x_0)] \quad \text{for all } x \text{ in } X .$$

Also, since $w_0 = (f(x_0), 0_z)$ is in A, it follows that all ordered pairs of the form $(f(x_0), z)$ are in A if $z \leqq 0_z$, which implies $z_0^*(z) \geqq 0$ for all $z \geqq 0$, i.e.,

(8) $$z_0^* \geqq 0 .$$

Similarly, because w_0 is in A, so are all pairs of the form $(y, 0_z)$ for $y \leqq f(x_0)$; this implies

(9) $$y_0^* \geqq 0 .$$

Now suppose $y_0^* = 0_y^*$ (the null functional). It follows from (7) that $z_0^*[g(x)] \leqq 0$ for all x in X, hence for x_*. Also, since w_0^* is non-null, $z_0^* \geq 0$. But then $z_0^*[g(x_*)] = 0$ because $g(x_*)$ was assumed positive and z_0^* is non-negative. However, since $g(z_*)$ is an interior point of the cone P_z and z_0^* is non-null, it must be that $z_0^*[g(x_*)] > 0$. (This follows from an extension of Prop. 5, Bourbaki [7], p. 75, to the case where the cone need not be pointed and the space is merely assumed linear topological; that this extension is valid follows directly from Bourbaki [7], Prop. 16, p. 52.) Hence we have established that

(10) $$y_0^* \geq 0 .$$

Now let $x = x_0$ in (7). It follows that

(11) $$z_0^*[g(x_0)] \leqq 0 .$$

On the other hand, since both $g(x_0)$ and z_0^* are non-negative,

(12) $$z_0^*[g(x_0)] \geqq 0 ,$$

hence

(13) $$z_0^*[g(x_0)] = 0 .$$

Since $z^*[g(x_0)] \geqq 0$ for all $z^* \geqq 0$, (13) implies the right-hand saddle-point inequality, while (7) and (13) yield the left-hand inequality. This completes the proof.

V.3.1.3. A case of particular interest is that of $\mathscr{Y}$ being the space of reals. Here $y_0^* \geq 0$ is equivalent to $y_0^* > 0$ and we have a strictly isotone functional of the type needed in Theorem V.1.1 (saddle-point implies maximality). When $\mathscr{Y}$ is multi-dimensional, however, y_0^* is not strictly isotone but merely isotone, which is inadequate in the context of Theorem V.1, for example, to establish "efficiency" of a given resource allocation.

V.3.1.4. It was shown by Slater that the condition $g(x_*) > 0$ cannot be dispensed with. [In his counter-example, all three spaces are one-dimensional (reals), $f(x) = x - 1$, $g(x) = -(x-1)^2$.] A slight modification of Slater's counter-example shows that the condition $g(x_*) \geq 0$ is not sufficient: we again take $f(x) = x - 1$, $\mathscr{Y}$ two-dimensional, with $g_1(x) = -(x-1)^2$, $g_2(x) = -x + 2$.

V.3.2. *Non-negative Lagrangian saddle-points: the linear non-homogeneous case.*[56]

V.3.2.1. Although linear non-homogeneous situations may be handled by theorems covering the non-linear situations as well, it seems more helpful and simpler to give the direct proofs based on the assumption of linearity.

V.3.2.2. We consider the problem of maximizing the linear non-homogeneous real-valued function on $\mathscr{X}$ to $\mathscr{Y}$ ($\mathscr{Y}$ reals)

[56] We call a function $\varphi(x) + y_0$ on $\mathscr{X}$ to $\mathscr{Y}$ *linear non-homogeneous* if $\varphi(x)$ is linear [i.e., if $\varphi(x)$ is homogeneous and additive]. The possibility that y_0 vanishes is not excluded. ("Affine" might be a more appropriate term.)

$$(1) \qquad f(x) = -x^*(x) + \nu \qquad (x^* \in \mathscr{X}^*)$$

subject to the linear constraints

$$(2.1) \qquad g(x) = U(x) - a \geqq 0_z ,$$

$$(2.2) \qquad x \geqq 0_x ,$$

where U is a linear transformation on $\mathscr{X}$ to $\mathscr{Z}$, it being assumed that $\mathscr{X}$ is a linear topological space and $\mathscr{Z}$ a locally convex space, and the convex cones $P_z = \{z: z \geqq 0_z\}$, $P_x = \{x: x \geqq 0_x\}$ are closed.

