GPO PRICE $ ____

CFSTI PRICE(S) $ ____

Hard copy (HC) ~~3.00~~

Microfiche (MF) ~~.75~~

ff 653 July 65

UNIVERSITY OF MARYLAND
COMPUTER SCIENCE CENTER

COLLEGE PARK, MARYLAND

FACILITY FORM 602

N66 29572 (ACCESSION NUMBER)	(THRU)
81 (PAGES)	1 (CODE)
CR-75898 (NASA CR OR TMX OR AD NUMBER)	19 (CATEGORY)

Technical Report TR-66-29 May 1966

NsG-398

PARTIALLY ORDERED SPACES
and
NEWTON'S METHOD FOR CONVEX OPERATORS

by

James S. Vandergraft
Research Associate
Computer Science Center
University of Maryland, College Park, Maryland

The work reported here was in part supported by the National Aeronautics and Space Administration under Grant NsG-398 to the Computer Science Center of the University of Maryland..

Ia

ABSTRACT

Title of Thesis: Partially Ordered Spaces and Newton's Method for Convex Operators

James S. Vandergraft, Doctor of Philosophy, 1966

Thesis directed by: Werner C. Rheinboldt, Research Professor

Some general theorems are proved concerning Newton's method applied to convex operators which are defined on partially ordered topological linear spaces. The spaces are examined and various relations between the partial ordering and the topology are discussed. A mean value theorem is proved and is then used to study convex operators. Several convergence theorems for Newton's method are obtained and applications to differential and integral equations are given. Finally, these results are used to simplify a theorem of Kantorovich concerning the convergence of Newton's method applied to operators which are defined on spaces with partially ordered norm.

Ib

ACKNOWLEDGMENTS

I would like to express my gratitude to my thesis advisor, Professor Werner C. Rheinboldt, for his invaluable suggestions and constant encouragement. Also, I would like to thank Dr. James Ortega, of the University of Maryland Computer Science Center, for many helpful discussions, comments, and criticisms.

Finally, I would like to thank the National Aeronautics and Space Administration for their generous financial support under Grant NsG-398 to the Computer Science Center of the University of Maryland.

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION

If F is a nonlinear operator from one topological linear space to another, then Newton's method consists of finding the sequence $\{x_n\}$ defined by

$$x_{n+1} = x_n - (F'[x_n])^{-1} F[x_n] \tag{1}$$

where $F'[x_n]$ is either the Frechet or Gateaux derivative of F at x_n. The first general theorem concerning the convergence of this method was given by Kantorovich in 1948 [13]. Under a number of assumptions on F, and the starting point x_0, he proved that the sequence exists and converges to a solution of the equation $F[x] = 0$. The conditions on F and x_0, which are needed for the proof, involve bounds on $F''[x]$ and on the inverse of $F'[x]$. Because of the complexity of these conditions, it is often impossible to apply the result to practical problems. Moreover, the restrictions on x_0 may limit it to a very small region, as is already shown by real functions. In a later paper [14], Kantorovich re-proved his basic theorem, using a completely different method of proof. He also generalized this proof to operators defined on spaces which have a partially ordered norm. In this generalized theorem, however, the assumption is made that, for a certain auxillary operator, Newton's method produces a sequence which is monotone, bounded, and convergent.

It is well-known that for many operators, the sequence defined by (1) is monotone and bounded. In this case, the proof of convergence may be simplified considerably. Such special theorems have been given by Collatz [7], Kalaba [11], and Greenspan and Parter [10]. All of these results use some sort of convexity assumption on the operator in order to get the monotonicity of the iterates.

In the present work, we have proven some very general convergence theorems for Newton's method applied to convex operators. These theorems contain the results of Collatz, Kalaba, and Greenspan and Parter cited above. The operators we consider are defined on certain classes of topological linear spaces in which a partial ordering is defined. Chapter I is devoted to a study of such spaces and of the operators defined on them. After a review of some basic notions about partially ordered spaces, we define partially ordered topological linear spaces (PTL spaces), and discuss some important relations between the ordering and the topology. The treatment here is similar to Krasnoselski's theory of partially ordered Banach spaces [18], however, we assume that the spaces are only locally convex. In the section on operators in PTL spaces, a mean value theorem is proved and is subsequently used to study convex operators.

Chapter II is then devoted to Newton's method. We first prove a general convergence theorem for convex operators defined on a class of spaces which includes the finite dimensional spaces and the L^p spaces. The results we prove differ from those of Kantorovich in several ways. For example, we assume that $F'[x_n]$ is only a Gateaux derivative, whereas, in Kantorovich's theorem, $F'[x_n]$ is the Frechet derivative. Also, we write (1) in the weaker form

$$F'[x_n](x_{n+1} - x_n) = -F[x_n]$$

and prove that $\left\{ x_n \right\}$ exists and converges, without proving that $F'[x_n]$ has an inverse. In fact, an example is given to show that under our conditions, $F'[x_n]$ need not have an inverse. Several examples are analyzed to show how the theorem can be applied, and some numerical calculations are given. The rate of convergence is shown to be quadratic, or super-linear, provided the derivative of the operator satisfies certain boundedness conditions. Finally, we discuss some possible modifications of Newton's method which still lead to monotone and convergent iterations. We also note the works of Baluev [1,2,3] and Slügin[24,25]. These papers, which are concerned with Chaplygin methods, contain results that are quite similar to some of ours. However, in all cases where a comparison can be made, it is clear that our results are stronger.

In the next section, we consider a class of spaces which includes the spaces of differentiable functions. Because of convergence problems here, we restrict ourselves to operators of the form

$$F[x] = f(x) - L[x]$$

where L is a linear operator and f is a convex operator. The theorems for these operators are of the same type as those of the previous section. Various differential equations are studied, and it is shown that the results of Kalaba in [11] are covered by these theorems.

The last section contains a modification of the second Kantorovich theorem mentioned earlier. Using the previous results, we can actually prove that the auxillary operator has a Newton sequence which is monotone, bounded, and convergent. Hence, the Kantorovich result is simplified considerably, and we obtain an interesting convergence theorem for general (i.e., not necessarily convex) operators.

CHAPTER I

PARTIAL ORDERINGS

Partially Ordered Linear Spaces

The basis for all our discussions will be partially ordered linear spaces. Many of the following results can be found in Birkhoff [6], Namioka [20], or Schaefer [21], but we include them here for the sake of completeness.

Definition 1. Let X be a real linear space in which a binary relation $\leq$ is defined between certain elements such that;

1) $x \leq x$ for all x in X
2) if $x \leq y$ and $y \leq x$ then $x = y$
3) if $x \leq y$ and $y \leq z$ then $x \leq z$
4) if $x \leq y$ then $x+z \leq y+z$ for any z in X
5) if $x \leq y$ then $\alpha x \leq \alpha y$ for any positive α.

Then X is called a partially ordered linear space, (PL space).

Conditions 1),2),3) say the ordering is reflexive, anti- symmetric, and transitive. The last two properties provide a connection between the order structure and the linear structure of the space.

If a and b are elements in a PL space, and $a \leq b$, then the closed interval $[a,b]$ is the set $\{x : a \leq x \leq b\}$. If S is any subset of X, then an element u in X is called an upper bound on S if $x \leq u$ for all x in S. Similarly, a lower bound on S is any v in X such that $v \leq x$ for all x in S. An element u in X is called a supremum of S if, u is an upper bound and, moreover, if v is another upper bound on S, then $u \leq v$. Similarly, an infimum of S is a lower bound v, such that $v \geq u$ where u is any other lower bound. It is clear from condition 2) of the definition that a set can have at most one infimum and supremum.

A subset is called <u>order-bounded</u> if it has both an upper bound and a lower bound.

In a PL space, the set $K = \{x : x \geq 0\}$ has the properties:

1) $K + K \subset K$,
2) $\alpha K \subset K$ for any positive α,
3) $K \cap (-K) = \{0\}$

These follow easily from definition 1. A subset of a linear space which satisfies 1) and 2) is called a (convex) <u>cone</u>, and if 3) is also fulfilled, the cone is called <u>proper</u>. The set K defined above is called the <u>positive cone</u> for the PL space.

It is important to observe the duality between proper cones and partial orderings. We have just seen that the order relation in a PL space defines a proper cone. Conversely, if we are given a proper cone K in a linear space X, then the order relation defined by

$$x \leq y \text{ if and only if } y - x \in K$$

satisfies all the conditions of definition 1, and hence makes X into a PL space. Thus we can completely define a PL space by giving the space X and the positive cone K. For this reason, we will write (X,K) for a PL space X with positive cone K.

Two examples of PL spaces are (E^2, K_1), where $K_1 = \{(x,y) : x \geq 0 \text{ and } y \geq 0\}$, and (E^2, K_2), where $K_2 = \{(x,y) : x = 0 \text{ and } y \geq 0\}$. That is, in the first space, $(x,y) \leq (u,v)$ means $x \leq u$ and $y \leq v$, whereas, in the second space, $(x,y) \leq (u,v)$ means $x = u$ and $y \leq v$. An important difference between these spaces is that in the ordering defined by K_1, any two points in E^2 have a supremum and an infimum, whereas, with the other ordering, the points (0,1) and (1,1) do not have a common upper or lower bound, and hence no supremum or infimum. PL spaces in which every pair of elements have an infimum and a supremum are called <u>vector lattices.</u> Hence, (E^2,K_1) is a vector lattice while (E^2,K_2) is not. If, however, we have a set S in (E^2,K_2) which is order bounded, then that set has an infimum and a supremum, since all members of S must have the same x-coordinate. A PL space in which every order bounded non-empty subset has a supremum and an infimum is said to be <u>complete</u>. If every countable order-bounded set has a supremum and an infimum, the space is called <u>σ-complete</u>.

Some basic relations which hold in a vector lattice are given in the following theorem. (The proof can be found in [6], p.219).

<u>Theorem 1</u> In a vector lattice X, the following relations hold for any x,y,z in X.

a) $\sup(x,y)+z=\sup(x+z,y+z)$

b) $x+y=\sup(x,y) + \inf(x,y)$

c) $\alpha\sup(x,y) = \sup(\alpha x,\alpha y)$, for $\alpha > 0$

d) $\alpha\sup(x,y) = \inf(\alpha x,\alpha y)$, for $\alpha < 0$

e) $\sup(\inf(x,y),z) = \inf(\sup(x,z),\sup(y,z))$.

Moreover, all of these relations remain valid if sup and inf are everywhere interchanged.

If the space is σ-complete, then these relations can be extended to bounded countable sets. That is, if $\{x_n\}$ is bounded,

a') $\sup(x_n) + z = \sup(x_n+z)$

c') $\alpha\sup(x_n) = \sup(\alpha x_n)$, for $\alpha > 0$

d') $\alpha\sup(x_n) = \inf(\alpha x_n)$, for $\alpha < 0$.

One of the most important properties of a vector lattice is the existence of an absolute value. For any x in X,we define $x^+= \sup(x,0)$, $x^-= \inf(x,0)$, and $|x|= \sup(x,-x)$. x^+ is the <u>positive part</u> of x, x^- is the <u>negative</u>

<u>part</u>, and $|x|$ is called the <u>absolute value</u> of x. Some properties of these quantities are given by

<u>Theorem 2</u>. Let X be a vector lattice. Then, for any x and y in X,

1) $x = x^+ + x^-$
2) $\inf(x^+, -x^-) = 0$
3) $|x| = x^+ - x^-$
4) $|x| \geq 0$ and $|x| = 0$ if and only if $x = 0$
5) $|\alpha x| = |\alpha| \cdot |x|$ for all real α
6) $|x+y| \leq |x| + |y|$

(The proof can also be found in [6], p. 220).

From the definition of x^+ and x^- it is clear that $x^+ \geq 0$, $x^- \leq 0$, so part 1) of this theorem says that any element in a vector lattice can be written as a difference of elements in the positive cone. That is, $X = K - K$. Such a cone is called <u>reproducing</u> or <u>generating</u>.

We will now discuss a natural topology that can be introduced into a PL space. The topology is natural in the sense that it is defined in terms of order concepts and is a vector topology. First, recall from the theory of linear topological spaces the following basic theorem. (See [17], pp.34,35).

> Let X be a topological space , and let B be a local base. Then
> 1) for U and V in B there is a W in B such that $W \subset U \cap V$;
> 2) for U in B there is a member V of B such that $V + V \subset U$;
> 3) for U in B there is a member V of B such that $\alpha V \subset U$ for each scalar α with $|\alpha| \leq 1$;
> 4) for x in X and U in B there is a scalar α such that $x \in \alpha U$;
> 5) for U in B there is a V in B and a circled set W such that $V \subset W \subset U$.
> 6) If X is a Hausdorff space, then $\cap \{U : U \in B\} = \{0\}$

> Conversely, let X be a linear space and B a non-void family of subsets which satisfy 1) through 4), and let T be the family of all sets W such that, for each x in W, there is U in B with $x + U \subset W$. Then T is a vector topology for X, and, B is a local base for this topology. If, further, 6) holds, then T is a Hausdorff topology.

Here we call a set S <u>circled</u> if $\alpha S \subset S$ for all α with $|\alpha| \leq 1$. A set S is said to <u>absorb</u> a set V if there exists an $\alpha_0 > 0$ such that $\alpha V \subset S$ for all $0 \leq \alpha \leq \alpha_0$. The next definition is due to Namioka [20].

<u>Definition 2</u>. Let X be a PL space, and let B be the family of all order bounded subsets of X. Let $\mathcal{U}$ be the family of all subsets of X which are convex, circled, and absorb every member of B. Then the topology for which $\mathcal{U}$ is a local base is called the <u>order bound topology</u>.

To justify this definition, it must be proven that $\mathcal{U}$ does in fact define a unique topology.

<u>Lemma 1</u>. The family $\mathcal{U}$ in the preceeding definition is a local base for a unique locally convex vector topology on X which is finer than any other locally convex vector topology for which order bounded sets are topologically bounded.

<u>Proof</u>. It is sufficient to verify that $\mathcal{U}$ satisfies conditions 1) through 4) of the theorem stated above. But, 1) is true since $U \cap V$ is convex, circled, and given $S \subset X$, if $\alpha_1 S \subset U$, $\alpha_2 S \subset V$ then $\alpha_3 S \subset U \cap V$ where $\alpha_3 = \min(\alpha_1, \alpha_2)$. Condition 2) holds with $V = \frac{1}{2} U$, since U is convex. Finally, 3) and 4) follow from the circled and absorbtion properties. The last part of the theorem is clear.

Very closely related to the order bound topology is the concept of relative uniform convergence, as defined by Birkhoff [6].

Definition 3. A sequence $\{x_n\}$ in a PL space X is said to converge relative uniformly to x* if there exists an element $u \geq 0$ in X and a sequence $\{\alpha_n\}$ of real numbers, such that $\alpha_1 \geq \alpha_2 \geq \ldots \geq 0$, $\lim_{n \to \infty} \alpha_n = 0$, and

$$-\alpha_n u \leq x_n - x^* \leq \alpha_n u.$$

The relation between this convergence and the order bound topology is given by a theorem of Gordon ([9],p.421).

Theorem 3. The order bound topology is the finest locally convex topology such that if $\{x_n\}$ converges relative uniformly to x* then $\{x_n\}$ also converges to x* in the order bound topology.

Proof. Let T be any topology such that order bounded sets are topologically bounded. Let $\{x_n\}$ converge to x* relative uniformly. Then $-\alpha_n u \leq x_n - x^* \leq \alpha_n u$ for some $u \geq 0$ and $\alpha_n \to 0$. Let V be any open set in T and let $B = \{x \in X : -u \leq x \leq u\}$. Then $x_n - x^* \in \alpha_n B$. But B is order bounded, so V absorbs B, i.e., $\alpha B \subset V$ for all small α. Hence, $x_n - x^* \in V$ for large n.

It is not true, in general, that order bound convergence implies relative uniform convergence. Take, for example, the space (E^2, K_2) considered above, with $x_n = (\frac{1}{n}, \frac{1}{n})$. Since the order bound topology is a vector topology, it follows that $x_n \to 0$ in this topology. But, $\{x_n\}$ is not order bounded so clearly there is no u such that $-\alpha_n u \leq x_n \leq \alpha_n u$, where $\alpha_n \to 0$. There is, however, a very wide class of PL spaces in which order bound convergence is equivalent to relative uniform convergence. Included in this class are those spaces which have an order unit;

Definition 4. An element z in a PL space is called an order unit if $z \geq 0$ and, for any x in the space there is a real $\alpha > 0$ such that $-\alpha z \leq x \leq \alpha z$.

In other words, we call z an order unit if [-z,z] is radial at zero.

It is a well-known fact that the Minkowski functional for a convex, circled set which is radial at zero is a semi- norm. This fact allows a very useful characterization of the order bound topology:

Theorem 4. If (Z,K) is a PL space which has an order unit z, then the order bound topology on Z is the semi- norm topology given by the Minkowski functional

$$p(x) = \inf\{\alpha : x \in \alpha[-z,z]\} .$$

Proof. As noted above, p is a semi- norm on Z. Let $S_r = \{x : p(x) < r\}$. Then $S_r \subset r[-z,z] \subset S_{r+\varepsilon}$, $\varepsilon > 0$. Hence, since $[-z,z]$ is absorbing, so is S_r. Also, S_r is convex and circled so $S_r \in \mathcal{U}$, where $\mathcal{U}$ is the local base for the order bound topology. Furthermore, each S_r is order bounded, and every $U \in \mathcal{U}$ absorbs order bounded sets, so every $U \in \mathcal{U}$ contains some S_r. Therefore, the topology determined by $\mathcal{U}$ is the same as the topology determined by $\{S_r\}$.

