QUALITATIVE METHODS IN NONLINEAR MECHANICS

J. Tinsley Oden

QUALITATIVE METHODS IN NONLINEAR MECHANICS

J. Tinsley Oden
The University of Texas at Austin

Prentice-Hall, Inc., Englewood Cliffs, New Jersey 07632

Library of Congress Catalog Card Number: 85-61486
ISBN: 0-13-744996-8

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

ISBN 0-13-744996-8 01

Prentice-Hall International (UK) Limited, *London*
Prentice-Hall of Australia Pty. Limited, *Sydney*
Prentice-Hall Canada Inc., *Toronto*
Prentice-Hall Hispanoamericana, S.A., *Mexico*
Prentice-Hall of India Private Limited, *New Delhi*
Prentice-Hall of Japan, Inc., *Tokyo*
Prentice-Hall of Southeast Asia Pte. Ltd., *Singapore*
Editora Prentice-Hall do Brasil, Ltda., *Rio de Janeiro*
Whitehall Books Limited, *Wellington, New Zealand*

PREFACE

This book was the outgrowth of class notes I prepared over a period of several years for a graduate course for students of mechanics and mathematics, and occasionally, physics, on nonlinear boundary-value problems in mechanics. The aim of the course was to present a systematic introduction to the mathematical foundations of the theory of nonlinear boundary-value problems which focused on qualitative features of the subject, e.g., on questions of existence, uniqueness, and regularity of solutions, and which approached this subject in a way accessible to the student equipped with some graduate level mechanics and introductory functional analysis. While the emphasis here is on qualitative properties of nonlinear problems, many of the basic theorems are, to a great extent, constructive, so that the theory allows one not only to derive useful information on nonlinear operators, but also to identify in the proofs effective ways to construct solutions. In particular, I have found the material collected here to be extremely valuable in developing approximation theories and computational methods for nonlinear problems in mechanics.

The book is divided into three major parts. Part I - Optimization Theory and Variational Methods - contains treatments of convex and non-convex optimization, saddle point problems, duality, relaxation, and penalty methods, and some applications to contact problems in elastostatics and finite elasticity. Part II - Nonlinear Operator Theory - provides an introduction to monotone and pseudomonotone operators and variational inequalities (the treatment of this latter topic being essentially a revision of my expository paper with N. Kikuchi on this subject which was published in the International Journal of Engineering Science, and I acknowledge permission of Pergamon Press to use portions of that work). Part III - Local Analysis - is a brief introduction to degree theory, classical bifurcation theory and some generalizations to problems of bifurcation from a simple eigenvalue, and a very brief look at nonlinear eigenvalue problems. The overall scope of the book has been intentionally limited so as to provide material for a one-semester course.

During the writing of this book, I have been fortunate to have had the help of many students and colleagues in proofreading various versions of the notes. I wish to express special thanks to Noboru Kikuchi, Luis de Campos and Leszek Demkowicz for reading portions of the work and making helpful suggestions, to Phillip Ciarlet for his comments on a late draft of this work, and to Ruth Dye for typing early versions of the notes. I also thank Ralph Showalter, friend and colleague, who introduced me to several of the subjects dealt with here and who has frequently shared with me his exceptional knowledge of differential equations and analysis. Finally, I wish to register a special note of gratitude to Linda Calvin who so expertly and professionally prepared the final copy of the manuscript for publication.