In this case, the Lagrangian expression (cf. II.3.3) can be written as[57]

$$(3) \qquad \Phi(x, z^*) = [-x^*(x) + \nu] + z^*[U(x) - a] .$$

V.3.2.3. THEOREM V.3.2. *Let $\mathscr{X}$ be a linear topological space, and $\mathscr{Z}$ a locally convex linear space, the convex cones P_z, P_x closed, and assume the regular convexity of*

$$(4) \qquad W_T^* = \{w^* \in \mathscr{W}: w^* = T^*(v^*), v^* \geqq 0, v^* \in \mathscr{V}^*\} ,$$

where T is the linear continuous transformation on the topological linear product space $\mathscr{W}$ of the pairs $w = (\rho, x)$, ρ real, $x \in \mathscr{X}$, into the topological product space $\mathscr{Z} \times \mathscr{W} = \mathscr{V}$ given by

$$(5) \qquad T((\rho, x)) = (-a\rho + U(x), (\rho, x)) \qquad \textit{for all } \rho \textit{ real and all } x \in \mathscr{X} .$$

Then, for x_0 to maximize $f(x)$ subject to the constraints (2), *it is necessary and sufficient that $\Phi(x, z^*)$ have a non-negative saddle-point at (x_0, z_0^*); i.e., for Φ defined by* (3),

$$(6.1) \qquad \Phi(x, z_0^*) \leqq \Phi(x_0, z_0^*) \qquad \textit{for all } x \geqq 0 ,$$

$$(6.2) \qquad \Phi(x_0, z_0^*) \leqq \Phi(x_0, z^*) \qquad \textit{for all } z^* \geqq 0 ,$$

if and only if

$$(7.1) \qquad U(x_0) - a \geqq 0_z ,$$

$$(7.2) \qquad x_0 \geqq 0_x ,$$

and, for any $x \in \mathscr{X}$,

(8) *if* (2.1) *and* (2.2) *hold, then* $-x^*(x) + \nu \leqq -x^*(x_0) + \nu$.

PROOF. In view of Theorem V.1, we need only prove the necessity. Inequality (6.1) may be rewritten as

$$(6.1') \qquad -x^*(x) + z_0^*[U(x) - a] \leqq -x^*(x_0) + z_0^*[U(x_0) - a] \qquad \text{for all } x \geqq 0$$

or as

$$(6.1'') \qquad z_0^*[U(x) - a] \leqq x^*(x) - x^*(x_0) + z_0^*[U(x_0) - a] \qquad \text{for all } x \geqq 0 .$$

Now write

$$(8) \qquad x^*(x_0) = \beta .$$

Then, for any x satisfying (2), (8) yields

[57] In general, the first term of the right member of (3) is $y_0^*[-x^*(x) + \lambda]$. In this case, since we shall always take $y_0^* > 0$, we may put $y_0^* = 1$ without loss of generality. [I.e., $y_0^*(y) = y$ for all $y \in \mathscr{Y}$.]

(9′) $$-x^*(x) \leqq -\beta ,$$
or
(9″) $$x^*(x) - \beta \geqq 0 .$$
Hence the hypotheses of Theorem IV.3 are satisfied[58] and therefore there exists a $z_0^* \geqq 0$ such that
(10.1) $$z_0^*[U(x) - a] \leqq x^*(x) - \beta \qquad \text{for } x \geqq 0$$
and
(10.2) $$z_0^*[U(x_0) - a] = 0$$
since x_0 satisfies the hypotheses of IV (34). Equations (10.1) and (10.2) imply (6.1″), hence (6.1).

Inequality (6.2) may be written as
(6.2′) $$-x^*(x_0) + z_0^*[U(x_0) - a] \leqq -x^*(x_0) + z^*[U(x_0 - a)] \qquad \text{for all } z^* \geqq 0 ,$$
i.e., because of (10.2),
(6.2″) $$z^*[U(x_0) - a] \geqq 0 \qquad \text{for all } z^* \geqq 0 .$$
But (6.2″) must hold because of (7.1) and $z^* \geqq 0$.

V.3.3. *Non-negative Lagrangian saddle-points and quasi-saddle-points : the differentiable case.*

V.3.3.1. In this section all spaces are Banach and the functions f and g are assumed to possess Fréchet differentials. We shall call them *differentiable.* The convex cones P_x, P_y, P_z are assumed closed.