In order for $p(x)$ to be a norm, we need an additional hypothesis.

Definition 5. A PL space is called almost Archimedian if $-\alpha x \leq y \leq \alpha x$ for some $x \geq 0$, and all $\alpha > 0$, implies $y = 0$.

Corollary. If a PL space has an order unit, then the order bound topology is a norm topology if and only if the space is almost Archimedian.

Proof. If the space is almost Archimedian, then $p(x) = 0$ implies $-\alpha z \leq x \leq \alpha z$ for all $\alpha > 0$, hence $x = 0$. Conversely, if p is a norm, then the topology is Hausdorff. But $-\alpha x \leq y \leq \alpha x$ for all $\alpha > 0$ implies that $y \in U$ for every U in the local base for the order bound topology. Since this topology is Hausdorff, $y = 0$.

To illustrate these results, consider E^2 with the positive cone $K = \{(x,y) : x > 0 \text{ or } (x = 0 \text{ and } y \geq 0)\}$ Then the point (1,1) is an order unit, and so the order bound topology is given by the semi-norm

$$p(x,y) = \inf\{\alpha : (x,y) \in \alpha[(-1,-1),(1,1)]\}$$
$$= \inf\{\alpha : (-\alpha,-\alpha) \leq (x,y) \leq (\alpha,\alpha)\}$$

hence $p(x,y) = |x|$. This space is not almost Archimedian because, if $a = (1,1)$ and $b = (0,1)$ then $-\alpha a \leq b \leq \alpha a$ for all $\alpha > 0$.

Using the semi-norm defined in theorem 4, we can now prove the equivalence of order bound convergence and relative uniform convergence.

Theorem 5. If a PL space has an order unit z, then relative uniform convergence is equivalent to order bound convergence.

Proof. Theorem 3 shows that if $x_n \to x$ relative uniformly, then $x_n \to x$ in the order bound topology. Now, if $x_n \to x$ in the order bound topology, then $p(x_n - x) \to 0$ as $n \to \infty$. But, $p(x_n - x) = \inf\{\alpha : (x_n - x) \in \alpha[-z,z]\}$, hence there exists a sequence $\alpha_n \to 0$ so that $x_n - x \in \alpha_n[-z,z]$. i.e., $-\alpha_n z \leq x_n - x \leq \alpha_n z$.

As an indication of the close connection between the ordering in the space and the order bound topology, we prove a theorem in which an order concept (σ-completeness) implies a topological concept (topological completeness).

Theorem 6. If a PL space with order unit is σ-complete, then the order bound topology is topologically complete.

Proof. By theorem 4, the order bound topology is a semi-norm topology, hence completeness is equivalent to sequential completeness. Let $\{x_n\}$ be a Cauchy sequence; i.e., $x_n - x_m \to 0$ as $n,m \to \infty$. By theorem 5, this implies

$$-\alpha_{n,m} z \leq x_n - x_m \leq \alpha_{n,m} z$$

where $\alpha_{n,m} \to 0$ as $n,m \to \infty$. But then

$$x_1 - \alpha_{n,1} z \leq x_n \leq \alpha_{n,1} z + x_1$$

so $\{x_n\}$ is order bounded, and

$$\sup_{n \geq m} \{-\alpha_{n,m} z\} \leq \sup_{n \geq m} (x_n - x_m) \leq \sup_{n \geq m} \{\alpha_{n,m} z\}.$$

By theorem 1, this implies

$$\sup_{n \geq m} \{-\alpha_{n,m}\} z \leq \sup_{n \geq m} \{x_n\} - x_m \leq \sup_{n \geq m} \{\alpha_{n,m}\} z$$

so

$$0 \leq \sup_{n \geq m} \{x_n\} - x_m \leq \sup_{n \geq m} \{\alpha_{n,m}\} z.$$

Let $u_m = \sup_{n \geq m} \{x_n\}$. Then $u_1 \geq u_2 \geq \ldots \geq \inf_n \{x_n\}$, so if $x^* = \inf_m \{u_m\}$, then $0 \leq u_m - x_m \leq \sup_{n \geq m} \{\alpha_{n,m}\} z$ and

$$x^* - x_m \leq u_m - x_m \leq \alpha_m z$$

where $\alpha_m = \sup_{n \geq m} \{\alpha_{n,m}\} \to 0$ as $m \to \infty$. Similarly, taking infima, we can show

$$x_m - x_* \leq \alpha_m z$$

where $x_* = \sup_m \inf_{n \geq m} \{x_n\}$. But then

$$x^* - x_* = (x^* - x_m) + (x_m - x_*) \leq 2\alpha_m z$$

where $\alpha_m \to 0$. Thus $x^* = x_*$ and

$$-\alpha_m z \leq x^* - x_m \leq \alpha_m z$$

hence $x_m \to x^*$ relative uniformly. Again applying theorem 5, the proof is complete.

As a final comment on PL spaces which have an order unit, we observe that these spaces are exactly those used by Schroeder [23] in his work on operators with positive inverses. In this paper, Schroeder defines another type of ordering by setting $x \gg 0$ if, for every z in the space, there exists an $\alpha > 0$ such that $-\alpha x \leq z \leq \alpha x$. Hence "$z \gg 0$" is equivalent to "$z$ is an order unit." This ordering is not a partial ordering, in the sense of definition 1, because it is not true that $0 \gg 0$. Some basic properties of this ordering are;

<u>Lemma 2</u>. If (Z,K) is a PL space with an order unit,

then

1) $x \gg 0$ and $y \gg 0$ imply $\alpha x + \beta y \gg 0$ for $\alpha \geq 0$, $\beta \geq 0$, and $\alpha + \beta > 0$.

2) $x \geq 0$ and $y \gg x$ imply $y \gg 0$

3) $x \gg 0$ and $y \geq x$ imply $y \gg 0$.

Proof.

1) Let $z \in Z$. Then for some $\eta_1, \eta_2 \geq 0$, $-\eta_1 x \leq z \leq \eta_1 x$ and $-\eta_2 y \leq z \leq \eta_2 y$. If either α or β are zero, the result is clear. Assume both are non- zero and let $\eta_3 = \max(\eta_1/2\alpha, \eta_2/2\beta)$. Then

$$-\eta_3(\alpha x + \beta y) \leq -\frac{\eta_1}{2}x - \frac{\eta_2}{2}y \leq z \leq \frac{\eta_1}{2}x + \frac{\eta_2}{2}y \leq \eta_3(\alpha x + \beta y).$$

2) If $z \in Z$, then for some $\eta \geq 0$, $-\eta(y-x) \leq z \leq \eta(y-x)$. But $x \geq 0$ so $-\eta y \leq -\eta y + \eta x \leq z \leq \eta(y-x) \leq \eta y$, hence $y \gg 0$.

3) Let $z = y-x$. Then $z \geq 0$ and $y-z = x \gg 0$, so $y \gg z$. Now apply part 2) to get $y \gg 0$.

Partially Ordered Topological Linear Spaces

In the previous section, it was shown that we can start with a PL space and define a topology on it which will give a topological linear space. In most applications however, a linear space is given which already has one or more well- known topologies, and it is not always clear how these topologies are related to the order bound topology. An alternative procedure for developing this theory, which avoids this problem, is to start with a topological linear space in which a partial ordering is defined, and then try to prove the necessary relations between the topology and the ordering. This is the method used by Krasnoselski [18] for Banach spaces and by Schaefer [21] for locally convex spaces. It will be shown, however, that even the most basic relations cannot be proven without some additional assumptions. One such relation which is always very helpful is that the limit of a sequence of positive elements should also be positive. For this reason, we introduce the following definition.

Definition 6. A partially ordered topological linear space (PTL space) is a PL space with a locally convex vector topology such that the positive cone is a closed set.

We will use (Z,K,T) to denote a PTL space Z with closed cone K and locally convex vector topology T. The fact that the cone is closed also implies that intervals [a,b] are closed sets.

A PL space with its order bound topology is not necessarily a PTL space. Take, for example, the space E^2 with $K = \{(x,y) : x > 0 \text{ or}, x = 0 \text{ and } y \geq 0\}$. In the previous section, we showed that the order bound topology for this space is determined by the semi- norm $p(x,y) = |x|$ hence K is not closed.

We will now list some common PTL spaces which will provide us, throughout the remainder of this chapter, with examples and counter- examples.

1) C[0,1], real valued functions, continuous on [0,1], T the topology given by the norm $\|f\| = \max|f(t)|$, $K = \{f : f(t) \geq 0 \text{ for } t \in [0,1]\}$.

2) $C^n[0,1]$, real functions with n continuous derivatives on [0,1],

T_0 the norm topology with $\|f\| = \max|f(t)|$,

T_n the norm topology with $\|f\| = \sum_{k=0}^{n} \max|f^{(k)}(t)|$

$K = \{f : f(t) \geq 0 \text{ for } t \in [0,1]\}$

3) BV[0,1], functions of bounded variation on [0,1], T given by the norm $\|f\| = |f(0)| + V(f)$, where V(f) is the total variation of $|f|$,

$K = \{f : f(t) \geq 0 \text{ for } t \in [0,1]\}$

4) $L^p[0,1]$, $0 < p < \infty$

T given by the norm $\|f\| = \left\{ \int_0^1 |f|^p dt \right\}^{1/p}$

$K = \{f : f(t) \geqslant 0 \text{ a.e.}\}$

5) $L^\infty[0,1]$, bounded measurable functions on $[0,1]$

T given by the norm $\|f\| = \text{ess sup } |f(t)|$

$K = \{f : f(t) \geqslant 0 \text{ a.e.}\}$

6) E^n, real n-dimensional space

T given by $\|x\| = \left\{ \sum_{i=1}^{n} |x_i|^2 \right\}^{1/2}$

$K_1 = \{(x_1, x_2, \ldots, x_n) : x_i \geqslant 0,\ i = 1,2,\ldots,n\}$

$K_2 = \{(x_1, x_2, \ldots, x_n) : x_i \geqslant 0,\ i = 1,2,\ldots,n-1,\ x_n = 0\}$

7) $H(S^2)$, real valued functions which are harmonic in the open unit sphere $S^2 = \{(x,y) : x^2 + y^2 < 1\}$, and bounded on the closed unit sphere in E^2.

T given by the norm $\|f\| = \lim_{r \to 1} \max_{x^2+y^2 \leqslant r} |f(x,y)|$

$K = \{f : f(x,y) \geqslant 0 \text{ for } x^2 + y^2 \leqslant 1\}$

It is simple to check that all of the above examples are PTL spaces. Using them, we can easily show that the closedness of the positive cone is not, in general, a strong enough connection between the ordering and the topology. Consider, for example, the following properties of sequences of real numbers:

A) $x_1 \leqslant x_2 \leqslant x_3 \leqslant \ldots \leqslant x^*$ and $\sup\{x_n\} = x^*$ implies $\lim x_n = x^*$.

B) $\lim x_n = 0$ implies that there exists $\{y_n\}$ with $y_1 \geqslant y_2 \geqslant \ldots \geqslant 0$, $\inf\{y_n\} = 0$, and $-y_n \leqslant x_n \leqslant y_n$.

C) $0 \leqslant x_n \leqslant y_n$ and $\lim y_n = 0$ implies $\lim x_n = 0$.

Unfortunately, these statements are not true for all PTL spaces:

a) In $(C[0,1],K,T)$ let $x_n(t) = -t^n$. Then $x_1 \leqslant x_2 \leqslant \ldots \leqslant 0$, and $\sup\{x_n\} = 0$, but $\|x_n\| = 1$, all n, so $\lim x_n$ does not exist. Hence A) does not hold.

b) In $(L^1[0,1],T,K)$ let $x_n(t) = n$ for $\frac{1}{n+1} \leqslant t \leqslant \frac{1}{n}$ and zero elsewhere. Then $\lim \|x_n\| = 0$ but clearly, property B) does not hold.

c) In $(C^1[0,1],T_1,K)$ let $x_n(t) = \frac{1}{n}t^n$, $y_n(t) = \frac{1}{n}$. Then $0 \leq x_n \leq y_n$, and $\lim y_n = 0$, but in the T_1 norm $\|x_n\| = \max|\frac{1}{n}t^n| + \max|t^{n-1}| = \frac{1}{n} + 1 > 1$, hence x_n does not converge to zero.

The remainder of this section will be devoted to a brief discussion of certain types of PTL spaces in which some of the above statements are true.

Definition 7. A PTL space is called regular if every order bounded increasing sequence has a limit.

Examples of regular PTL spaces are (E^n,T,K_1), (E^n,T,K_2), and $(L^p[0,1],T,K)$, whereas $(C^n[0,1],T_0,K)$ and $(C^n[0,1],T_n,K)$ are not regular, as was shown in example a) above.

If $\{z_n\}$ is a monotone increasing sequence, and $\lim z_n = z^*$ exists, then for any k_0, $n \geq k_0$ implies $z_n \geq z_{k_0}$. Hence $z^* = \lim z_n \geq z_{k_0}$, i.e., z^* is an upper bound on $\{z_n\}$. Moreover, if w is any other upper bound, then $z_n \leq w$ and hence $z^* = \lim z_n \leq w$, i.e., $z^* = \sup\{z_n\}$. We have shown that in any PTL space, the closedness of the positive cone guarantees that, if a monotone increasing sequence has a limit, then it also has a supremum. In a regular space, the converse of this is true, i.e., if a monotone increasing sequence has a supremum, then it also has a limit. It is important to note that the definition of regularity involves both an order concept (monotone boundedness) and a topological concept (limit).

Definition 8. If the positive cone has an interior point, then the PTL space is called solid.

Examples of solid spaces are $(C[0,1],T,K)$ and (E^n,K_1,T). The first has $f(t)\equiv 1$ as an interior point of K, while $(1,1,\ldots,1)$ is an interior point of K_1. The spaces

$(L^p[0,1],T,K)$, $0<p<\infty$, are not solid because, given any $f\in K$, and any $\varepsilon>0$, there is a $g\notin K$ such that $\|f-g\|\leq\varepsilon$. An equivalent characterization of a solid space is given by

<u>Lemma 3</u>. A space is solid if and only if there exists an open set Θ containing the origin, with $\Theta\subset[-a,a]$, for some $a\geqslant 0$.

<u>Proof</u>. If $0\in\Theta\subset[-a,a]$, then $a\in\Theta+a\subset[0,2a]\subset K$, hence a is an interior point of K. Conversely, if $a\in\Theta\subset K$, then $0\in[\Theta-a\cap(-\Theta+a)]\subset[-a,a]$.

The theory concerning order units which was developed in the preceeding section can be applied to solid spaces because of the following result.

<u>Lemma 4</u>. If (Z,T,K) is a solid PTL space, then z_0 is an order unit if and only if z_0 is an interior point of K.

<u>Proof</u>. If z_0 is in the interior of K, then $z_0\in\Theta\subset K$, where Θ is an open set. Let z be arbitrary in Z. Then $\eta z\to 0$ as $\eta\to 0$ and since $\Theta-z_0$, $-\Theta+z_0$ are both neighborhoods of 0, for some $\eta_1>0$,

$$\eta z\in(\Theta-z_0)\cap(-\Theta+z_0)\subset[-z_0,z_0]$$

for all $\eta\leq\eta_1$. That is, $\eta z\in[-z_0,z_0]$ and hence z_0 is an order unit. Conversely, if z_0 is an order unit, and z_1 is an interior point of K, then for some η, $\eta z_1\leq z_0$. But $z_1\in\Theta\subset K$ so $\eta z_1\in\eta\Theta\subset K$ and

$$z_0=\eta z_1+(z_0-\eta z_1)\subset\eta\Theta+(z_0-\eta z_1)\subset K,$$

where $\eta\Theta+(z_0-\eta z_1)$ is an open set. Thus z_0 is an interior point of K.

A consequence of this lemma is that if a space is solid, then it has an order unit. The converse is not true, in general, as is illustrated by the following example. Let $Z=C[0,1]$, K the usual ordering, and T the topology given by the $L^1[0,1]$ norm. Then $f(t)\equiv 1$ is an order unit, but there is no interior point in K, as was pointed out earlier, hence (Z,T,K) is not solid.

It is now possible to relate the topology in a solid PTL space to the order bound topology. First, however, we need some basic facts about these spaces.

Lemma 5. If (Z,T,K) is a solid PTL space, then

1) K is reproducing;

2) every topologically bounded set is order bounded;

3) every compact set is order bounded.

Proof.

1) Let z_0 be an interior point of K and let z be arbitrary. By lemma 4, $[-z_0,z_0]$ is radial at 0, so for some $\eta > 0$, $\eta z_0 \geqslant z$. Hence $z = \eta z_0 - (\eta z_0 - z)$ where ηz_0 and $(\eta z_0 - z)$ are in K.

2) If S is topologically bounded, and z_0 is an interior point of K, then $\eta S \subset [-z_0,z_0]$ for some $\eta > 0$, because $[-z_0,z_0]$ contains an open set. (See the proof of lemma 3).

3) Every compact set is topologically bounded, hence is order bounded because of part 2).

Theorem 7. If (Z,T,K) is a solid PTL space, then T is finer than the order bound topology.