V.3.3.2. *Definitions.* We shall say that the function g on $\mathscr{X}$ into $\mathscr{Z}$ is *regular* at a point $\bar{x} \in \mathscr{X}$ if and only if, for every
(1) $$\xi \in \mathscr{X}, \qquad \xi \neq 0_x$$
such that the equality
(2) $$x = \bar{x} + \xi$$
implies the two inequalities
(3) $$x \geqq 0$$
and
(4) $$\delta g(\bar{x}; \xi) + g(\bar{x}) \geqq 0 ,$$
there exists a function Ψ on the closed (real) interval $[0, 1]$ into $\mathscr{X}$, say $x' = \Psi(t)$ $(0 \leq t \leq 1)$, with the following properties :
(5) (a) $\delta\Psi(t; \tau)$ exists for all $0 \leq t \leq 1$

(b) $\bar{x} = \Psi(0)$

(c) $\Psi(t) \geqq 0$ for $0 \leq t \leq 1$

(d) $g[\Psi(t)] \geqq 0$ for $0 \leq t \leq 1$

(e) $\xi = \delta\Psi(0; \tau)$ with $\tau > 0$.

[58] Note that $\bar{x}$ required in IV.3 exists, for x_0 has this property by definition of maximality.

It is easily seen that the condition of regularity is closely related to the Kuhn and Tucker "constraint qualification" ([31], p. 483). In fact, the assertion concerning Ψ is identical with the corresponding assertion in Kuhn and Tucker, while the conditions (1), (2), (3), (4) under which Ψ must exist are not weaker[59] than the corresponding conditions (5), [31], *loc. cit.* Therefore $\bar{x}$ is necessarily regular in our sense if the Kuhn-Tucker "constraint qualification" is satisfied.

It should also be noted that our condition of regularity is closely related to Goldstine's hypothesis (a) ([15], p. 145) whose relationship to the condition in Bliss ([4], Lemma 76.1, p. 210) is similar to that of our regularity concept to the Kuhn-Tucker "constraint qualification."

V.3.3.3. THEOREM V.3.3.1. A. *Let f be a real-valued differentiable function on the Banach space $\mathscr{X}$, g a differentiable function on $\mathscr{X}$ into the Banach space $\mathscr{Z}$. The cones $P_x = \{x: x \geqq 0\}$ and $P_z = \{z: z \geqq 0\}$ are assumed closed.*

Let x_0 maximize $f(x)$ subject to the constraints $x \geqq 0$, $g(x) \geqq 0$ and suppose g is regular at x_0.

B. *It then follows that the relations*

(6.1) $$x \geqq 0$$

(6.2) $$\delta g(x_0; \xi) + g(x_0) \geqq 0 \qquad (\xi = x - x_0)$$

imply

(7) $$-\delta f(x_0; \xi) \geqq 0.$$

PROOF. Consider the real-valued function $h(t)$, $0 \leq t \leq 1$, of the real variable t, defined by

(8) $$h(t) = f[\Psi(t)] \qquad (0 \leq t \leq 1).$$

[The function Ψ exists since, by virtue of (6), the relations (1), (2), (3), (4) are satisfied and g is assumed regular at x_0.]

Because of 5(b), (c), (d) and the maximality of x_0, $h(t)$ must have a maximum at $t = 0$. It follows that[60] for

(9) $$\tau > 0 ,$$

(10) $$\delta h(0; \tau) = \delta f[\Psi(0) ; \delta\Psi(0, \tau)] \leqq 0 ,$$

whence by 5(e), (7) follows.[61]

THEOREM V.3.3.2. (This Theorem is a generalization of the Kuhn-Tucker Theorem 1 [31], p. 484.)

A. *Let all assumptions under* A *in Theorem* V.3.3.1 *hold. Assume*

[59] Since Kuhn and Tucker impose their conditions only on certain components of x and g, it should be noted that for those components g_i of g on which Kuhn and Tucker impose the constraint (5) ([31], *loc. cit.*), we have $g_i(\bar{x}) = 0$. Hence (4) is not weaker than the first part of Kuhn-Tucker (5).

[60] Using the "function of a function rule" as applied to Fréchet differentials, cf. II.5.2.