Proof. Let $\mathcal{U}$ be the local base for the order bound topology, and let $\mathcal{W}$ be a local base for T. By lemma 3, there must be order bounded sets in $\mathcal{W}$, and every $U \in \mathcal{U}$ absorbs such sets, hence every $U \in \mathcal{U}$ contains some $W \in \mathcal{W}$. Thus $\mathcal{W}$ is finer than $\mathcal{U}$.

An important consequence of this theorem is the following result.

Corollary. In a solid PTL space, all sequences have property B). That is, if $\lim x_n = 0$, then there exists $\{y_n\}$ with $y_1 \geqslant y_2 \geqslant \cdots \geqslant 0$, $\inf \{y_n\} = 0$, and $-y_n \leqslant x_n \leqslant y_n$.

Proof. If $\lim x_n = 0$, then $x_n \to 0$ in the order bound topology, hence by theorem 5, $x_n \to 0$ relative uniformly. Thus, there exists $\eta_1 \geq \eta_2 \geq \dots \geq 0$, and $u \geq 0$ such that $\inf \{\eta_n\} = 0$ and $-\eta_n u \leq x_n \leq \eta_n u$.

The last type of PTL space we will consider is defined by

<u>Definition 9</u>. A PTL space is <u>normal</u> if, given any local base $\mathcal{U}$ for the topology, there exists an $\eta > 0$ so that if $0 \leq z \in U \in \mathcal{U}$, then $[0,z] \subset \eta U$.

The spaces (E^n,T,K), $(L^p[0,1],T,K)$ and $(C[0,1],T,K)$ are normal, while $(C^n[0,1],T_n,K)$ and $(BV[0,1],T,K)$ are not. If T is a locally convex topology, then there is a family $\{p_\alpha\}$ of semi- norms such that the family of sets of the form $\{z : p_\alpha(z) < r\}$, for r real, is a local base for the topology T. Hence, the definition says that there exists an $\eta > 0$ so that if $0 \leq x \leq y$ and $p_\alpha(y) < r$ then $p_\alpha(x) < \eta r$. From this observation, we can prove several statements which are equivalent to normality.

<u>Theorem 8</u>. Let (Z,T,K) be a PTL space. Then each of the following statements is necessary and sufficient for the space to be normal. For any continuous semi- norm p,

1) there exists an $\eta > 0$, which is independent of p, such that $p(x) \leq \eta p(y)$ for $0 \leq x \leq y$.

2) $(S_p + K) \cap (S_p - K) \subset \eta S_p$ where η is independent of p, and $S_p = \{z : p(z) < 1\}$.

3) there is a continuous semi- norm q, equivalent to p, such that $q(x) \leq q(y)$ for $0 \leq x \leq y$.

<u>Proof</u>. We first observe that the collection of all continuous semi- norms defines a local base, hence the family $\{p_\alpha\}$, in the remark following the definition, can be this collection.

1) If Z is normal, and $0 \leq x \leq y$ then $p(y) < \frac{1}{\eta} p(x)$ would imply $p(x) < \eta p(y) < p(x)$ which is impossible, so

$p(y) \geqslant 1/\eta\, p(x)$. Conversely, if $0 \leq x \leq y$ implies $p(x) \leq \eta p(y)$, where η is independent of p, then clearly $p(y) < r$ implies $p(x) < \eta r$, hence normality.

2) If (Z,T,K) is normal, and $z \in (S_p + K) \cap (S_p - K)$, then $z = u+x = v-y$ where $p(u) < 1$, $p(v) < 1$, and $x,y \geqslant 0$. But then $0 \leq x \leq x+y$ so by 1),

$$p(x) \leq \eta p(x+y) = \eta p(v-u) \leq \eta(p(v)+p(u)) \leq 2\eta,$$

hence $p(z) = p(u+x) \leq p(u) + p(x) \leq 1+2\eta$, thus $z \in (1+2\eta)S_p$. Conversely, if $U_p = (S_p + K) \cap (S_p - K) \subset \eta S_p$ then $p(z) < \eta$ for all $z \in U_p$. But then, if $0 \leq x \leq y$ and $0 < p(y) < r$, letting $y^* = \frac{1}{r} y$, $x^* = \frac{1}{r} x$, we have $0 \leq x^* \leq y^*$, $p(y^*) < 1$, and $x^* = y^* + (x^* - -y^*)$ so $x^* \in U_p$ and hence $p(x^*) < \eta$. i.e., $p(x) < \beta r$ where β is independent of p.

3) If every p has an equivalent monotone q, then part 1) is true with $\eta = 1$. Conversely, if Z is normal, then $S_p \subset U_p = (S_p + K) \cap (S_p - K) \subset \eta S_p$. But S_p is radial at 0, and hence U_p is a convex, circled set, radial at 0. Thus the Minkowski functional $q(z) = \inf\left\{\eta : x \in \eta U_p\right\}$ is a continuous semi- norm. Furthermore,

a) $p(x) = 0$ implies $q(x) = 0$ since for such x, $x \in \eta U_p$ for all $\eta > 0$. Also, $p(x) > 0$ implies $q(x) = p(x) q\left(\frac{x}{p(x)}\right) < p(x)$ since $\frac{x}{p(x)} \in U_p$. Hence, $q(x) \leq p(x)$.

b) By part 2), $p(U_p) \leq k$, but $q(x) \geqslant \eta$ where $1/\eta\, x \in U_p$, hence $p(1/\eta\, x) \leq k$. That is $q(x) \geqslant \eta \geqslant \frac{1}{k} p(x)$.

Thus for any $x \in Z$, $\frac{1}{k} p(x) \leq q(x) \leq p(x)$ so q is equivalent to p. Finally, if $0 \leq x \leq y$, and $p(x) \leq 1$ then $1/\eta\, y \in U_p$ implies $1/\eta\, y = u+a = v-b$ where $u,v \in U_p$ and $a,b \in K$. But $x=y-(y-x)$ so $1/\eta\, x = v-b-1/\eta(y-x) = v-(b+1/\eta(y-x))$ where $b+1/\eta(y-x) \in K$. Hence, $1/\eta\, x \in U_p$ and so $\left\{\eta : y \in \eta S\right\} \subset \left\{\eta : x \in \eta S\right\}$ and therefore $q(y) \leq q(x)$.

The last part of this theorem has an interesting consequence. Given any family $\left\{p_\alpha\right\}$ of semi- norms, the theorem says that normality is equivalent to the existence of a family $\{q_\alpha\}$ of semi- norms with q_α equivalent to p_α and q_α monotone on K. Hence we have

Corollary. A PTL space is normal if and only if there exists a local base $\mathcal{U}$ for the topology, such that if $0 \leq x \in U \in \mathcal{U}$, and $0 \leq y \leq x$, then $y \in U$.

This is, in fact, the statement used by Schaefer to define a normal PTL space. (See [21], p.121). If the topology is a norm topology, then part 1) of the theorem says that there exists an η so that $0 \leq x \leq y$ implies $\|x\| \leq \eta \|y\|$. This is equivalent to the condition used by Kelley and Namioka ([17], p.227), and Krasnoselski ([18], pp. 20,24) to define a normal normed space. This theorem shows that these two definitions are the same.

Some further properties of normal spaces which will be needed in the next chapter are:

Theorem 9. In a normal PTL space,

1) every order bounded set is topologically bounded;

2) the topology is coarser than the order bound topology;

3) property C) holds. That is, if $0 \leq x_n \leq y_n$ and $\lim y_n = 0$ then $\lim x_n = 0$.

Proof.

1) If $S \subset [a,b]$ then $S' = S - a \subset [0,b-a]$. If $\{p_\alpha\}$ is a family of semi- norms which defines the topology, then $p_\alpha(S') \leq \eta p_\alpha(b-a)$, all α. Hence $p_\alpha(S) = p_\alpha(S'+a) \leq p_\alpha(S') + p_\alpha(a) \leq \eta p_\alpha(b-a) + p_\alpha(a)$, so S is topologically bounded.

2) The order bound topology is, by definition, finer than any topology for which 1) is true.

3) If $0 \leq x_n \leq y_n$ and $\lim y_n = 0$, then for any p_α, $p_\alpha(x_n) \leq \eta p_\alpha(y_n) \to 0$ as $n \to \infty$. Hence, $p_\alpha(x_n) \to 0$ for all α, and therefore, $\lim x_n = 0$.

Combining part 2) of this theorem with theorem 7, we see that the topology in a solid normal PTL space is equivalent to the order bound topology. Actually, we can prove a little more about such spaces, but first we need some more definitions.

Definition 10.

1) A PTL space in which the topology is given by a norm $\|\cdot\|$ which satisfies $\|x\| \leq k\|y\|$ for $0 \leq x \leq y$, where k is independent of x,y, is called a partially ordered normed linear space, (PNL space).

2) A PNL space which is also a Banach space is called a partially ordered Banach space, (PB space).

3) A PB space which is also a lattice in which the absolute value and the norm are related by $\| \, |x| \, \| = \|x\|$ for all x, is called a Banach lattice, (BL space).

Notice that PNL, PB, and BL spaces are all normal, since condition 1) of theorem 8 is satisfied. Furthermore, by this same theorem, any normal PTL space, where T is a norm topology, is a PNL space. That is;

Theorem 10. If T is a norm topology, then (Z,T,K) is a PNL space if and only if it is a normal PTL space.

Now we can summarize some of the characteristics of a solid and normal space.

Theorem 11. Let (Z,T,K) be a PTL space which is both solid and normal. Then

1) T is the order bound topology;

2) (Z,T,K) is a PNL space;

3) if (Z,K) is σ-complete, then (Z,T,K) is a PB space.

Proof. Part 1) was proven above, and 3) follows from theorem 6. To prove part 2) it suffices to show that the order bound topology is a norm topology. But by the corollary to theorem 4, this is true if (Z,K) is almost Archimedian. Suppose that $-\alpha x \leq y \leq \alpha x$ for some $x \geq 0$ and all $\alpha > 0$. Then since K is closed, and T is a vector topology, letting $\alpha \to 0$ gives $y \leq 0$ and $y \geq 0$, hence we must have $y = 0$. That is, the space is almost Archimedian.

We conclude this section with a summary of the spaces we have been using as examples.

	Lattice	σ-comp.	Regular	Solid	Normal	PB Space
(E^n, K_1, T)	yes	yes	yes	yes	yes	yes
(E^n, K_2, T)	no	yes	yes	no	yes	yes
(L^p, K, T)	yes	yes	yes	no	yes	yes
$(L\ , K, T)$	yes	yes	no	yes	yes	yes
(C, K, T)	yes	no	no	yes	yes	yes
$(C, K, T_1)^{1)}$	yes	no	no	no	yes	no$^{3)}$
(C^n, K, T_n)	no	no	no	yes	no	no$^{2)}$
(BV, K, T)	yes	yes	yes	no	no	no$^{2)}$
(H, K, T)	yes	yes	yes$^{4)}$	yes	yes	yes

1) T_1 is given by the L^1 norm; $\|f\| = \int_0^1 |f(t)|\,dt$.

2) The norm is these spaces is not monotone.

3) This space is not complete.

4) The regularity of this space follows from Harnack's theorem.

Table 1. PTL Spaces

Spaces with Partially Ordered Norm

Let X be a real linear space and suppose there is a mapping $\|\cdot\|$ from X into a PTL space Z, which satisfies:

1) $\|x\| \geqslant 0$ for all x in X;

2) $\|x\| = 0$ only if $x = 0$;

3) $\|\alpha x\| = |\alpha| \, \|x\|$ for all real α;

4) $\|x+y\| \leq \|x\| + \|y\|$.

Then X is said to be <u>normed by Z</u>. This concept has been used extensively by Kantorovich ([14], also see [4]), who has proven a convergence theorem for Newton's method applied to operators in such spaces. In order to apply the results of the next chapter to these spaces, we will now define a topology on X which is induced by the partially ordered norm.

Suppose the linear space X is normed by a <u>normal</u> PTL space Z. Let $\mathcal{U}$ be a local neighborhood base for the topology in Z, and, for each $U \in \mathcal{U}$ let

$$Q(U) = \{x \in X : \|x\| \in U\}$$

Then, the collection $\mathcal{W} = \{Q(U) : U \in \mathcal{U}\}$ is a local base for a vector topology in X. This is true because we can assume that $\mathcal{U}$ has the property: $0 \leqslant x \leqslant y \in U \in \mathcal{U}$ implies $x \in U$, for all $U \in \mathcal{U}$ and $y \in U$, in which case,

1) for any $Q(U), Q(V)$, there is a $W \subset U \cap V$ and hence $Q(W) \subset Q(U) \cap Q(V)$;

2) for any $Q(U)$ there is $V \in \mathcal{U}$ with $V+V \subset U$, hence $Q(V)+Q(V) \subset Q(U)$ because if $x_1, x_2 \in Q(V)$, then $\|x_1\| \in V$, $\|x_2\| \in V$ so $0 \leq \|x_1+x_2\| \leq \|x_1\| + \|x_2\|$ and $\|x_1\| + \|x_2\| \in U$, hence by normality, $\|x_1+x_2\| \in U$. i.e., $x_1+x_2 \in Q(U)$;

3) for any $Q(U)$, there is a $V \in \mathcal{U}$ with $\alpha V \subset U$ for all α with $|\alpha| \leqslant 1$. Hence $\alpha Q(V) \subset Q(U)$.

4) for any $x \in X$, and any $Q(U)$, $\|x\| \in \alpha U$ for some α, hence $x \in \alpha Q(U)$.

Thus, by the theorem on local bases quoted earlier, $\mathcal{W}$ is a local base for a locally convex vector topology. This topology will be called the norm topology induced by Z. If Z is Hausdorff, then so is this topology, because, if $x \in \bigcap Q(U)$ then $\|x\| \in \bigcap U = \{0\}$ and, by condition 2) of the definition, this implies that $x = 0$.

If Z is a PNL space, then the topology induced by Z is also a norm topology. To show this, we use the fact that Z has a norm p which is monotone. Hence, the function q defined on X by $q(x) = p(\|x\|)$ is a norm because:

a) $q(x) \geq 0$ for all x in X;

b) $q(x) = 0$ only if $x = 0$;

c) $q(\alpha x) = p(\|\alpha x\|) = p(|\alpha| \|x\|) = |\alpha| q(x)$;

d) $0 \leq \|x+y\| \leq \|x\| + \|y\|$, hence $p(\|x+y\|) \leq p(\|x\|+\|y\|) \leq p(\|x\|) + p(\|y\|)$, i.e., $q(x+y) \leq q(x) + q(y)$.

Furthermore, a local base for the topology induced by q consists of sets of the form $\{x : q(x) \leq r\} = \{x : p(\|x\|) \leq r\}$ But these sets also form a local base for the topology induced by Z, hence these two topologies on X are identical.

Operators on PTL Spaces

The purpose of this section is to define convex operators and prove some properties of them which will be needed in the next chapter. These proofs will use an integral theorem of the form $f(1)-f(0)=\int_0^1 f'(t)dt$, and so we will first have to define the derivative and integral of an operator. These definitions do not make use of the ordering in the space and hence are given for topological linear spaces. Much of this material can be found in Vainberg [27] and Kantorovich and Akilov [16].

Let X and Y be locally convex topological linear spaces, and let D be a convex subset of X. If x_0 is an interior point of D, ($x_0 \in D^0$), then for any $h \in X$, $x_0+th \in D^0$ for small enough t, and hence if $F:D \subset X \to Y$ then $F[x_0+th]$ is defined for small t. (We use $F:D \subset X \to Y$ to denote an operator F defined on a subset D of X, with range in Y)

<u>Definition 11</u>. If $F:D \subset X \to Y$ and $x_0 \in D^0$ and if the limit $VF[x_0,h]=\lim_{t \to 0} \frac{1}{t} \{F[x_0+th]-F[x_0]\}$ exists and is unique for all h in X, then F is said to have a <u>Gateaux differential</u> at x_0, and $VF[x_0,h]$ is called the <u>Gateaux differential in the direction h</u>.

From this definition, it is clear that $VF[x,\alpha h] = \alpha VF[x,h]$ for any real α. If $VF[x,h]$ is also additive in h, then the operator $F'[x]$ defined by

$$F'[x]h = VF[x,h]$$

is a linear operator. (It is not always true that $VF[x,h]$ is linear in h.) If $F'[x]$ is a <u>continuous</u> linear operator then it is called the <u>Gateaux derivative of F at x</u>. If $S \subset D^0$ and F has a Gateaux derivative (G-derivative) at every point in S, then F is <u>G-differentiable in S</u>. F is called <u>continuously G-differentiable in S</u> if the mapping $F':S \to E(X,Y)$ is continuous, where $E(X,Y)$ is the space of continuous linear operators from X into Y, with the weak operator topology.

The G-derivative at a point is a generalization of the directional derivative. For this reason, it is to be expected that certain important results cannot be proved using this type of derivative. For example, an operator can have a G-derivative at a point without being continuous there, and, in general, the composite function theorem

$$(FH)'[x] = F'(H[x])H'[x]$$

does not hold for G-differentiable operators. By adding another condition to the definition of the G-derivative, we get the Frechet derivative, which has all the properties needed of a derivative. This added condition can be given for an operator defined on any topological space [19], but since it will be used here only in normed spaces, we define it as follows;

Definition 12. Let $F:D\subset X\rightarrow Y$ be G-differentiable at $x_0 \in D^0$, where X and Y are normed linear spaces. If the operator $F'[x_0]$ satisfies

$$\lim_{\|h\|\to 0} \frac{1}{\|h\|} \|F[x_0+h] - F[x_0] - F'[x_0]h\| = 0$$

then F is called Frechet differentiable (F-differentiable) at x^0. The operator $F'[x_0]$ is called the Frechet derivative.