[61] Theorem V.3.3.1 is implicit in the Kuhn-Tucker proof of their Theorem 1. The proof is suggested (*mutatis mutandis*) by Goldstine [15]. The writer is indebted to Kenneth J. Arrow for clarification on this point.

further the regular convexity of the set

$$W_T^* = \{w^* \in \mathscr{W}^*: w^* = T^*(v^*),\ v^* \geqq 0,\ v^* \in \mathscr{V}^*\}\ ,$$

where T *is given by* V.3.2(4), *with* U *and* a *as in* (15) *below.*

B. *Then there exists a* $z_0^* \geqq 0$ *such that the Lagrangian expression* $\Phi(x, z^*) = f(x) + z^*[g(x)]$
has a non-negative quasi-saddle-point at $(x_0, z_0^*; y_0^*)$, $y_0^* = 1$, *i.e., it satisfies the following relations*:

(11.1) $\delta_x\Phi((x_0, z_0^*); \xi) \leqq 0$ *for all* $x \geqq 0, x = x_0 + \xi$

(11.2) $\delta_x\Phi((x_0, z_0^*); x_0) = 0$

(12.1) $\delta_{z^*}\Phi((x_0,z_0^*); \zeta^*) = \zeta^*[g(x_0)] \geqq 0$ *for all* $z^* \geqq 0, \zeta^* = z^* - z_0^*$,

(12.2) $\delta_{z^*}\Phi((x_0, z_0^*); z_0^*) = z_0^*[g(x_0)] = 0$.

PROOF. Since $\delta g(x; \xi)$ and $\delta f(x; \xi)$ are additive in ξ, x being fixed, the relations (6.1), (6.2), and (7) of Theorem V.3.3.1 may be rewritten respectively as

(13.1) $x \geqq 0$,

(13.2) $-\delta f(x_0; x) - [\delta g(x_0; x_0) - g(x_0)] \geqq 0$,

and

(14) $-\delta f(x_0; x) - [-\delta f(x_0; x_0)] \geqq 0$.

Since x_0 is assumed maximal, Theorem V.3.3.1 states that (13.1), (13.2) together imply (14). This corresponds to the implication (34) in Theorem IV.3, with the following correspondence:

(15.1) $U(x) = \delta g(x_0; x)$ for all $x \in \mathscr{X}$,

(15.2) $a = \delta g(x_0; x_0) - g(x_0)$,

(15.3) $x^*(x) = -\delta f(x_0; x)$ for all $x \in \mathscr{X}$,

(15.4) $\beta = -\delta f(x_0; x_0)$.

Since all the other hypotheses of Theorem IV.3 are satisfied (in particular, $\bar{x}$ of IV (31) exists since x_0 is maximal and hence has the required properties), there exists a $z_0^* \geqq 0$ such that

(16.1) $z_0^*[\delta g(x_0; x) - (\delta g(x_0; x_0) - g(x_0))] \leqq -\delta f(x_0; x) - (-\delta f(x_0; x_0))$

and

(16.2) $z_0^*[\delta g(x_0; x) - (\delta g(x_0; x_0) - g(x_0))] = 0$ for $x = x_0$.

Equation (16.2) immediately yields

(17) $z_0^*[g(x_0)] = 0$

which is (12.2) in Theorem V.3.3.2.

Since x_0 is maximal,

(18) $g(x_0) \geqq 0$;

hence

(19) $z^* \geqq 0$ implies $z^*[g(x_0)] \geqq 0$.

Hence, because of (17), (12.1) holds for any ζ^* such that $\zeta^* = z^* - z_0^*$, $z^* \geqq 0$.

Using (17) and the additivity of $\delta f(x; \xi)$ and $\delta g(x; \xi)$ as functions of ξ, we may rewrite (16.1) as

(20) $$z_0^*[\delta g(x_0; x - x_0)] \leqq -\delta f(x_0; x - x_0) \qquad \text{for all } x \geqq 0 ,$$

i.e.,[62]

(21) $$\delta f(x_0; x - x_0) + \delta z_0^* g(x_0; x - x_0) \leqq 0 \qquad \text{for all } x \geqq 0 ,$$

which is (11.1) in Theorem V.3.3.2. [$z_0^* g(x) = z_0^*[g(x)]$ for all x.]