If F is either G or F-differentiable in a convex set S, then F' is an operator from S into $E(X,Y)$. If this operator is also differentiable, in the same sense, then F is said to be twice G or F- differentiable. The derivative of F' at x_0 is denoted by $F''[x_0]$. Note that $F''[x_0]:X\rightarrow E(X,Y)$, hence $F''[x_0](x)$ is a linear operator from X into Y. That is, $F''[x_0]$ can be interpreted as a bi-linear operator from $X \times X$ into Y. Higher derivatives are defined similarly.

It is possible to define the integral of an operator $F:X\rightarrow Y$ where X and Y are locally convex spaces. For our purposes, however, it will suffice to consider only the special case where X is the unit interval [0,1]. Let $\{t_i\}$ $i=0,1,\ldots,n$, be a subdivision of [0,1], i.e., $0=t_0 \leq t_1 \leq t_2 \leq \ldots \leq t_n = 1$, and let ξ_k be any point in $[t_k,t_{k+1}]$. If the Riemann sums $\sum_{k=0}^{n-1} F[\xi_k](t_{k+1}-t_k)$ approach a unique limit, as $\max(|t_{k+1}-t_k|)$ goes to zero, independently of the choice of the subdivisions and the points ξ_k, then this limit is called the <u>Riemann integral of F</u>, and is denoted by $\int_0^1 F(t)\,dt$. The reason we define the integral of an operator is to allow us to prove the formula $F(1) - F(0) = \int_0^1 F'(t)\,dt$. If $F:[0,1] \to Y$, where Y is a Banach space, and F is continuously G-differentiable, then this formula is indeed true. (See [16], p. 666). However, using PTL spaces, we can give other conditions on F and Y which will allow us to prove a result more useful to us than the Banach space theorem. In order to state these conditions, we first need some properties of operators defined on PTL spaces.

<u>Definition 13</u>. If $F:Z \to W$, where Z and W are PTL spaces, then F is called <u>positive</u> <u>($F \geqslant 0$)</u> if $F[z] \geqslant 0$ for all $z \geqslant 0$.

If F and G are two operators from Z into W, then we write $F \geqslant G$ if $F - G$ is a positive operator. Thus, the ordering in the spaces induces an ordering between the operators.

<u>Definition 14</u>. An operator F is called <u>monotone</u> if $z_1 \leqslant z_2$ implies $F[z_1] \leqslant F[z_2]$.

If F is linear, then $F \geqslant 0$ is equivalent to the monotonicity of F.

<u>Definition 15</u>. An operator F is called <u>inverse positive</u> if $F[z] \geqslant 0$ implies $z \geqslant 0$.

If F has an inverse, then F is inverse positive if and only if the inverse of F is positive. A linear operator which is inverse positive is also one-to-one because, if $F[z^*]=0$ then, by definition, $z^* \geqslant 0$ and also $z^* \leqslant 0$, hence $z^*=0$.

<u>Definition 16</u>. An operator is <u>order bounded</u> if it maps order bounded sets into order bounded sets.

Every monotone operator is clearly order bounded, but not conversely.

If x and y are elements in a linear space X, then the <u>segment (x, x+y)</u> is the set of points of the form x+ty, for $0 \le t \le 1$.

<u>Definition 17</u>. An operator $F:Z \to W$ is <u>uniformly differentiable on the segment (z,z+h)</u> if F is G- differentiable at every point in the segment, and there is an operator $w:[0,1] \to W$ such that $(-w)$ is monotone, $\lim_{t \to 0} w(t) = 0$, and

$$-w(|\Delta t|) \le \frac{1}{\Delta t}(F[z+(t+\Delta t)h]-F[z+th])-F'[z+th]h \le w(|\Delta t|)$$

for all t, Δt with $0 \le t \le 1$ and $0 \le t+\Delta t \le 1$.

This concept has been used by Kantorovich [14] for operators in spaces normed by a PTL space. In this case, the condition becomes;

$$\left\| \frac{1}{\Delta t}(F[z+(t+ \ t)h]-F[z+th])-F'[z+th]h \right\| \le w(|\Delta t|),$$

where the range of w is in the space which norms the range of F. The relation between uniform differentiability and continuous differentiability is given by;

<u>Theorem 12</u>. Let $F:D \subset Z \to W$ where Z and W are PTL spaces, D is convex, and F is continuous in D. If W is normal and F is uniformly differentiable on every segment in D^0, then F is continuously differentiable in D^0. If W is solid and F is continuously G- differentiable in D^0 then F is uniformly differentiable on every segment in D^0.

<u>Proof</u>. Let h be arbitrary in Z and let $z_1, z_2 \in D^0$. Then

$$\begin{aligned} F'[z_1]h - F'[z_2]h = {} & \tfrac{1}{\Delta t}(F[z_2+\Delta th]-F[z_2])-F'[z_2]h - \\ & \tfrac{1}{\Delta t}(F[z_1+\Delta th]-F[z_1])-F'[z_1]h + \\ & \tfrac{1}{\Delta t}(F[z_1+\Delta th]-F[z_2+\Delta th]+F[z_2]-F[z_1]) \\ \le {} & w_2(|\Delta t|) + w_1(|\Delta t|) + \\ & \tfrac{1}{\Delta t}(F[z_1+\Delta th]-F[z_2+\Delta th] + F[z_2]-F[z_1]) \end{aligned}$$

where Δt is small enough that $z_1+\Delta th$ and $z_2+\Delta th$ are in D^0. Similarly, we can show that

$$F'[z_1]h - F'[z_2]h \geq -w_2(|\Delta t|) - w_1(|\Delta t|) + \frac{1}{\Delta t}(F[z_1+\Delta th]-F[z_2+\Delta th]+F[z_2]-F[z_1]).$$

But, $w_i(\Delta t) \to 0$ as $\Delta t \to 0$, F is continuous, and W is normal, so it follows that $F'[z]h$ is continuous in z. Conversely, if W is solid and F is continuously differentiable then, for $(z,z+h) \subset D^0$, the function

$$G(t,\Delta t) = \frac{1}{\Delta t}(F[z+(t+\Delta t)h]-F[z+th])-F'[z+th]h$$

is continuous in t. Thus, for fixed Δt, the set

$$\{G(t,\Delta t) \;:\; 0 \leq t \leq 1\}$$

is compact. But compact sets are topologically bounded, and by lemma 3, there is an order bounded open set in W. Thus there is a positive w_0 in W; and a real function f, such that $\{G(t,\Delta t): 0 \leq t \leq 1\} \subset f(|\Delta t|)\cdot[-w_0,w_0]$. i.e., $-f(|\Delta t|)w_0 \leq G(t,\Delta t) \leq f(|\Delta t|)w_0$. Since $G(t,\Delta t) \to 0$ as $\Delta t \to 0$, uniformly for $0 \leq t \leq 1$, we can assume f is monotone decreasing and $\lim f(\Delta t) = 0$. Setting $w(\Delta t) = f(\Delta t)w_0$ completes the proof.

The hypothesis that W is solid in the last part of this theorem is necessary. To show this, consider the following example. Let $F:[0,1] \to L^1[0,1]$ be given by $F[t](x) = (t-x)^{1/3}$. Then F is continuously differentiable, with $F'[t] = \frac{1}{3}(t-x)^{-2/3}$, but the difference $\frac{1}{\Delta t}(F[t+\Delta t]-F[t])-F'[t]$ is not order bounded, hence no $w(\Delta t)$ can exist. That is, F is not uniformly differentiable.

We can now prove the important result;

<u>Theorem 13</u>. Let $F:[0,1] \to W$ be uniformly differentiable on $[0,1]$, where W is a normal PTL space. Then $\int_0^1 F'(t)dt$ exists, and $F(1) - F(0) = \int_0^1 F'(t)dt$.

<u>Proof</u>. Let $\{t_i\}$ be a partition of $[0,1]$. Then

$$F(1)-F(0)-\sum_{k=0}^{n-1} F'(t_k)(t_{k+1}-t_k) = \sum_{k=0}^{n-1}\left(F(t_{k+1})-F(t_k)-F'(t_k)(t_{k+1}-t_k)\right)$$
$$=\sum_{k=0}^{n-1}\left(\frac{1}{t_{k+1}-t_k}(F(t_{k+1})-F(t_k))-F(t_k)\right)(t_{k+1}-t_k)$$

Thus, $F(1)-F(0)-\sum_{k=0}^{n-1}F'(t_k)(t_{k+1}-t_k) \leq \sum_{k=0}^{n-1}w(t_{k+1}-t_k)(t_{k+1}-t_k)$
$\leq w(\max|t_{k+1}-t_k|)$

Similarly, we can show
$F(1)-F(0)-\sum_{k=0}^{n-1}F'(t_k)(t_{k+1}-t_k) \geq -w(\max|t_{k+1}-t_k|)$. Hence, by normality, as the size of the partition goes to zero, the summation converges to $F(1)-F(0)$, and the conclusion of the theorem follows.

The form in which this result will be most often used is;

Corollary. Let $F:D\subset Z\to W$ where Z and W are PTL spaces, W is normal, and F is uniformly differentiable on the segment $(z,z+h)\subset D^0$. Then

$$F[z+h]-F[z]=\int_0^1 F'[z+th]h\,dt.$$

Proof. Apply the theorem to the operator $f(t)=F[z+th]$.

Using the above theorems, we can now prove some important results about convex operators. First consider a real valued function $f(t)$ defined on some interval $[a,b]$ of the real line. If f is convex, then it has the following properties, any of which may be used to define convexity.

a) $f(\eta x_1+(1-\eta)x_2) \leq \eta f(x_1)+(1-\eta)f(x_2)$ for $0\leq\eta\leq 1$ and $x_1,x_2\in[a,b]$.

b) If $f'(x)$ exists for $x\in[a,b]$, then for $x,x+h\in[a,b]$, $f'(x)\leq \frac{1}{h}[f(x+h)-f(x)]$.

c) If $f''(x)$ exists for $x\in[a,b]$, then $f''(x)\geq 0$ for all $x\in[a,b]$.

These statements can be interpreted in any PTL space. Let $F:D\subset Z\to W$ where Z and W are PTL spaces, and D is a convex subset.

a') $F[\eta z_1+(1-\eta)z_2] \leq \eta F[z_1]+(1-\eta)F[z_2]$ for $0\leq\eta\leq 1$ and $z_1,z_2\in D$.

b') If F is G- differentiable in D, then for $z,z+h\in D$, $F'[z]h\leq F[z+h]-F[z]$.

c') If F is twice G- differentiable in D, then $F''[z] \geq 0$ for all z D.

For differentiable functions, a) and b) are equivalent, and we also have

Lemma 6. If F is G- differentiable in D then a') and b') are equivalent.

Proof. If b') holds, then for $x, y \in D$, $0 \leq \eta < 1$, we have $z = \eta x + (1-\eta)y \in D$ and

$$F[y] - F[z] \geq F'[z](y-z)$$
$$F[x] - F[z] \geq F'[z](x-z).$$

Multiplying the first by $(1-\eta)$, the second by η, and adding, gives, $\eta F[x] + (1-\eta)F[y] - F[z] \geq F'[z](\eta x + (1-\eta)y - z)$

$$= F'[z](z-z) = 0$$

hence a') holds. Conversely, if a') is assumed to be true, then for $0 \leq t \leq 1$ and $x, y \in D$,

$$F[tx + (1-t)y] \leq tF[x] + (1-t)F[y].$$

Hence, $\frac{1}{t}(F[tx + (1-t)y] - F[y]) \leq F[x] - F[y]$. Letting $t \to 0$, the left side converges to $F'[y](x-y)$, hence b') holds.

Condition b') is used by Collatz [7] to define a convex operator. In view of the above lemma, however, we use the more basic property a'). That is,

Definition 18. If $F: D \subset Z \to W$ where Z and W are PTL spaces, D is a convex subset, and F satisfies a'), then F is convex in D.

In the next chapter, we will need a property of convex operators which, in certain spaces, is equivalent to a').

Lemma 7. If Z is a PTL space, W a normal PTL space, and F is uniformly differentiable in every segment in D, then F is convex in D if and only if

d') $F'[u](v-u) \leq F'[v](v-u)$ for all u,v in D.

Proof. If F is convex then b') holds and hence we have,

$F'[u](v-u) \leq F[v]-F[u] \leq F'[v](v-u)$. Conversely, if d') holds, then by the corollary to theorem 13,

$$F[u+z]-F[u]=\int_0^1 F'[u+tz]zdt.$$

But, by d'), $F'[u]tz \leq F'[u+tz]tz$, hence, if $t \geq 0$, then $\int_0^1 F'[u+tz]zdt \geq \int_0^1 F'[u]zdt = F'[u]z$. Thus b') holds.

Condition c'), (i.e., $F''[z] \geq 0$), is not, in general, related to convexity. In E^n for example, a twice differentiable functional $F:D \subset E^n \to E^1$ is usually called convex if the quadratic form $\sum_{i,j=1}^{n} \frac{\partial^2 F}{\partial x_i \partial x_j} y_i y_j$ is positive definite, whereas, the condition $F''[z] \geq 0$ means that $\frac{\partial^2 F}{\partial x_i \partial x_j} \geq 0$ for all $i,j=1,2,\ldots,n$. These conditions are clearly not equivalent. In fact, they are not even related in the sense of one implying the other. The condition of positive definiteness can be generalized to

e') F is twice G- differentiable in D and $F''[z]hh \geq 0$ for all z, z+h in D.

Lemma 8. Let $F:D \subset Z \to W$ be twice G- differentiable in the convex set D, where Z and W are PTL spaces.

1) If F is convex, then e') holds.

2) If W is normal, and $F'[z]$ is uniformly differentiable on every segment in D, and e') holds, then F is convex in D.

Proof.

1) $F''[z]hh= \lim_{t \to 0} \frac{1}{t}(F'[z+th]h - F'[z]h)$. But convexity implies, by lemma 7, that $F'[z+th]th - F'[z]th \geq 0$, hence e').

2) $F'[u+h]h-F'[u]h = \int_0^1 F''[u+th]hhdt \geq 0$, and hence, again using lemma 7, we conclude that F is convex.

Notice that a'),b'),d'), and e') make sense even if the domain space has no order relation, while c') does not. Operators satisfying c') do have some properties which are very similar to b') and d').

Lemma 9. Let $F:D \subset Z \to W$ where W is a normal PTL space, Z is a PTL space, F is twice uniformly differentiable in every segment in the convex subset D, and $F''[z] \geqslant 0$ for all z in D. Then

1) $F'[u] \leqslant F'[v]$ for $u \leqslant v$ in D.

2) $F'[u](v-u) \leqslant F'[v](v-u)$ for $u \leqslant v$ or $v \leqslant u$, u and v in D.

3) $F'[u](v-u) \leqslant F[u]-F[v]$ for $u \leqslant v$ or $v \leqslant u$, u and v in D.

<u>Proof</u>.

1) Let $z \geqslant 0$ and use theorem 13 to write

$$F'[v]z-F'[u]z=\int_0^1 F''[u+t(v-u)]z(v-u)dt \geqslant 0.$$

2) If $u \leqslant v$ then 1) implies $F'[u](v-u) \leqslant F'[v](v-u)$. If $u \geqslant v$ then $F'[v](u-v) \leqslant F'[u](u-v)$, so $F'[v](v-u) \geqslant F'[u](v-u)$.

3) By part 2), $F'[u+t(v-u)](v-u) = \frac{1}{t}F'[u+t(v-u)](u+t(v-u)-u) \geqslant \frac{1}{t}F'[u](u+t(v-u)-u) = F'[u](v-u)$. Hence by theorem 13, $F[v]-F[u] = \int_0^1 F'[u+t(v-u)](v-u)dt \geqslant F'[u](v-u)$.

CHAPTER II

NEWTON'S METHOD

onvex Operators.

Let $F:D \subset Z \to W$ where Z and W are PTL spaces and F is a nonlinear operator. If F is G-differentiable in D^0 and if there is a sequence $\{z_n\}$ in D^0 which satisfies the linear equations

$$F'[z_n](z_{n+1} - z_n) = -F[z_n],$$

$n = 1,2,\ldots,$ then F is said to have a Newton sequence at z_0. We are interested in finding conditions on a convex operator which will be sufficient to guarantee that it has a Newton sequence which converges to a zero of the operator. Because of the similarities pointed out in chapter I (lemma 9), we will also consider operators whose second derivative is positive.

The theorems which will be proved differ from the Newton method theorems of Kantorovich [13] in several ways. First, we use only the G-derivative, whereas the results of Kantorovich require the operator to be twice F-differentiable. Secondly, instead of Banach spaces, we will use various types of PTL spaces, and by using these spaces, we are able to replace bounds on the norms of the first and second derivatives by hypotheses of the form $-F'[z] \leq \Gamma$, where Γ is a linear invertible operator. On the other hand, Kantorovich's theorems imply that $F'[z_k]$ has a continuous inverse, a fact which is not true in our case, as will be shown later by an example. Some related results of Baluev [2,3] and Slugin [24,25] will also be discussed. These results, which pertain to Chaplygin methods, can also be applied to Newton's method, however, it will easily be seen that they are weaker theorems than those which will now be proven.

Before giving the main result, we will prove three lemmas. The first two show that in order for a convex operator, or an operator with positive second derivative, to have a Newton sequence, it is necessary for there to be a point at which the operator takes a positive value.

Lemma 1. Let $F:D \subset Z \to W$ where Z and W are PTL spaces, F is convex and G-differentiable in D^0, and z_0, z_1 are points in D^0 which satisfy

$$F'[z_0](z_1 - z_0) = -F[z_0]. \tag{1}$$

Then $F[z_1] \geqslant 0$.

Proof. Convexity implies $F'[z_0](z_1 - z_0) \leqslant F[z_1] - F[z_0]$. Subtracting this from (1) gives $F[z_1] \geqslant 0$.

A similar but slightly weaker result holds for operators with positive second derivative.

Lemma 2. Let $F:D \subset Z \to W$, where Z and W are PTL spaces, W is normal, F is twice uniformly differentiable on every segment in D, and $F''[z] \geqslant 0$ for all $z \in D^0$. If $z_0, z_1 \in D^0$ and either $z_0 \leqslant z_1$ or $z_1 \leqslant z_0$, then equation (1) implies $[z_1] \geqslant 0$.

Proof. By lemma 9, chapter I, the equation

$$F'[z_0](z_1 - z_0) \leqslant F[z_1] - F[z_0]$$

is valid. Hence, as in lemma 1, $F[z_1] \geqslant 0$.

The next lemma is a fixed point theorem of Kantorovich [12] which will be used extensively throughout this chapter.

Lemma 3. Let $V:Z \to Z$ where Z is a regular PTL space, and V satisfies

1) V is continuous;
2) $V[0] \geqslant 0$
   $V[z^*] \leqslant z^*$ for some $z^* \geqslant 0$;
3) if $\Delta z \geqslant 0$ then $V[z+\Delta z] \geqslant V[z]$ for any $z \in Z$.

Then V has a fixed point in the interval $[0,z^*]$.

Proof. Let $z_0=0$, $z_{n+1}= V[z_n]$. Then $z_1=V[z_0] \geq 0 = z_0$ and $z_1= V[z_0] \leq V[z^*] \leq z^*$, hence $z_1 \in [z_0,z^*]$. If z_n is in $[z_{n-1},z^*]$, then $z_{n+1}= V[z_n] \geq V[z_{n-1}] = z_n$, and

$$z_{n+1}= V[z_n] \leq V[z^*] \leq z^*,$$

hence $z_{n+1} \in [z_n,z^*]$. Therefore, by induction, $0 \leq z_0 \leq z_1 \leq z_2 \leq \ldots \leq z_n \leq z_{n+1} \leq \ldots \leq z^*$. Since Z is regular, $\lim_{n\to\infty} z_n = z'$ exists, $z' \in [0,z^*]$ and V is continuous so $V[z'] = z'$.

We are now ready to state the main result.

Theorem 1. Let $F:D \subset Z \to W$ where Z and W are PTL spaces, Z is regular and normal, and $[z_0^*,z_1^*]$ is an interval in D^0 such that

1) F is continuous, G-differentiable, and convex in $[z_0^*,z_1^*]$;

2) $F[z_0^*] \geq 0 \geq F[z_1^*]$;

3) there is a linear operator $\Gamma :Z \to W$ which has a continuous positive inverse; and $-F'[z] \leq \Gamma$ holds for all $z \in [z_0^*,z_1^*]$.

Then, F has a Newton sequence at z_0^* which is monotone increasing and converges to $z^* \in [z_0^*,z_1^*]$, where $F[z^*] = 0$.

Proof. Let $z_0= z_0^*$ and $V[z] = z + \Gamma^{-1}(F[z_0]+F'[z_0]z)$. Then $V:Z \to Z$, V is continuous, and

$$V[0] = \Gamma^{-1}F[z_0^*] \geq 0, \text{ since } \Gamma^{-1} \text{ is positive,}$$

$$\begin{aligned} V[z_1^* - z_0] &= z_1^* - z_0 + \Gamma^{-1}(F[z_0]+F'[z_0](z_1^* - z_0)) \\ &\leq z_1^* - z_0 + \Gamma^{-1}F[z_1^*], \text{ by convexity,} \\ &\leq z_1^* - z_0 \text{ since } F[z_1^*] \leq 0 \text{ and } \Gamma^{-1} \text{ is positive.} \end{aligned}$$

Furthermore, if $\Delta z \geq 0$ then

$$V[z +\Delta z] = V[z] +\Delta z + \Gamma^{-1}F'[z_0]\Delta z \geq V[z],$$

since $\Gamma^{-1}F'[z_0]\Delta z \geq -\Delta z$. By lemma 3, V has a fixed point $z' \in [0,z_1^*-z_0]$. Let $z_1= z_0+ z'$. Then $z_0 \leq z_1 \leq z_1^*$ and $F'[z_0](z_1-z_0) = -F[z_0]$. By lemma 2, $F[z_1] \geq 0$, so we can replace z_0^* by z_1 and use induction to show that $\{z_n\}$ exists, and $z_0 \leq z_1 \leq \ldots \leq z_n \leq z_{n+1} \leq \ldots \leq z_1^*$. Since Z is regular, $\lim_{n\to\infty} z_n = z^*$ exists, and $z^* \in [z_0^*, z_1^*]$. Finally,

$$0 \leq F[z_n] = -F'[z_n](z_{n+1} - z_n) \leq \Gamma\,(z_{n+1} - z_n)$$

so $0 \leq \Gamma^{-1}F[z_n] \leq z_{n+1} - z_n$. But Z is normal, and $z_{n+1} - z_n \to 0$, hence $\Gamma^{-1}F[z_n] \to 0$. By continuity, $F[z^*] = 0$.

Before discussing the hypotheses and implications of this theorem, we will consider a simple application. The equation $u(x) = \int_0^1 T(x,y)[\alpha^2 \sin u(y) - f(y)]dy$ arises in the study of the forced oscillations of finite amplitude of a pendulum [27]. The function T is given by

$$T(x,y) = \begin{cases} x(1-y) & , 0 \leq x \leq y \\ y(1-x) & , y \leq x \leq 1 \end{cases}$$

and we assume that f is a continuous function, with $0 \leq f \leq M$. The spaces we use here are $Z = W = L^1[0,1]$, with the usual ordering and topology. Let

$$F[u](x) = -u(x) + \int_0^1 T(x,y)[\alpha^2 \sin u(y) - f(y)]dy.$$

Then

$$F'[u]h(x) = -h(x) + \alpha^2 \int_0^1 T(x,y)\cos[u(y)]h(y)dy.$$

Thus, F is G-differentiable and, for $-\pi/2 \leq x \leq 0$, $\sin(x)$ is convex, so F is convex in the interval $[-\pi/2, 0]$. Assume, furthermore, that $\alpha^2 + M \leq 4\pi$, in which case

$$\begin{aligned} F[-\pi/2] &= \pi/2 + \int_0^1 T(x,y)[-\alpha^2 - f(y)]dy \\ &\geq \pi/2 - (\alpha^2 + M)\int_0^1 T(x,y)dy \geq \pi/2 - 1/8(\alpha^2 + M) \geq 0, \end{aligned}$$

$$F[0] = -\int_0^1 T(x,y)f(y)dy \leq 0.$$

Finally, if $h \geq 0$ and $u \in [-\pi/2, 0]$, then $-F'[u]h \leq h$, hence we can let $\Gamma = I$, the identity operator. All the conditions of the theorem are satisfied, so there must exist a Newton sequence, starting at $u_0(x) \equiv -\pi/2$, which converges, monotonically, to a solution $u^*(x)$ which satisfies $-\pi/2 \leq u^* \leq 0$.

Referring to the table in chapter I, we see that the space Z in this theorem can be E^n, H, or $L^p[0,1]$, $0 < p < \infty$. That the theorem does not hold in $C^n[0,1]$ or $L^\infty[0,1]$, is shown by the following example. Let $F: D \subset C^n[0,1] \to W$, where

$D = \{x : 0 \leq x(t) \leq 1\}$ $W = \{y : (1-t)^2 y(t) \in L^1[0,1]\}$,
$F[x](t) = (1 - x(t))^{1/1-t}$. If the topology in W is given by the norm $\|y\| = \int_0^1 |y(t)| (1-t)^2 dt$, then F is continuous, twice G-differentiable, and

$$F'[x]h = -\tfrac{1}{1-t}(1 - x(t))^{t/1-t}\, h(t)$$
$$F''[x]hh = \tfrac{t}{(1-t)^2}(1 - x(t))^{\frac{2t-1}{t}}\, h^2(t)$$

Hence, $F''[x]hh \geq 0$ for all h, so by a result in chapter I, F is convex. Moreover, $-F'[x] \leq \frac{1}{1-t}$, and the operator $\Gamma[y](t) = \frac{1}{1-t} y(t)$ has a positive continuous inverse. Finally, $F[0] \geq 0 \geq F[1]$, so F satisfies the conditions of the theorem, in the interval [0,1]. But, the Newton sequence for F at 0 is $x_n(t) = 1 - t^n$, which does not converge in $C^n[0,1]$. The same example can be used in the space $L^\infty[0,1]$.

From the proof of the theorem, we see that one method for solving the sequence of linear equations is to apply successive approximation to the operator V. That is, if z_n has been found, then let

$$V_n[z] = z + \Gamma^{-1}(F[z_n] + F'[z_n]z),$$

and compute the sequence $x_k = V_n[x_{k-1}]$, where $x_0 = 0$. Then, from the proof of the theorem, $0 \leq x_1 \leq x_2 \leq \ldots \leq z_1^* - z_n$, $\lim x_k = x^*$ exists, and $z_{n+1} = z_n + x^*$.

The proof also shows that hypothesis 3) can be replaced by the slightly weaker hypothesis

3') $-\Gamma^{-1}F'[z] \leq I$, for $z \in [z_0^*, z_1^*]$, where I is the identity operator, and Γ^{-1} is a positive continuous linear operator.

This is weaker than 3) because we have assumed $\Gamma^{-1} \geq 0$, but not $\Gamma \geq 0$. In the case of a real function of a real variable, condition 3) means that the derivative is bounded in the interval.

Because of the monotonicity of the Newton sequence, the same proof holds if convexity is replaced by $F''[z] \geq 0$ for $z \in [z_0^*, z_1^*]$, provided W is normal and F is twice uniformly differentiable, so that lemma 2 is applicable. In this case, we also have $-F'[z] \leq -F'[z_0^*]$ for all $z \in [z_0^*, z_1^*]$, hence the result

Corollary. Let $F: D \subset Z \to W$ where Z is a regular PTL space, W is a normal PTL space, F is twice uniformly differentiable on every segment in an interval $[z_0^*, z_1^*] \subset D^0$, and

1) $F''[z] \geq 0$ for all z in the interval;
2) $F[z_0^*] \geq 0 \geq F[z_1^*]$;
3) $-F'[z_0^*]$ has a continuous positive inverse.

Then, F has a Newton sequence at z_0^* which converges, monotonically, to a solution $z^* \in [z_0^*, z_1^*]$ of $F[z] = 0$.

An example of a problem which can be solved using this corollary, but cannot be handled by theorem 1, is the Chandrasekhar H-equation, which arises in radiative transfer [1,8];

$$H(x) = 1 + xH(x)\int_0^1 \frac{f(t)\,H(t)}{x+t}\,dt$$

where f(t) is a given function. If we assume that $f(t) \geq 0$, and let

$$F[z](x) = 1 + xz(x)\int_0^1 \frac{f(t)\,z(t)}{x+t}\,dt - z(x)$$

then

$$F'[z]h(x) = xh(x)\int_0^1 \frac{f(t)\,z(t)}{x+t}\,dt + xz(x)\int_0^1 \frac{f(t)h(t)}{x+t}\,dt - h(x).$$

Convexity of F would require that $xh(x)\int_0^1 \frac{f(t)\,h(t)}{x+t}\,dt \geq 0$, for all h such that z and z+h lie in some interval. This is not true, but we do have

$$F''[z]hk(x) = xh(x)\int_0^1 \frac{f(t)\,k(t)}{x+t}\,dt + xk(x)\int_0^1 \frac{f(t)\,h(t)}{x+t}\,dt$$

and hence $F''[z] \geq 0$ for all z. Finally, $F[0] = 1 \geq 0$, $-F'[0]h = h$ and, if $0 \leq f(t) \leq 1/4$, then $F[2] \leq 0$. Hence, the corollary shows that there is a solution to the equation in the interval [0,2] which can be found by applying Newton's method. The bound on f which guarantees that $F[2] \leq 0$ can be relaxed. For example, it is known [8] that if $f(t) \equiv \frac{1}{2}$ then the equation has a positive solution, and we can use this solution as the upper bound, instead of $z(t) \equiv 2$.

The following table shows the results of some numerical calulations for this equation. The iteration was started at $z_0(t) \equiv 0$, and the linear equations were solved by replacing the integral by a numerical quadrature formula, using 24 subdivisions, and then solving the resulting linear system by Gaussian elimination.

n	$z_n(.25)$	$z_n(.50)$	$z_n(.75)$	$z_n(1.0)$
		$f(t) = 0.25$		
	1.0	1.0	1.0	1.0
2	1.1271957	1.1825743	1.2175054	1.2419763
3	1.1296431	1.1877202	1.2249139	1.2512373
4	1.1296444	1.1877242	1.2249208	1.2512471
5	1.1296445	1.1877242	1.2249208	1.2512471
		$f(t) = 0.50$		
1	1.0	1.0	1.0	1.0
2	1.3532370	1.5429468	1.6749577	1.7736915
3	1.4616778	1.7905519	2.0628780	2.2945761
4	1.5051736	1.9031322	2.2626224	2.5933702
9	1.5426216	2.0006399	2.4412844	2.8730492
10	1.5426290	2.0006590	2.4413195	2.8731045
11	1.5426288	2.0006585	2.4413185	2.8731029
		$f(t) = 0.25t$		
1	1.0	1.0	1.0	1.0
2	1.0413517	1.0637265	1.0783764	1.0888112
3	1.0414704	1.0640055	1.0787985	1.0893541
4	1.0414704	1.0640055	1.0787985	1.0893541

Table 2. Newton Sequence for Chandrasekhar Equation.

The above corollary is similar to the following result of Baluev [3].

Theorem (Baluev). Let $F:D\subset Z\to W$ where Z is a normal, regular PTL space, W is a normal PTL space, and D^0 contains an interval $[x_0,y_0]$ such that

1) $F[x_0]\geqslant 0\geqslant F[y_0]$;

2) F is continuous and twice uniformly differentiable on every segment in $[x_0,y_0]$;

3) the operator $-F'[x]$ has a positive inverse for every $x\in[x_0,y_0]$;

4) $F''[x]\geqslant 0$ for all $x\in[x_0,y_0]$;

5) $(F'[x_0])^{-1}$ is continuous.

Then, the equation $F[x] = 0$ has a unique solution x^* in $[x_0,y_0]$. The elements x_n and y_n, determined by the formulae

$$x_n = x_{n-1} - (F'[x_{n-1}])^{-1}F[x_{n-1}],$$
$$y_n = y_{n-1} - (F'[x_{n-1}])^{-1}F[y_{n-1}]$$

satisfy $x_{n-1}\leqslant x_n\leqslant x^*\leqslant y_n\leqslant y_{n-1}$, $F[x_n]\geqslant 0\geqslant F[y_n]$, and

$$\lim x_n = \lim y_n = x^*.$$

(The statement of this theorem has been altered slightly to allow a direct comparison with the corollary. Baluev's proof remains unchanged.)

The important difference between these two results is Baluev's assumption that $F'[x]$ has an inverse. To illustrate the implications of this, consider the operator $F:E^2\to E^2$ defined by $F(x,y) = (f(x,y),g(x,y))$ where

$$f(x,y) = \begin{cases}(1+x)^4, & x\leqslant -1\\ 0, & x\geqslant -1\end{cases}$$

and $g(x,y) = y^2$. Then for $(-2,-1)\leqslant(x,y)\leqslant(0,0)$, $F''(x,y)\geqslant 0$ and $F(-2,-1)\geqslant 0\geqslant F(0,0)$. We use here the usual ordering in E^2, given by the cone K_1 as defined in chapter 1. Furthermore, $F'(x,y)$ is defined by the matrix

$$\begin{pmatrix} h(x) & 0\\ 0 & 2y\end{pmatrix}$$

where

$$h(x) = \begin{cases} 4(1+x)^3 & , x \leq -1 \\ 0 & , x \geq -1. \end{cases}$$

Hence,

$$-F'(-2,-1) = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}$$

which has a positive inverse, and so the hypotheses of the corollary are satisfied. But, if $x \geq -1$, then $F'(x,y)$ is singular, so Baluev's theorem cannot be applied.

The above example also shows that the hypotheses of theorem 1 are not sufficient to insure that $F'[z_n]$ has an inverse.

The rate of convergence for the Newton sequence in theorem 1, and its corollary, is similar to the convergence rate for the real case. That is, convergence is super-linear if F is continuously G-differentiable, and is quadratic if F is twice uniformly differentiable. More precisely, we have

<u>Theorem 2</u>. Let $F:D \subset Z \to W$ where Z and W are PNL spaces, F is continuously G-differentiable and convex in some interval $[z_0^*, z_1^*] \subset D^0$. Suppose F has a monotone increasing Newton sequence $\{z_n\} \subset [z_0^*, z_1^*]$ which converges to z^* where $F[z^*] = 0$. Furthermore, suppose $-F'[z_n] \geq \Gamma_1$ where $\Gamma_1 : Z \to W$ is a linear operator with a positive continuous inverse. Then

$$\|z^* - z_{n+1}\| \leq \|\Gamma_1^{-1}\| \, r(z^* - z_n) \, \|z^* - z_n\|$$

where $\lim_{z_n \to z^*} r(z^* - z_n) = 0$.