Now setting $x = 0$ in (21), we get

(22) $$-\delta_x \Phi((x_0, z_0^*); x_0) \leqq 0 .$$

Rewrite (21) as

(21′) $$\delta_x \Phi((x_0, z_0^*); x) - \delta_x \Phi((x_0, z_0^*); x_0) \leqq 0 \qquad \text{for all } x \geqq 0$$

and suppose that

(23) $$\delta_x \Phi((x_0, z_0^*); x_0) > 0 .$$

But using in (21′) $x = 2x_0$, we reach a contradiction since $\delta_x \Phi((x_0, z_0^*); x)$ is homogeneous in x. Hence the equality sign must hold in (22), and (11.2) in Theorem V.3.3.1 follows.

V.3.3.5. THEOREM V.3.3.3. (This Theorem is a generalization of the "only if" part of the Kuhn-Tucker Theorem 3. The converse—the "if" part of the Kuhn-Tucker Theorem 3—follows from V.1.1.) *Let all the assumptions under* A *in Theorem* V.3.3.2 *hold, and assume further that f and g are concave. Then $\Phi(x, z^*)$ has a non-negative saddle-point at (x_0, z_0^*) where x_0 is the maximal point of the hypothesis.*

PROOF. The Kuhn-Tucker proof of the "only if" part of Theorem 3 [31], p. 487, is valid under our assumptions. For the sake of completeness we reproduce its major steps in our notation. First, if $h(x)$, $x \in \mathscr{X}$, $\mathscr{X}$ Banach, is a concave function with values in a Banach space V, and the ordering relation is given by a *closed*[63] convex cone P_v, we have, for $0 < \theta \leq 1$

(24) $$h(x'') - h(x') \leqq \frac{1}{\theta} \{h[x' + \theta(x'' - x')] - h(x')\} .$$

Now

(25′) $$\delta h(x'; x'' - x') = \lim_{\theta \to 0} \frac{1}{\theta} \{h[x' + \theta(x'' - x')] - h(x')\} .$$

Then, because P_v is closed,

(25″) $$h(x'') - h(x') \leqq \delta h(x'; x'' - x')$$

which corresponds to Lemma 3 in Kuhn and Tucker [31], p. 485. Hence, for $x \geqq 0$ and $z_0^* \geqq 0$ and using (25), since f and g are concave,

---

62 We use the function of a function rule and the fact that, since z^* is linear, $\delta z^*(z_0; \zeta) = z^*(\zeta)$. (Cf. II.5.2 and II.5.1, footnote 29.)

63 Closedness is not used for (24) or (25′), but only for (25″).

$$(26)\qquad \begin{aligned}\Phi(x, z_0^*) &\leqq f(x_0) + z_0^*[g(x_0)] + \delta f(x_0; x - x_0) + z_0^*[\delta g(x_0; x - x_0)]\\ &= \Phi(x_0, z_0^*) + \delta_x\Phi((x_0, z_0^*); x - x_0)\\ &\leqq \Phi(x_0, z_0^*)\end{aligned}$$

where the last inequality is based on (11.1) in Theorem V.3.3.2.

On the other hand, for $z^* \geqq 0$,

$$(27)\qquad \Phi(x_0, z^*) - \Phi(x_0, z_0^*) = (z^* - z_0^*)[g(x_0)] \geq 0$$

by (12.1) in Theorem V.3.3.2. Relations (26) and (27) together imply that Φ has a non-negative saddle-point at (x_0, z_0^*).

V.3.3.6. THEOREM V.3.3.4. (This Theorem is a generalization of the Kuhn-Tucker Theorem 4.)

A. *Let $\mathscr{X}$, $\mathscr{Z}$, P_x, P_z, and g be as in Theorem* V.3.3.2 (*including the assumption of regular convexity of X_T^* and regularity of g but not that of concavity*) *while $\mathscr{Y}$ is a Banach space possessing the property stated at the beginning of Theorem* V.2.2 (*e.g., it would suffice to assume $\mathscr{Y}$ separable*). *Assume further that f is a differentiable function on $\mathscr{X}$ into $\mathscr{Y}$ and also that x_0 is properly maximal.*

B. *Then for some $y_0^* > 0$,*

$$(28)\qquad \Phi_1(x, z^*) = y^*[f(x)] + z^*[g(x)]$$

has a non-negative quasi-saddle-point at $(x_0, z_0^*; y_0^*)$; i.e., the relations (11), (12) *hold with $f(x)$ replaced by $y_0^*[f(x)]$.*

PROOF. Using Theorem V.2.2 with $y_0 = f(x_0)$ we obtain y_0^* such that

$$(29)\qquad y \in Y \quad \text{implies} \quad y_0^*(y) \leqq y_0^*(y_0)$$

for

$$(30)\qquad Y = f(P_x \cap g^{-1}(P_z))\ ;$$

i.e., the function y_0^*f, given by

$$(31)\qquad F(x) = y_0^*[f(x)] \quad \text{for all } x \in \mathscr{X}$$

has a maximum at x_0 subject to

$$(32)\qquad x \geqq 0, \quad g(x) \geqq 0\ .$$

Thus we may use Theorem V.3.3.2 as applied to $F(x)$ and the Theorem follows. ($F(x)$ is differentiable since f is differentiable and so is y_0^*.)

V.3.3.7. THEOREM V.3.3.5. (Generalization of the "only if" part of the Kuhn-Tucker Theorem 6.) *Let all the assumptions under* A *in Theorem* V.3.3.4 *hold, and assume further that f and g are concave. Then the function $\Phi_1(x, z^*)$ as defined by* (28) *has a non-negative saddle-point at $(x_0, z_0^*\ ; y_0^*)$ for some $y_0^* > 0$.*

PROOF. Use Theorem V.3.3.4, then Theorem V.3.3.3 as applied to Φ_1. (Note that F, defined in (31), is concave if f is.)

Note. The converse is found in Theorem V.1 (the "if" part of the Kuhn-Tucker Theorem 6).

V.3.3.8. In Section V.3.3 the spaces have so far been assumed Banach and the differentials Fréchet. It appears, however, that by using a more general concept of a differential (to be called here the MF differential) one can validate the results of V.3.3 for that class of linear topological spaces for which the auxiliary results from previous sections are valid (i.e., locally convex linear, or the alternatives mentioned in II.1.4.1).

The MF differential is that called μ (or μ^*) differential in Michal [38], p. 82, and also defined (later but independently) by Fréchet [12], pp. 64–65.[64] We shall denote this differential by $df(x_0; h)$ if evaluated at x_0 with increment $h \in \mathscr{X}$; $f(x)$ is in $\mathscr{Y}$; $\mathscr{X}$ and $\mathscr{Y}$ are Hausdorff linear spaces. The MF differential is additive and continuous (hence linear) in the increment h and is further characterized by the following *property* (c): There exists a fixed neighborhood W of 0_x such that, given any neighborhood V of 0_y, there is a neighborhood U of 0_x (U depends on W) such that if

$$h \in U, nh \in W ,$$

then

$$n[f(x_0 + h) - f(x_0) - df(x_0; h)] \in V$$

for all positive integers n (or all positive real numbers n).[65] As partly stated in [38] and shown in [12], $df(x_0; h)$ has the important properties of the Fréchet differential; in particular, the "function of a function" rule is valid and the partial differentials are defined in the usual way and are additive. Furthermore, from the remarks and theorems in Michal [36], [37] and [38], it follows that when $df(x_0; h)$ exists, then the Gâteaux differential, i.e.,

$$\lim_{\theta \to 0} \frac{1}{\theta}[f(x_0 + \theta h) - f(x_0)] ,$$

exists and the two are equal.[66] Now let $\mathscr{X}$ and $\mathscr{Z}$ be (say) locally

[64] One could probably also use the slightly more general F or M differentials; cf. Hyers [22], pp. 14–15.

[65] In linear normed spaces the MF differential exists if and only if the Fréchet differential exists, and the two are equal. Property (c) is equivalent to that given at the end of II.5.1. This is shown in [12], pp. 62–64.