<u>Proof</u>.
$$\begin{aligned} \Gamma_1 (z^* - z_{n+1}) &\leq -F'[z_n](z^* - z_{n+1}) \\ &= -F'[z_n](z^* - z_n) - F'[z_n](z_n - z_{n+1}) \\ &= -F'[z_n](z^* - z_n) - F[z_n]. \end{aligned}$$

But, by convexity,

$$-F'[z^*](z^* - z_n) \leq F[z_n] - F[z^*] = F[z_n]$$

so

$$\Gamma_1 (z^* - z_{n+1}) \leq -F'[z_n](z^* - z_n) + F'[z^*](z^* - z_n)$$

Therefore,

$$0 \leq (z^* - z_{n+1}) \leq \Gamma_1^{-1}(-F'[z_n] + F'[z^*])(z^* - z_n),$$

and

$$\| z^* - z_{n+1} \| \leq \| \Gamma_1^{-1} \| \cdot \| F'[z_n] - F'[z^*] \| \cdot \| z^* - z_n \|$$

Letting $r(z_n - z^*) = \| F'[z_n] - F'[z^*] \|$ gives the result.

Theorem 3. In addition to the hypotheses of theorem 2, assume that F is twice uniformly differentiable on every segment in D^0, and $\| F''[z] \| \leq M$ for all $z \in [z_0^*, z_1^*]$. Then

$$\| z^* - z_{n+1} \| \leq M \; \| \Gamma_1^{-1} \| \cdot \| z^* - z_n \|^2.$$

Proof. In the previous proof, we have shown that

$$0 \leq z^* - z_{n+1} \leq \Gamma_1^{-1}(-F'[z_n](z^* - z_n) + F'[z^*](z^* - z_n)).$$

But, by the integral theorem of chapter I,

$$-F'[z_n](z^* - z_n) + F'[z^*](z^* - z_n) =$$

$$= \int_0^1 F''[z^* - t(z^* - z_n)](z^* - z_n)(z^* - z_n)\,dt.$$

So we have,

$$\| z^* - z_{n+1} \| \leq \quad \| \Gamma_1^{-1} \| \int_0^1 \| F''[z^* - t(z^* - z_n)] \| \, dt \, \| z^* - z_n \|$$

$$\leq M \quad \| \Gamma_1^{-1} \| \cdot \| z^* - z_n \|^2.$$

These two theorems also hold if convexity is replaced by $F''[z] \geq 0$. In this case, we can set $\Gamma_1 = -F'[z_1^*]$, provided this operator has a positive continuous inverse.

The problem of finding the endpoints z_0^*, z_1^* of the interval can be difficult at times. If the equation is known to have a solution, then this solution may often be used for z_1^*. For example, the equation of Bratu [8];

$$y(x) = \eta \int_0^1 T(x,t) \; e^{y(t)} dt,$$

where

$$T(x,t) = \begin{cases} t(1-x) & , \; 0 \leq t \leq x \leq 1 \\ x(1-t) & , \; 0 \leq x \leq t \leq 1 \end{cases}$$

is known to have two continuous positive solutions, if $0 < \eta < \eta_0$, where η_0 is a positive constant, approximately equal to 3.497. If we let

$$F[y] = \eta \int_0^1 T(x,t) \; e^{y(t)} dt - y(x),$$

and $Z = W = L^2[0,1]$, then for any bounded y, F is continuous, G-differentiable, convex, and

$$\begin{aligned} -F'[y]h(x) &= h(x) \quad \eta\int_0 T(x,t)e^{y(t)}h(t)\,dt \\ &\leq h(x) \end{aligned}$$

for all $h \geq 0$, hence we can let $\Gamma[h] = h$. Furthermore, $F[0] = \eta\int_0^1 T(x,t)\,dt = \eta\tfrac{1}{2}x(1-x) \geq 0$, so we can set $z_0^* = 0$ and $z_1^* = z^*$, where z^* is either of the solutions.

It was observed earlier that theorem 1 does not guarantee the existence of $F'[z_n]^{-1}$. However, by strengthening the hypotheses slightly, we can prove that $F'[z_n]$ is invertible, and hence the Newton sequence can be written in the more usual form

$$z_{n+1} = z_n - F'[z_n]^{-1}F[z_n].$$

This is the content of the next theorem.

Theorem 4. Let $F{:}D \subset Z \to W$ where Z is regular and normal, and W is solid. Let $[z_0^*, z_1^*] \subset D^0$ be an interval such that

1) F is continuous, G-differentiable, and convex in $[z_0^*, z_1^*]$;

2) $F[z_0^*] \geq 0 \geq F[z_1^*]$;

3) $-F'[z] \leq \Gamma$, for all $z \in [z_0^*, z_1^*]$, where Γ is a linear operator with a continuous positive inverse;

4) either

i) $Z = W = E^n$

or

ii) $F'[z]$ is one-to-one, for all $z \in [z_0^*, z_1^*]$;

5) either

i) $F[z_1^*] \ll 0$ (i.e., $-F[z_1^*]$ is an interior point of the positive cone.)

or

ii) $-F'[z] \geq \Delta_z$ where Δ_z is a linear operator with a continuous positive inverse.

Then, F has a Newton sequence $\{z_n\}$ at z_0^* which is monotone increasing and converges to a solution of $F[z] = 0$. Moreover, $-F'[z_n]^{-1}$ exists and is positive.

Proof. Conditions 1),2),3) ensure the existence of the Newton sequence. We will use condition 5) to show that $F'[z_n]$ maps Z onto W and then, by 4), $F'[z_n]^{-1}$ is defined. Since W is solid, it is also reproducing. Hence it suffices to show that $-F'[z_n]$ maps Z onto the positive cone of W. Let $w_0 \geqslant 0$, and let

$$V[z] = z + \Gamma^{-1}(w_0 + F'[z_n]z).$$

If 5i) holds, let $z' = \alpha(z_1^* - z_n)$ where $\alpha > 0$ is such that $-\alpha F[z_1^*] \geqslant w_0$. If 5ii) holds, let $z' = \Delta_{z_n}^{-1} w_0$. In the first case, we have

$$\begin{aligned} F'[z_n]z' &= \alpha F'[z_n](z_1^* - z_n) \\ &\leqslant \alpha F[z_1^*] - \alpha F[z_n] \\ &\leqslant \alpha F[z_1^*] \leqslant -w_0 \end{aligned}$$

and in the second case,

$$\begin{aligned} F'[z_n]z' &= F'[z_n]\Delta_{z_n}^{-1} w_0 \\ &\leqslant -\Delta_{z_n}(\Delta_{z_n}^{-1} w_0) = -w_0 \end{aligned}$$

hence, in either case, $z' \geqslant 0$ and $F'[z_n]z' + w_0 \leqslant 0$. Now, V is continuous, and

1) $V[0] = \Gamma^{-1} w_0 \geqslant 0$,
2) $V[z'] = z' + \Gamma^{-1}(w_0 + F'[z_n]z') \leqslant z'$,
3) if $\Delta z \geqslant 0$ then $V[z+\Delta z] = V[z] + \Delta z + \Gamma^{-1}F'[z_n]\Delta z \geqslant V[z]$.

Thus, by lemma 3, V has a fixed point, $V[z^*] = z^*$, and $0 \leqslant z^* \leqslant z'$. But then $-F'[z_n]z^* = w_0$. Hence, $-F'[z_n]$ maps Z onto the positive cone of W, and this, together with condition 4) implies that $-F'[z_n]^{-1}$ exists and is positive.

In practice, it is often very difficult to find even an approximate solution to the set of linear equations which determines the Newton sequence. This fact leads to two interesting questions; will a sequence of approximate solutions converge to a solution of the nonlinear problem, and, is it possible to replace the set of linear equations by another set which can be solved more easily and such that the solutions to these new equations still converge to a solution of the nonlinear problem? One simple answer to the first question is given by the observation that, if $\{z_n\}$ is a monotone Newton sequence for F which converges to z^*, where $F[z^*] = 0$, and $\{z'_n\}$ is an approximate sequence which satisfies $z_n \leq z'_{n+1} \leq z^*$, then $0 \leq z^* - z'_{n+1} \leq z^* - z_n$, so, by the normality of Z, $\lim z'_n = z^*$. This means that if we solve the linear equations by successive approximations, as suggested earlier, it is only necessary to carry out a few steps of this process, since any of the successive approximations will give a z'_{n+1} which satisfies the above inequality.

To answer the second question, we consider equations of the form $\Gamma_n(z_{n+1} - z_n) = -F[z_n]$, and look for conditions on the operators Γ_n which will guarantee the existence and convergence of the sequence $\{z_n\}$ to a solution of $F[z]=0$. Theorems of this kind have been proven by Slugin [24,25] and Baluev [2,3]. Their theorems are concerned with two-sided approximations. However, using the same type of proof, the following result is easily proven.

Theorem (Slugin). Let $F:D \subset Z \to W$, where Z is a normal regular PTL space, W is a normal PTL space, and $[z_0^*, z_1^*] \subset D^0$ is an interval such that

1) F is continuous, G-differentiable, and convex in the interval;

2) $F[z_0^*] \geq 0 \geq F[z_1^*]$;

3) for every z in the interval there exists a linear invertible operator Γ_z such that Γ_z^{-1} is positive and

$$-F'[z] \leq \Gamma_z \leq \Gamma$$

where Γ is also linear with a positive continuous inverse; Then, there is a sequence $\{z_n\} \subset [z_0^*, z_1^*]$ which is defined by

$$z_0 = z_0^*, \; z_{n+1} = z_n + \Gamma_{z_n}^{-1} F[z_n],$$

and satisfies $z_0 \leq z_1 \leq \ldots \leq z_n \leq z_{n+1} \leq \ldots \leq z_1^*$. Moreover, $\lim z_n = z^*$ exists and $F[z^*] = 0$.

Using the methods of Slugin and Baluev, the hypothesis that Γ_z^{-1} exists is certainly needed. However, with the ideas of theorem 1, we can remove this condition. The resulting theorem is;

<u>Theorem 5</u>. Let $F: D \subset Z \to W$ where Z is a regular normal PTL space, W is a PTL space, and $[z_0^*, z_1^*] \subset D^0$ is an interval such that

1) F is continuous, G-differentiable, and convex in the interval;

2) $F[z_0^*] \geq 0 \geq F[z_1^*]$;

3) for every z in the interval, there exists a Γ_z which is continuous, linear, and satisfies

$$-F'[z] \leq \Gamma_z \leq \Gamma$$

where Γ is a linear operator which has a continuous positive inverse.

Then, there is a sequence $\{z_n\} \subset [z_0^*, z_1^*]$ which satisfies

$$\Gamma_{z_n}(z_{n+1} - z_n) = F[z_n]$$

and $z_0^* = z_0 \leq z_1 \leq \ldots \leq z_n \leq z_{n+1} \leq \ldots \leq z_1^*$, $\lim z_n = z^*$ where $F[z^*] = 0$.

<u>Proof</u>. The proof is identical to that of theorem 1, with $-F'[z_n]$ replaced by Γ_{z_n}.

An illustration of this result is the Liebmann iteration as described by Greenspan and Parter [10]. When the mildly nonlinear elliptic partial differential equation

$$\Delta u = g(u,x,y)$$

is discretized, in the usual way, the resulting system of equations has the form $AX = f(X)$, where $X = (x_1, x_2, \ldots, x_n)$ and A is an n by n matrix with a positive inverse and non-positive off- diagonal elements. If $g(u,x,y)$ is convex as a function of u, then $f(x) = (f_1(x_1), \ldots, f_n(x_n))$ where $f_i(x)$ is convex. Newton's method applied to this system of equations gives

$$AX^{(k+1)} - D(X^{(k)})X^{(k+1)} = D(X^{(k)})X^{(k)} + f(X^{(k)})$$

where $D(X)$ is a diagonal matrix with elements $f_i'(x_i)$. Thus, finding $X^{(k+1)}$ requires the inverting of the matrix $A-D(X^{(k)})$. The process can be simplified considerably if we assume $-f_i'(x) \leq m$, and write $A = D - L - U$, where D is a diagonal matrix, L is lower triangular, and U is upper triangular. If $F(X) = f(X) - AX$, then

$$\begin{aligned} -F'(X) &= A - f'(X) \leq A + mI \\ &= D + mI - L - U \\ &\leq D + mI - L \end{aligned}$$

since $U \geq 0$. Thus, letting $\Gamma_z = \Gamma = D + mI - L$, theorem 5 gives the sequence of equations

$$(D + mI - L)X^{(k+1)} = UX^{(k)} + f(X^{(k)}) + mX^{(k)}$$

which is easily solvable since the matrix on the left is lower triangular. Theorem 5 says that if $F(X^{(0)}) \geq 0 \geq F(\bar{X}^{(1)})$ then the sequence $\{X^{(k)}\}$ is monotone increasing and converges to a solution of $AX = f(X)$.

Mildly Nonlinear Equations

Because the spaces $C^n[0,1]$ are not regular, we cannot use theorem 1 directly to solve nonlinear differential equations. However, if the equation is a boundary value problem, of the form $L[u] = f(u)$, with appropriate boundary conditions, where L is a linear differential operator which has a positive Green's function, then the differential equation can be replaced by the equivalent integral equation

$$u(x) = \int_R G(x,t)f(u(t))dt.$$

We can now investigate this equation in the regular space $L^2[D]$. This method can be applied, for example, to the problem

$$u''(x) = -e^{u(x)}, \quad x \in [0,1]$$
$$u(0) = u(1) = 0.$$

which is equivalent to the equation of Bratu

$$u(x) = \int_0^1 T(x,t)e^{u(t)}dt$$

which we solved in the preceeding section. Consider, however, the equation

$$u''(x) = -e^{-u(x)}, \quad x \in [0,1]$$
$$u(0) = u(1) = 0.$$

Proceeding as before, we have the equivalent equation

$$u(x) = \int_0^1 T(x,t)e^{-u(t)}dt.$$

We let $Z = W = L^2[0,1]$, $D = \left\{ u : u(0) = u(1) = 0 \right\}$, and

$$F[u](x) = \int_0^1 T(x,t)e^{-u(t)}dt - u(x).$$

But then

$$-F'[u]h(x) = h(x) + \int_0^1 T(x,t)e^{-u(t)}h(t)dt,$$

and if $u(t) \geqslant 0$, then

$$-F'[u]h(x) \leq h(x) + \int_0^1 T(x,t)h(t)dt.$$

Letting $\Gamma[h] = h(x) + \int_0^1 T(x,t)h(t)dt$, then Γ has an inverse given by

$$\Gamma^{-1}[k](x) = k(x) - \int_0^1 R(x,y)k(t)dt$$

where $R(x,t)$ is the resolvent kernel for T. But T is positive, so R is also positive, and hence Γ'^{-1} is not positive. Returning to the differential equation, and ignoring for a moment the problems of convergence, let

$$F[u] = e^{-u} + u''.$$

Then formally, we have

$$-F'[u]h = -h'' + e^{-u}h \leqslant -h'' + \alpha h,$$

for all u such that $e^{-u(t)} \leqslant \alpha$. But, it is known [5] that the operator $G[h] = -(h'' - \alpha h)$ has a positive Green's function provided $\alpha \geqslant 0$. Furthermore, F is convex and $F[0] = 1 \geqslant 0 \geqslant F[\frac{1}{2}x(1-x)]$. Thus, it appears that this operator satisfies all the conditions of theorem 1, except that the domain space, $C^2[0,1]$, is not regular, and F is not continuous. Note that $F'[u]$ as defined above is not a continuous operator, hence is not a G-derivative. In the remainder of this chapter, we will study operators which can be written in the form $F = f - L$, where L is linear, but not continuous, and f is G-differentiable. Then $F'[u]$ will denote the (dis-continuous) linear operator $f'(u) - L$, where $f'(u)$ is the G-derivative of f at u.

In the next theorem, we prove a result similar to theorem 1, in which the space Z does not have to be regular.

<u>Theorem 6</u>. Let $F:D \subset Z \to W$ where Z is normal, W is regular and normal, and $F = f - L$ where f and L are operators which satisfy the following conditions.

1) L is linear and has a completely continuous positive inverse on W;

2) f is continuous, G-differentiable, and convex in some interval $[z_0^*, z_1^*] \subset D^0$

3) $L[z_0^*] \leqslant f(z_0^*)$
$L[z_1^*] \geqslant f(z_1^*)$;

4) there is a continuous linear operator $g:Z \to W$ such that, for every z in the interval, $f'(z) \geqslant g$, and moreover, the operator $\Gamma = L - g$ has a positive continuous inverse. Then, F has a Newton sequence $\{z_n\}$ at z_0^* which is monotone increasing and converges to a solution of $L[z] = f(z)$.

<u>Proof</u>. Even though $F = f - L$ is not G-differentiable, if we let $F'[u]h = f'(u)h - L[h]$, then

$F'[u]h \leqslant f(u+h) - f(u) - L[u+h] + L[u] = F[u+h] - F[u]$, hence F formally satisfies the convexity condition. Now, let

$$V[w] = w + F[z_0] + F'[z_0]\Gamma^{-1}w$$

where $z_0 = z_0^*$. Then $V: W \to W$, continuously, and

$$V[0] = F[z_0] \geqslant 0, \text{ (by 3)},$$

$$\begin{aligned} V[\Gamma(z_1^* - z_0)] &= \Gamma(z_1^* - z_0) + F[z_0] + F'[z_0](z_1^* - z_0) \\ &\leqslant \Gamma(z_1^* - z_0) + F[z_1^*] \\ &\leqslant \Gamma(z_1^* - z_0), \end{aligned}$$

and $\Gamma(z_1^* - z_0) \geqslant -F'[z_0](z_1^* - z_0) \geqslant F[z_0] - F[z_1^*] \geqslant 0$.