[66] In [38], p. 82, it is stated that in what we called linear spaces the μ^* differential (equivalent to the MF differential) is equivalent to what in [38] is called the M_1 differential. In [37], Theorem V, the existence of the M_1 differential (called "the differential") is asserted to imply the existence of the M differential of [36] and the two are equal. Finally, in Theorem 4 of [36] it is stated that the existence of the M differential implies that of the Gâteaux differential and the two are equal. The precise meaning of the limit in the definition of the Gâteaux differential is as follows. Write $g(\theta) = [f(x_0 + \theta h)]/\theta$ and denote by $d_G f(x_0; h)$ the Gâteaux differential at x_0 with increment h. Then for a given neighborhood V of 0_y there exists a real number $\delta > 0$ such that

$$g(\theta) \in d_G f(x_0;\ h) + V \qquad \text{for all } 0 < |\theta| < \delta .$$

The equality $df(x_0;\ h) = d_G f(x_0;\ h)$ follows easily from property (c) above when the "starlike" neighborhood system $\mathscr{U}$ is used (see Bourgin [9], pp. 638–39).

convex Hausdorff linear spaces (cf. II.1.4.1 for possible alternative assumptions) and consider the results of V.3.3 with the MF differential $df(x_0; \xi)$, etc., substituted for the Fréchet differential, $\delta f(x_0; \xi)$, etc., throughout. Theorem V.3.3.1 obviously remains valid, since the MF differential is linear and obeys the "function of a function" rule. Theorem V.3.3.2 also retains its validity since the partial MF differentials have the required properties. In the proof of Theorem V.3.3.3, the two relations (25) remain valid because the MF differential, like the Fréchet differential, equals the Gâteaux differential, and Theorem V.3.3.3 follows. (It can also be shown that relations (11) and (12) in Theorem V.3.3.2 are satisfied at a non-negative saddle-point.) The remaining two theorems of V.3.3 also go through.

It may finally be noted that the MF differential is defined and retains many of its important properties when the domain and the range of the function whose differential is taken are in topological groups (not necessarily linear spaces). Thus a possibility appears of studying a broader class of spaces and their Lagrangian expressions from the viewpoint of differentiability.

REFERENCES

[1] K. J. Arrow, E. W. Barankin, and D. Blackwell, "Admissible Points of Convex Sets," in H. W. Kuhn and A. W. Tucker (eds.), *Contributions to the Theory of Games*. Princeton: Princeton University Press, 1953, pp. 87–91.

[2] K. J. Arrow and L. Hurwicz, "Dynamic Aspects of Achieving Optimal Allocation of Resources" (Abstract), *Econometrica*, XX (1952), 86–87.

[3] S. Banach, *Théorie des Operations Linéaires*. Monografje Matematyczne, Warszawa, Tom I, 1932.

[4] C. A. Bliss, *Lectures on the Calculus of Variations*. Chicago: University of Chicago Press, 1946.

[5] N. Bourbaki, *Topologie Générale* (Eléments de Mathématique, II, Première Partie, Livre III, Chapters I and II). Paris: Hermann et Cie, 1940.

[6] N. Bourbaki, *Topologie Générale* (Eléments de Mathématique, III, Première Partie, Chapters III and IV). Paris: 2d ed., Hermann et Cie, 1951.

[7] N. Bourbaki, *Espaces Vectoriels Topologiques* (Eléments de Mathématique, XV, Première Partie; Livre V, Chapters I and II). Paris: Hermann et Cie, 1953.

[8] N. Bourbaki, *Espaces Vectoriels Topologiques* (Eléments de Mathématique, XVIII, Première Partie; Livre V, Chapters III, IV, V). Paris: Hermann et Cie, 1955.

[9] D. G. Bourgin, "Linear Topological Spaces," *Amer. J. Math.*, LXV (1943), 637–59.

[10] G. B. Dantzig, "A Proof of the Equivalence of the Programming Problem and the Game Problem," in [28], pp. 330–35.

[11] M. Fréchet, "La Notion de Différentielle dans l'Analyse Générale," *Annales Scientifiques de l'École Normale Superieure*, 3me Serie, XLII (1925), 293–323.

[12] M. Fréchet, "La Notion de la Différentielle sur un groupe Abelian," *Portugaliae Mathematica*, VII (1948), 59–72.

[13] M. Fréchet, "La Notion de Différentielle dans l'Analyse Générale," in Maurice Fréchet, *Pages Choisies d'Analyse Générale*. Paris-Louvain, 1953, pp. 180–204.

[14] D. Gale, "Convex Cones and Linear Inequalities," in [28], pp. 287–97.

[15] H. H. Goldstine, "Minimum Problems in the Functional Calculus," *Bull. Amer. Math. Soc.*, XLVI (1940), 142–49.

[16] L. M. Graves, "Riemann Integration and Taylor's Theorem in General Analysis," *Trans. Amer. Math. Soc.*, XXIX (1927), 163–77.