Finally, if $\Delta w \geqslant 0$, then

$$V[w+\Delta w] = V[w] + \Delta w + F'[z_0]\Gamma^{-1}\Delta w \geqslant V[w],$$

since, by 4),

$$F'[z_0]\Gamma^{-1}\Delta w \geqslant -\Gamma\Gamma^{-1}\Delta w.$$

Thus, V satisfies the conditions of lemma 3, except it is not continuous. Nevertheless, we still have a sequence $\{w_n\}$ defined by

$$w_0 = 0, \quad w_{n+1} = V[w_n],$$

which satisfies $0 = w_0 \leqslant w_1 \leqslant \ldots \leqslant w_n \leqslant w_{n+1} \leqslant \ldots \leqslant \Gamma(z_1^* - z_0)$.

Hence, by regularity, $w^* = \lim w_n$ exists, and

$$0 \leqslant w^* \leqslant \Gamma(z_1^* - z_0).$$

To show that $w^* = V[w^*]$, note that

$$V[w] = w + F[z_0] + f'(z_0)\Gamma^{-1}w - L\Gamma^{-1}w$$

so,

$$L^{-1}w_{n+1} = L^{-1}w_n + L^{-1}F[z_0] + L^{-1}f'(z_0)\Gamma^{-1}w_n - \Gamma^{-1}w_n.$$

Now, taking limits, as $n \to \infty$, we have

$$L^{-1}w^* = L^{-1}w^* + L^{-1}F[z_0] + L^{-1}f'(z_0)\Gamma^{-1}w^* - \Gamma^{-1}w^*,$$

hence, $w^* = V[w^*]$. But, if we set $z_1 = z_0 + \Gamma^{-1}w^*$, then $z_0 \leqslant z_1 \leqslant z_1^*$ and $F'[z_0](z_1 - z_0) = -F[z_0]$. By the proof of lemma 1, we have $F[z_1] \geq 0$, and so proceeding by induction, we get a monotone increasing bounded sequence $\{z_n\}$. But this sequence satisfies

$$L[z_{n+1}] = f'(z_n)(z_{n+1} - z_n) + f(z_n)$$

or, equivalently,

$$z_{n+1} = L^{-1}[f'(z_n)(z_{n+1} - z_n) + f(z_n)]$$

Let

$$w_n = f'(z_n)(z_{n+1} - z_n) + f(z_n).$$

Then, $w_n \geq g(z_{n+1} - z_n) + f(z_n)$ and, by convexity, $w_n \leq f(z_{n+1})$.

But, also by convexity,

$$f'(z_{n+1})(z_1^* - z_{n+1}) \leq f(z_1^*) - f(z_{n+1})$$

so

$$f(z_{n+1}) \leq f(z_1^*) - f'(z_{n+1})(z_1^* - z_{n+1})$$
$$\leq f(z_1^*) - g(z_1^* - z_{n+1})$$

and

$$f'(z_0^*)(z_n - z_0^*) \leq f(z_n) - f(z_0^*)$$

so

$$f(z_n) \geq f'(z_0^*)(z_n - z_0^*) + f(z_0^*)$$
$$\geq g(z_n - z_0^*) + f(z_0^*).$$

Hence

$$g(z_{n+1} - z_0^*) + f(z_0^*) \leq w_n \leq f(z_1^*) - g(z_1^* - z_{n+1})$$

or

$$-g(z_0^*) + f(z_0^*) \leq w_n - g(z_{n+1}) \leq f(z_1^*) - g(z_1^*).$$

Thus, the set $\{w_n - g(z_{n+1})\}$ is order bounded. But W is normal, so this set is, in fact , bounded. Now, $\{z_n\}$ is order bounded, hence bounded, and g is linear and continuous, so $\{g(z_n)\}$ is bounded. Therefore, $\{w_n\}$ is bounded, since

$$\{w_n\} \subset \{w_n - g(z_{n+1})\} + \{g(z_n)\}$$

L^{-1} is completely continuous, so $\{z_{n+1}\} = L^{-1}\{w_n\}$ is totally bounded. Thus, there is a convergent subsequence $\{z_{n_k}\}$, with $\lim_{k\to\infty} z_{n_k} = z^*$. But since this sequence is monotone, and Z is normal, $z^* = \lim_{n\to\infty} z_n$. Finally, $F[z^*] = 0$, because

$$g(z_{n+1} - z_n) \leq f'(z_n)(z_{n+1} - z_n) \leq f(z_{n+1}) - f(z_n)$$

so clearly, $\lim_{n\to\infty} [f'(z_n)(z_{n+1} - z_n)] = 0$. Applying this to the formula

$$z_{n+1} = L^{-1}[f'(z_n)(z_{n+1} - z_n) + f(z_n)]$$

gives $z^* = L^{-1}[f(z^*)]$, or

$$L[z^*] = f(z^*).$$

Comparing this theorem to theorem 1, the only assumption that has been added, aside from the special form of the equation, is the condition that the operator $\Gamma \geq -F'[z]$ can be written as $\Gamma = L - g$, where $g \leq f'(z)$.

We have already observed that the equation $u'' = -e^{-u}$, with boundary conditions $u(0) = u(1) = 0$, satisfies the hypotheses of this theorem, with $Z = C^2[0,1]$, $W = L^2[0,1]$, $D = \{z : z(0) = z(1) = 0\}$, $L[u] = -u''$, $f(u) = e^{-u}$, $z_0^* = 0$, $z_1^* = \frac{1}{2}x(1-x)$ and $g(h) = -e^{-1/8} h$, where $1/8 = \max z_1^*(x)$. Since we know that the operator $H[h] = -(h'' - \alpha h)$ has a positive continuous inverse, provided $\alpha \geq 0$, clearly the operator $L - g$ has a positive continuous inverse.

As another application, consider the Ricatti equation

$$u'(x) = u^2 + a(x)u + b(x) \quad , \quad u(0) = c.$$

The spaces we use here are $Z = C^1[0,1]$, $W = L^2[0,1]$, $D = \{u : u(0) = c\}$, with the usual topology and ordering. Letting $L[u] = u'$, $f(u) = u^2 + a(x)u + b(x)$, we have L^{-1} is positive and completely continuous. If $b(x) \geq 0$ then $L[0] = 0 \leq f(0) = b(x)$, and if u^* is a positive solution, then in the interval $[0,u^*]$ we have

$$f'(u)h = 2uh + a(x)h \geq a(x)h = g(x)$$

and $(L - g)h = h' - a(x)h$ has the inverse

$$(L - g)^{-1}k(x) = \int_0^x e^{\int_s^x a(t)dt} k(s)ds.$$

Hence $\Gamma^{-1} = (L - g)^{-1}$ is continuous and positive, provided $a(x)$ is integrable on $[0,1]$. Summarizing this, we have

<u>Corollary</u>. If $b(x) \geq 0$, $a(x)$ is integrable, and the equation

$$u'(x) = u^2 + a(x)u + b(x),\ u(0) = c$$

has a positive solution, then this equation has a monotone Newton sequence which converges to a positive solution.

Kalaba [11] has computed a Newton sequence for this equation in the special case $a(x) \equiv 0$, $b(x) \equiv 1$, and his results clearly show the monotone character of the sequence.

As with many of our previous theorems, the hypothesis that f is convex can be replaced by $f'' \geq 0$, provided f is twice uniformly differentiable. The resulting theorem is

<u>Theorem 7</u>. Let $F: D \subset Z \to W$ where Z is normal, W is regular and normal, $F = f - L$, and $[z_0^*, z_1^*] \subset D^0$ where

1) L is linear and has a completely continuous positive inverse;

2) f is continuous, twice uniformly differentiable and $f''(z) \geq 0$ for all z in the interval;

3) $F[z_0^*] \geq 0 \geq F[z^*_1]$;

4) $L - f'(z_0^*) \leq L - g$ where $(L - g)^{-1}$ is positive and continuous.

Then, the conclusions of theorem 6 hold.

An interesting application of this theorem is the following integro-differential equation of Volterra [8];

$$y'(t) = ay(t) + by^2(t) + y(t)\int_0^t K(t,s)y(s)ds,$$

where $a \geq 0$, $b \geq 0$, $K \geq 0$, and $y(0) = y_0 > 0$. Let $L[y] = y'$ and

$$f(y) = ay + by^2 + y\int_0^t K(t,s)y(s)ds.$$

Then

$$f'(y)h = ah + 2byh + y\int_0^t K(t,s)h(s)ds + h\int_0^t K(t,s)y(s)ds$$

and

$$f''(y)hk = 2bkh + k\int_0^t K(t,s)h(s)ds + h\int_0 K(t,s)k(s)ds.$$

Hence $f''(y) \geqslant 0$ for all y. Using the spaces $Z = C^1[0,1]$, $W = L^2[0,1]$, $D = \{ z : z(0) = y_0 \}$, we have $L^{-1}h = \int_0^x h(t)dt$, so L^{-1} is positive and completely continuous, and if $z_0^*(t) = y_0$ then $L - f'(z_0^*) \leqslant (L - a)$ and

$$(L - a)^{-1}h = \int_0^x e^{a(x-s)}h(s)ds \geqslant 0.$$

Finally, it is known that the equation has a positive solution y^*, and if $K(t,s) \geqslant 0$ then $y^* \geqslant z_0^*$, hence we can use the interval $[z_0^*, y^*]$.

To estimate the rate of convergence, we can use theorems 2 and 3, however, because of the special form of these equations, a slightly better result can be proven.

<u>Theorem 8</u>. Let $F:D \subset Z \to W$ where Z and W are PNL spaces, $F = f - L$ where L has a continuous inverse, f is continuous, convex, and continuously G-differentiable. If F has a Newton sequence $\{z_n\}$ which is monotone increasing and converges to a solution z^* of $F[z] = 0$, and if

$$\|L^{-1}\| \cdot \|f'(z_n)\| < 1$$

then

$$\|z_{n+1} - z^*\| \leqslant \frac{\|L^{-1}\|}{1 - \|L^{-1}\| \cdot \|f'(z_n)\|} \cdot r(z_n - z^*) \cdot \|z_n - z^*\|$$

where $r'(z_n - z^*) \to 0$ as $z_n \to z^*$.

If f is twice uniformly differentiable, and

$$m = \max[\|f'(z_n)\|, \|f''(z_n)\|] < \frac{1}{\|L^{-1}\|},$$

then

$$\|z_{n+1} - z^*\| \leqslant \frac{\|L^{-1}\| m}{1 - \|L^{-1}\| m} \|z_n - z^*\|^2$$

<u>Proof</u>.

$$-F'[z_n](z^* - z_{n+1}) = -F'[z_n](z^* - z_n) + F'[z_n](z_{n+1} - z_n),$$

hence,

$$-F'[z_n](z^* - z_{n+1}) = -F'[z_n](z^* - z_n) - F[z_n]$$
$$\leq -F'[z_n](z^* - z_n) + F'[z^*](z^* - z_n)$$

since

$$-F'[z^*](z^* - z_n) \leq F[z_n] - F[z^*] = F[z_n].$$

But

$$-F'[z_n](z^* - z_{n+1}) = L(z^* - z_{n+1}) - f'(z_n)(z^* - z_{n+1})$$

so

$$L(z^* - z_{n+1}) \leq f'(z_n)(z^* - z_{n+1}) + f'(z_n)(z_n - z^*) - f'(z^*)(z_n - z^*)$$

and hence

$$0 \leq z^* - z_{n+1} \leq L^{-1}[f'(z_n)(z^* - z_{n+1}) + [f'(z_n) - f'(z^*)](z_n - z^*)].$$

Since the norm is monotone, this gives

$$\| z^* - z_{n+1} \| \leq \| L^{-1} \| \cdot \left\{ \| f'(z_n) \| \cdot \| z^* - z_{n+1} \| + \| f'(z_n) - f'(z^*) \| \; \| z_n - z^* \| \right\}$$

Setting $r(z_n - z^*) = \| f'(z_n) - f'(z^*) \|$, the first estimate follows easily. If f'' is continuous, then by the integral theorem of chapter I,

$$f'(z^*)(z_n - z^*) - f'(z_n)(z_n - z^*) =$$
$$\int_0^1 f''(z_n + t(z_n - z^*))(z_n - z^*)(z_n - z^*)dt,$$

so

$$\| f'(z^*) - f'(z_n) \| \leq \int_0^1 \| f''(z_n + t(z_n - z^*)) \| \cdot \| z_n - z^* \| dt$$
$$\leq m \| z_n - z^* \|$$

and the last estimate follows easily .

Most of the results of Kalaba [11] are included in the above theorems. (When applied to differential equations, Newton's method is often called "quasi-linearization.") Kalaba has shown that this method can be an effective technique for solving certain types of ordinary and partial differential equations. (See also [5].) It should be noted, however, that several types of equations, which Kalaba considers, cannot be solved using this technique, except in very special circumstances. For example, suppose a Newton sequence for the Dirichlet problem

$$u = f(u,x,y) \ , \ (x,y) \in D$$
$$u(x,y) = 0 \ , \ (x,y) \in \partial D$$

is monotone and bounded. Then the sequence will converge in $L^2[0,1]$. But, without further information, it cannot be concluded that this limit function is a solution to the differential equation. In fact, the limit function for such a problem may not even be continuous. Another type of equation which presents a problem is

$$u''(x) = f(u,u',x)$$
$$u(0) = u(b) = 0.$$

The Newton equations for this problem are

$$u''_{n+1} = f(u_n,u'_n) + \frac{\partial f}{\partial u}(u_n,u'_n)(u_{n+1}-u_n) + \frac{\partial f}{\partial u'}(u_n,u'_n)(u'_{n+1}-u'_n)$$
$$u_{n+1}(0) = u_{n+1}(b) = 0.$$

Hence, even if it can be shown that the sequence $\{u_n\}$ exists, is monotone and bounded, the presence of the derivatives u'_{n+1} and u'_n in these equations prevents us from concluding that the sequence converges to a solution. In the context we have been using, if we write the equation as

$$L[u] = g(u),$$

where $g(u) = f(u,u',x)$, then, in the topology given by the maximum norm, g is neither continuous nor differentiable.

The remarks made in the previous section concerning the convergence of an approximate sequence apply also to these equations. The modification of theorem 6 which corresponds to theorem 5 is;

Theorem 9. In addition to the hypotheses of theorem 6, for every $z \in [z_0^*, z_1^*]$ let g_z be a continuous linear operator which satisfies

$$f'(z) \geq g_z \geq g$$

If $\Gamma_z = L - g_z$ then there exists a monotone increasing sequence $\{z_n\} \subset [z_0^*, z_1^*]$ which satisfies

$$\Gamma_{z_n}(z_{n+1} - z_n) = F[z_n]$$

and which converges to a solution of $F[z] \quad 0$.

Proof. In the proof of theorem 6, replace $-F'[z_n]$ by Γ_{z_n}.

When solving, for example, the equation $u'' = -f(u)$, the Newton sequence is given by the equations

$$z''_{n+1} = f'(z_n)(z_{n+1} - z_n) + f(z_n).$$

If f is convex, then we can replace $f'(z_n)$ by the divided difference

$$\xi_n = \frac{f(z_n) - f(z_{n-1})}{z_n - z_{n-1}}$$

and, by convexity, $f'(z_n) \geq \xi_n$. Hence, if we set $\Gamma_{z_n} h = h'' - \xi_n h$ then the Newton equations have the form

$$z''_{n+1} = \xi_n(z_{n+1} - z_n) + f(z_n).$$

If $f(u)$ has a complicated derivative, it may be much easier to compute ξ_n then to find $f'(z_n)$, however, theorem 9 still ensures the convergence of the sequence to a solution of the equation.

Equations in Spaces with Partially Ordered Norm

As a final application of the preceeding results, we will prove a convergence theorem for Newton's method applied to operators on spaces which are normed by a PTL space. The theorem is an extension of a theorem of Kantorovich ([14], see also [1]). In the following, if $F:X \to Y$ and $Q:Z \to W$, where X and Y are normed by Z and W, respectively, then we write $\|F\| \leq Q$ if $\|F[x]\| \leq Q[z]$ for all $x \in X$, $z \in Z$, such that $\|x\| \leq z$.

The Kantorovich theorem is;

<u>Theorem</u>. Let Z be a regular, solid PN space, W a solid PB space, and let X and Y be normed by Z and W respectively. Assume that X is complete in the Z-norm topology. Let $F:D \subset X \to Y$ satisfy the following;

1a) F is twice continuously G-differentiable in a sphere $S(x_0,r_0) = \{ x : \|x-x_0\| \leq r_0 \} \subset D^0$.

1b) $\Gamma_0 = F'[x_0]$ maps X onto Y, and Γ_0^{-1} is continuous.