[17] L. M. Graves, "Topics in the Functional Calculus," *Bull. Amer. Math. Soc.*, XLI (1935), 641–62.

[18] F. Hausdorff, "Zur Theorie der linearen metrischen Räume," *J. für die Reine und Angewandte Mathematik*, CLXVII (1932), 294–311.

[19] T. H. Hildebrandt and L. M. Graves, "Implicit Functions and Their Differentials in General Analysis," *Trans. Amer. Math Soc.*, XXIX (1927), 127–53.

[20] E. Hille, *Functional Analysis and Semi-Groups*, Amer. Math. Soc. Colloquium Publications (New York), XXXI (1948).

[21] H. Hotelling, "The General Welfare in Relation to Problems of Taxation and Railway Rates," *Econometrica*, VI (1938), 248–56.

[22] D. H. Hyers, "Linear Topological Spaces," *Bull. Amer. Math. Soc.* LI, 1 (1945), 1–21.

[23] J. L. Kelley, *General Topology*. New York: Van Nostrand, 1955.

[24] V. L. Klee, Jr., "The Support Property of a Convex Set in a Linear Normed Space," *Duke Math. J.*, XV (1948), 767–72.

[25] V. L. Klee, Jr., "Convex Sets in Linear Spaces," *Duke Math. J.*, XVIII (1951), 443–66.

[26] V. L. Klee, Jr., "Convex Sets in Linear Spaces, III," *Duke Math. J.*, XX (1953), pp. 105–11 (esp. p. 106).

[27] T. C. Koopmans, "Analysis of Production as an Efficient Combination of Activities," in [28], pp. 33-97.

[28] T. C. Koopmans (ed.), *Activity Analysis of Production and Allocation*, Cowles Commission Monograph No. 13. New York: Wiley, 1951.

[29] M. G. Krein and M. A. Rutman, "Linear Operators Leaving Invariant a Cone in a Banach Space," *Uspehi Matem. Nauk* (N.S.) 3 (1948), No. 1 (23), pp. 3–95. *Am. Math. Soc. Translation* No. 26 (New York, 1950).

[30] M. Krein, and V. Šmulian, "On Regularly Convex Sets in the Space Conjugate to a Banach Space," *Ann. Math.*, Second Series, XLI (1940), 556-83.

[31] H. W. Kuhn and A. W. Tucker, "Nonlinear Programming," *Proc. Second Berkeley Symposium on Mathematical Statistics and Probability*, 1951, pp. 481–92.

[32] C. Kuratowski, *Topologie* I, Deuxième Edition, Monografie Matematyczne, Tom XX, Warszawa-Wroclaw, 1948.

[33] O. Lange and F. M. Taylor, *On the Economic Theory of Socialism* (ed. B. E. Lippincott). Minneapolis: University of Minnesota Press, 1938.

[34] S. Lefschetz, *Algebraic Topology*, Amer. Math. Soc. Colloquium Publications (New York), XXVII, (1942).

[35] A. P. Lerner, *The Economics of Control*. New York: Macmillan, 1944.

[36] A. D. Michal, "Differential Calculus in Linear Topological Spaces," *Proc. Nat. Acad. Sc. U.S.A.*, XXIV (1938), 340–42.

[37] A. D. Michal, "Differentials of Functions with Arguments and Values in Topological Abelian Groups," *Proc. Nat. Acad. Sc. U.S.A.*, XXVI (1940), 357–59.

[38] A. D. Michal, "Functional Analysis in Topological Group Spaces," *Math. Magazine*, XXI (1947–48), 80–90.

[39] P. C. Rosenbloom, "Quelques Classes de Problèmes Extrémaux," *Bull. de la Soc. Math. de France*, LXXIX (1951), 1–58; LXXX (1952), 183–215.

[40] W. Rudin, *Principles of Mathematical Analysis*. New York: McGraw-Hill, 1953.

[41] M. Slater, "Lagrange Multipliers Revisited: A Contribution to Non-Linear Programming," Cowles Commission Discussion Paper, Math. 403, November 1950.

[42] E. C. Titchmarsh, *The Theory of Functions*, 2d ed. Oxford: Oxford University, Press, 1939

[43] A. Tychonoff, "Ein Fixpunktsatz," *Math. Annalen*, CXI (1935), 767–76.