Assume there exists an operator $Q:Z \to W$ which is twice continuously G-differentiable on $[z_0, z_1^*]$ where $r_0 \geqslant z_1^* - z_0$, and

2a) $\Delta_0 = -Q'[z_0]$ has a continuous inverse, and $\|\Gamma_0^{-1}\| \leq \Delta_0^{-1}$

2b) $\|\Gamma_0^{-1} F[x_0]\| \leq \Delta_0^{-1} Q[z_0]$;

2c) $\|\Gamma_0^{-1} F''[x]\| \leq \Delta_0^{-1} Q''[z]$ for all x,z with $\|x-x_0\| \leq z-z_0 \leq z_1^* - z_0$;

3) the sequence $\{z_n\}$ defined by $z_{n+1} = z_n - Q'[z_n]^{-1} Q[z_n]$ exists, remains in $[z_0, z_1^*]$, and converges to a solution $z^* \in [z_0, z_1^*]$ of $Q[z] = 0$.

Then, there exists a sequence $\{x_n\}$ in $S(x_0,r_0)$, defined by

$$x_{n+1} = x_n - F'[x_n]^{-1} F[x_n]$$

which converges to a solution x^* of $F[x] = 0$. Furthermore,

$$\|x_n - x^*\| \leq z^* - z_n.$$

In stating the differentiability conditions on F and Q, we have used the fact that theorem 12 of chapter I also holds for these spaces.

Condition 3), in most cases, is very difficult to verify. The theorem which follows modifies conditions 2b), and 2c) slightly, and replaces 3) by a much simpler hypothesis.

The conclusion remains unchanged. Before doing this however, we need some preliminary results. In particular, we must prove a generalization of the well-known Banach theorem:

> If $F:X\to X$, where X is a Banach space, and $\|F\| \leq q < 1$, then $I - F$ has an inverse, with $\|(I - F)^{-1}\| \leq (1-q)^{-1}$

<u>Lemma 4</u>. Let X be normed by a regular and normal PTL space Z, and assume the Z-norm topology is complete. If $F:X\to X$ and $Q:Z\to Z$ are continuous linear operators with $\|F\| \leq Q$, and if $I - Q$ has a positive inverse, then $I - F$ has an inverse, and $\|(I - F)^{-1}\| \leq (I - Q)^{-1}$.

<u>Proof</u>. Let $x \in X$ be arbitrary. For any $z \in Z$ with $\|x\| \leq z$, we have $\|F[x]\| \leq Q[z]$, and by induction, also, $\|F^n[x]\| \leq Q^n[z]$. Let

$$S_n x = x + F[x] + F^2[x] +\ldots+ F^n[x],$$

$$T_n z = z + Q[z] + Q^2[z] +\ldots+ Q^n[z].$$

Then $0 \leq T_n z \leq T_{n+1} z \leq (I - Q)^{-1} z$ because

$$(I - Q)T_{n+1} z = z - Q^{n+2}[z] \leq z,$$

hence,

$$T_{n+1} z \leq (I - Q)^{-1} z.$$

But Z is regular, so $z^* = \lim T_n z$ exists and $z^* \leq (I - Q)^{-1} z$. Furthermore,

$$0 \leq \| S_n x - S_m x \| \leq T_n z - T_m z$$

and the sequence $T_n z - T_m z \to 0$, so by normality, $S_n x - S_m x \to 0$. By completeness, $x^* = \lim S_n x$ exists. In fact,

$$x^* = (I - F)^{-1} x$$

because,

$$\| (I - F)S_n x - x \| = \| F^{n+1} x \| \leq Q^{n+1}[z] \to 0,$$

so $\lim_{n\to\infty} (I - F)S_n x = x$. Since F is continuous, this gives

$(I - F)x^* = x$, and, if $(I - F)x^{**} = x$, then letting $x' = x^{**} - x^*$, we have $\|x'\| = \|F[x']\| \leq Q[\|x'\|]$, hence

$$(I - Q)[\|x'\|] \leq 0$$

so $\|x'\| \leq 0$. Thus $x' = 0$ and $x^{**} = x^*$. Therefore, $(I - F)^{-1}$ is defined on all of X, and

$$\|S_n x\| \leq T_n z \leq z^* \leq (I - Q)^{-1} z$$

so

$$\lim_{n\to\infty} \|S_n x\| \leq (I - Q)^{-1} z,$$

hence

$$\|(I - F)^{-1} x\| \leq (I - Q)^{-1} z.$$

The integral theorem of chapter I will be used in the following form;

Lemma 5. Let $u: [0,1] \to Y$ where Y is normed by a PB space W. If u is uniformly differentiable on [0,1], then

$$\int_0^1 \|u'(t)\| \, dt$$

exists, and

$$\|u(1) - u(0)\| \leq \int_0^1 \|u'(t)\| \, dt.$$

Proof. Using the proof of theorem 13, chapter I, we can show that $\int_0^1 u'(t)dt$ exists, and

$$u(1) - u(0) = \int_0^1 u'(t)dt.$$

Furthermore, the function $v(t) = \|u'(t)\|$ is continuous and maps [0,1] into W, where W is a Banach space. Hence v is integrable. That is, $\int_0^1 \|u'(t)\| \, dt$ exists, and thus

$$\|u(1) - u(0)\| = \left\| \int_0^1 u'(t)dt \right\| \leq \int_0^1 \|u'(t)\| \, dt.$$

Finally, we note that if $F: Z \to W$ is a positive linear operator from a solid PN space Z to a PN space W, then F is continuous because, if $z_n \to z^*$ then, by theorems 5 and 7 of chapter I,

$$-\eta_n u \leq z_n - z^* \leq \eta_n u$$

for some $u \geq 0$, and $\eta_n \to 0$. Hence,

$$-\eta_n F[u] \leq F[z_n] - F[z^*] \leq \eta_n F[u]$$

so, by normality, $F[z_n] \to F[z^*]$.

We can now state and prove the main result of this section.

Theorem 10. Let Z be a regular, solid PN space, W a solid PB space, and let X and Y be normed by Z and W, respectively. Assume that X is complete in the Z-norm topology. Let $F: D \subset X \to Y$ satisfy the following conditions.

1a) F is twice continuously G-differentiable in a sphere $S(x_0, r_0) = \{ x : \| x - x_0 \| \leq r_0 \} \subset D^0$;

1b) $\Gamma_0 = F'[x_0]$ maps X onto Y, and Γ_0^{-1} is continuous.

Assume there exists an operator $Q: Z \to W$ which is twice continuously G-differentiable on $[z_0, z_1^*]$ where $r_0 \geq z_1^* - z_0$, and

2a) $\| \Gamma_0^{-1} \| \leq \Delta_0^{-1}$ where $\Delta_0 = -Q'[z_0]$;

2b) $\| F[x_0] \| \leq Q[z_0]$;

2c) $\| F''[x] \| \leq Q''[z]$ for all x,z with

$$\| x - x_0 \| \leq z - z_0 \leq z_1^* - z_0;$$

2d) $Q'[z]$ is one-to-one, for all $z \in [z_0, z_1^*]$;

2e) $Q[z_1^*] \ll 0$.

Then, there exists Newton sequences $\{x_n\}$, $\{z_n\}$ where

$$x_{n+1} = x_n - F'[x_n]^{-1} F[x_n]$$

$$z_{n+1} = z_n - Q'[z_n]^{-1} Q[z_n]$$

which converge to x^*, z^* respectively, where $x^* \in S(x_0, r_0)$ $z^* \in [z_0, z_1^*]$, $F[x^*] = 0$, $Q[z^*] = 0$, and

$$\| x^* - x_n \| \leq z^* - z_n.$$

Proof. Q is twice uniformly differentiable, $Q'' \geq 0$, $-Q'[z]$ is inverse positive, $-Q'[z_0]$ has a positive inverse which is therefore continuous, and $Q[z_0] \geq 0 \gg Q[z_1^*]$. Hence, by an obvious corollary to theorem 4, Q has a Newton sequence at z_0 which converges monotonically to a solution

$z^* \in [z_0, z_1^*]$ of $Q[z] = 0$. Furthermore, $Q'[z_n]$ has a positive (hence continuous) inverse. Now,

$$x_1 = x_0 - \Gamma_0^{-1} F[x_0]$$

exists, and

$$\| x_1 - x_0 \| = \| \Gamma_0^{-1} F[x_0] \| \leq \Delta_0^{-1} Q[z_0] = z_1 - z_0 \leq z_1^* - z_0 .$$

So $x_1 \in S(x_0, r_0)$. We will show that all of the hypotheses of the theorem are satisfied if x_0 and z_0 are replaced by x_1 and z_1. First we prove

2b') $\| F[x_1] \| \leq Q[z_1$

Let

$u(t) = F[x_0 + t(x_1 - x_0)] + (1-t)F'[x_0 + t(x_1 - x_0)](x_1 - x_0)$.

Then

$$\begin{aligned}
(t) &= F'[x_0 + t(x_1 - x_0)](x_1 - x_0) + \\
&\qquad + (1-t)F''[x_0 + t(x_1 - x_0)](x_1 - x_0)(x_1 - x_0) - \\
&\qquad - F'[x_0 + t(x_1 - x_0)](x_1 - x_0) \\
&= (1-t)F''[x_0 + t(x_1 - x_0)](x_1 - x_0)(x_1 - x_0).
\end{aligned}$$

By lemma 5,

$$\| u(1) - u(0) \| \leq \int_0^1 \| u'(t) \| \, dt,$$

so

$$\begin{aligned}
&\| F[x_1] - F[x_0] - F'[x_0](x_1 - x_0) \| \leq \\
&\qquad \leq \int_0^1 \| (1-t)F''[x_0 + t(x_1 - x_0)](x_1 - x_0)(x_1 - x_0) dt \| .
\end{aligned}$$

But $F'[x_0](x_1 - x_0) = -F[x_0]$, so

$$\begin{aligned}
\| F[x_1] \| &\leq \int_0^1 (1-t) \| F''[x_0 + t(x_1 - x_0)](x_1 - x_0)(x_1 - x_0) \| \, dt \\
&\leq \int_0^1 (1-t) Q''[z_0 + t(z_1 - z_0)](z_1 - z_0)(z_1 - z_0) dt \\
&= Q[z_1] - Q[z_0] - Q'[z_0](z_1 - z_0) \\
&= Q[z_1].
\end{aligned}$$

Next we show that $F'[x_1]$ has a continuous inverse, and

2a') $\|\Gamma_1^{-1}\| \leq \Delta_1^{-1}$

Let $u(t) = -F'[x_0 + t(x_1 - x_0)]h$, where $h \in X$ is arbitrary. Then $u:[0,1] \to Y$, u is continuously differentiable, and lemma 5 holds. Hence, if $\|h\| \leq k$,

$$\begin{aligned}\|(F'[x_0] - F'[x_1])h\| &\leq \int_0^1 \|F''[x_0 + t(x_1 - x_0)]h(x_1 - x_0)\|\,dt \\ &\leq \int_0^1 Q''[z_0 + t(z_1 - z_0)]k(z_1 - z_0)\,dt \\ &= Q'[z_1]k - Q'[z_0]k \\ &= (\Delta_0 - \Delta_1)k,\end{aligned}$$

where $\Delta_1 = -Q'[z_1]$. Thus, by 2a)

$$\|\Gamma_0^{-1}(\Gamma_0 - \Gamma_1)\| \leq \Delta_0^{-1}(\Delta_0 - \Delta_1)$$

where $\Gamma_1 = F'[x_1]$. Now, $Q''[z] \geq 0$, so $\Delta_1 \leq \Delta_0$, and so $\Delta_1^{-1}\Delta_0 \geq 0$. Let $G = \Gamma_0^{-1}(\Gamma_0 - \Gamma_1)$, $P = \Delta_0^{-1}(\Delta_0 - \Delta_1)$. Then $G:X \to X$, $P:Z \to Z$, $\|G\| \leq P$, and $(I - P)^{-1} = \Delta_1^{-1}\Delta_0 \geq 0$. So by lemma 4, $(I - G)^{-1}$ exists and is continuous. Also

(2) $$\|(I - G)^{-1}\| \leq (I - P)^{-1} = \Delta_1^{-1}\Delta_0$$

But,

$$(I - G)^{-1} = (\Gamma_0^{-1}\Gamma_1)^{-1} = \Gamma_1^{-1}\Gamma_0$$

so Γ_1^{-1} exists and is continuous. We still have to show

$$\|\Gamma_1^{-1}\| \leq \Delta_1^{-1}$$

or, equivalently,

$$\|\Gamma_1^{-1}y\| \leq \Delta_1^{-1}w$$

for all y, w with $\|y\| \leq w$. Let $y = \Gamma_0 x$. Then if $\|y\| \leq w$, 2a) implies $\|\Gamma_0^{-1}y\| \leq \Delta_0^{-1}w$. That is $\|x\| \leq \Delta_0^{-1}w$. Thus, by

(2) $$\|\Gamma_1^{-1}\Gamma_0 x\| \leq \Delta_1^{-1}\Delta_0(\Delta_0^{-1}w)$$

or $$\|\Gamma_1^{-1}y\| \leq \Delta_1^{-1}w$$

By induction then, we have proven that $\{x_n\}$ exists and

$$\|x_{n+1} - x_n\| \leq z_{n+1} - z_n,$$

therefore,

$$\|x_{n+p} - x_n\| \leq z_{n+p} - z_n.$$

As $n \to \infty$, $z_{n+p} - z_n \to 0$, so $\| x_{n+p} - x_n \| \to 0$. Since the topology in X is complete, $x^* = \lim x_n$ exists. Finally, $F[x^*] = 0$ because, $0 \leq \| F[x_n] \| \leq Q[z_n]$ and $Q[z_n] \to 0$ so $F[x_n] \to 0$. Since F is continuous, $0 = \lim F[x_n] = F[x^*]$

Note that in the important case where $Z = W = E^n$, condition 2d) can be eliminated. In this case, the only major addition to the Kantorovich hypotheses is 2e). However, implicit in condition 3) of the Kantorovich theorem is the existence of a point $z^* \in [z_0, z_1^*]$ such that $Q[z^*]=0$, and 2e) is only a strengthening of this.

REFERENCES

1. P.M. Anselone (ed.), Nonlinear Integral Equations, Univ. of Wisconsin, Madison, Wisc., 1964.

2. A.N. Baluev, On the Method of Chaplygin, D.A.N., 83 (1952), 781-784.

3. __________, On the Method of Chaplygin for Functional Equations, Vest. Leng., 13 (1956), 27-42.

4. __________, Applications of Semi-ordered Norms in the Approximate Solution of Nonlinear Equations, Vest. Leng., 271 (1958), 18-27.

5. E.F. Beckenbach and R. Bellman, Inequalities, 2nd rev. ed., Springer-Verlag, New York, 1965.

6. G. Birkhoff, Lattice Theory, rev. ed., Am. Math. Soc. Colloq. Publ. 25, Am. Math. Soc., New York, 1948.

7. L. Collatz, Funktionalanalysis und numerische Mathematik, Springer-Verlag, Berlin, 1965.

8. H.T. Davis, Introduction to Nonlinear Differential and Integral Equations, U.S. AEC, Wash., D.C., 1960.

9. H. Gordon, Relative Uniform Convergence, Math. Annalen, 153 (1964), 418-427.

10. D. Greenspan and S. Parter, Mildly Nonlinear Elliptic Partial Differential Equations and their Numerical Solution II, Num. Math., 7, no. 2 (1965), 129-147.

11. R. Kalaba, On Nonlinear Differential Equations, the Maximum Operation, and Monotone Convergence, J. Math. Mech., 8, no. 4 (1959), 519-574.

12. L.V. Kantorovich, The Method of Successive Approximations for Functional Equations, Acta. Math., 71 (1939), 63-77.

13. L.V. Kantorovich, Functional Analysis and Applied Mathematics, Uspehi. Mat. Nauk., 3, no. 6 (1948), 89-185, transl. C.D. Benster, National Bur. Stand., 1952.

14. ________________, The Majorant Principle and Newton's Method, D.A.N., 76 (1951), 17-20.

15. ________________, Some Further Applications of Newton's Method, Vest. Leng., 2, no. 7 (1957), 68-103.

16. ________________ and G.P. Akilov, Functional Analysis in Normed Spaces, transl. D.E. Brown, A.P. Robertson, MacMillan, New York, 1964.

17. J.L. Kelley, I. Namioka, and co-authors, Linear Topological Spaces, Van Nostrand, Princeton, 1963.

18. M.A. Krasnoselski, Positive Solutions of Operator Equations, transl. R.E. Flaherty, Noordhoff, Groningen, 1964.

19. S. Lang, Introduction to Differentiable Manifolds, Interscience, New York, 1962.

20. I. Namioka, Partially Ordered Linear Topological Spaces, Memoirs Am. Math. Soc., 24, Am. Math. Soc., New York, 1957.

21. H. Schaefer, Halbgeordnete lokalkonvexe Vektorraume, Math. Annalen, 135 (1958), 115-141.

22. J.W. Schmidt, Konvergenzuntersuchungen und Fehlerabschatzungen fur ein verallgemeinertes Iterationsverfahren, Arch. Rat. Mech. Analy., 6, no. 4 (1960), 261-276.

23. J. Schroeder, Monotonie- Eigenschaften bei Differentialgleichungen, Arch. Rat. Mech. Analy., 14 (1963), 38-60.

24. S.N. Slugin, Approximate Solution of Operator Equations by the Method of Chaplygin, D.A.N., 103 (1955), 565-568.

25. S.N. Slugin, On the Theory of the Methods of Newton and Chaplygin, D.A.N., 120, no. 3 (1958), 472-474.

26. F.G. Tricomi, Integral Equations, Interscience, New York, 1957.

27. M.M. Vainberg, Variational Methods for the Study of Nonlinear Operators, transl. A. Feinstein, Holden-Day, San Francisco, 1964.