J. Tinsley Oden
Austin

Table of Contents

PART I

OPTIMIZATION THEORY AND VARIATIONAL METHODS

CHAPTER 1. PRELIMINARIES TO OPTIMIZATION THEORY

1. Introduction

Throughout Chapter 1 of this monograph, we will be concerned with the classical optimization problem: Given a functional F defined on a subset K of a metric space V, find u in K for which $F(u)$ is the smallest possible value F can assume over all of K. The element u is thus a minimizer of F on K. Of course, the functional F should be bounded below on K for if a lower bound doesn't exist ($\inf F = -\infty$) on the set K, then the minimization problem would have little meaning. Being bounded below, the infimum $\mu = \inf\{F(v) \mid v \in K\}$ is finite, but, in general, we cannot conclude that there actually exists a $u \in K$ such that $\mu = F(u)$. Our minimization problem can thus be stated as follows:

$$\text{Find } u \in K \text{ such that } F(u) = \inf_{v \in K} F(v) \tag{1.1}$$

We will not be as much concerned with actually identifying the solution u of problem (1.1) as with establishing conditions on F and K sufficient to guarantee that a solution exists. These conditions are important not only for determining if the problem itself is a meaningful one but also in exposing other properties of the solutions when they exist and for designing methods to calculate them. We are also interested in the characterization of solutions to (1.1) - a subject which provides the connection between (1.1) and variational calculus. These ideas are of great importance in the study of nonlinear partial differential equations and we will describe concrete applications in later chapters.

The famous theorem of Weierstrass provides a classic solution to problem (1.1): If F is continuous on a compact subset K of V, then F achieves its minimum on K. These conditions on F and K are far too strong for this result to be of much value in the solution of interesting optimization problems. We will develop extensions of this classical theorem in which virtually all of its conditions are weakened. In place of real-valued functionals, we will frequently allow F to take on the value $+\infty$; in place of continuity, we will use the concept of lower semicontinuity, and, when V is a Banach space, we will consider cases in which K is only weakly sequentially closed and F is weakly lower semicontinuous. All of these terms are defined explicitly in subsequent sections. Finally, we will consider cases in which no solution at all exists to problem (1.1) but from which we can still extract useful information by constructing, in a natural way, approximations to (1.1).

Our aim in the present chapter is to lay down some essential mathematical preliminaries to our study. Many of the ideas surveyed here are standard in optimization theory and additional details can be found elsewhere; in particular, see EKELAND and TEMAM [26] and ROCKAFELLAR [74], and other references therein.

2. Extended Real-Valued Functionals

Optimization theory can be considerably enriched by allowing functionals defined on a space V to take on the values $\pm\infty$. Thus, we will frequently consider minimization problems in which $F: V \to \overline{\mathbb{R}}$, where $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$ is the extended real line.

There are several advantages to allowing infinite-valued functionals into our theory, at least from the point of view of simplicity in the formulation of certain minimization problems. For instance, if we are given a function $F_0: K \to \mathbb{R}$, we can always construct a new function F according to

$$F(v) = \begin{cases} F_0(v) & v \in K \\ +\infty & v \notin K \end{cases} \tag{2.1}$$

which is defined on *all* of V. Thus,

$$\inf_{v \in K} F_0(v) = \inf_{v \in V} F(v)$$

This means that, by allowing functions to assume the value $+\infty$, we make it necessary to consider only functions defined everywhere on V. If $K \neq \varnothing$, then only the elements in K are truly candidates for minimizers, whereas if $K = \varnothing$, then we can be certain that the minimization problem is infeasible.

Another way of expressing the problem of minimizing a functional $F_1 : V \to \overline{\mathbb{R}}$ on a subset $K \subset V$ is to introduce the *indicator function* ψ_K for the set K:

$$\psi_K(v) = \begin{cases} 0 & \text{if } v \in K \\ +\infty & \text{if } v \notin K \end{cases} \tag{2.2}$$

Then we construct the functional $F : V \to \overline{\mathbb{R}}$ given by

$$F(v) = F_1(v) + \psi_K(v), \quad v \in V \tag{2.3}$$

and, again, seek minimizers of F throughout all of V.

Of course, the use of devices such as (2.1) and (2.2) does not avoid the necessity of considering the constraint set K. However, many aspects of the problem depending on properties of F can be investigated at certain stages of the analysis without concentrating on details having to do with K since the constraint is effectively built into the definition of F.

The minimization of any extended real-valued functional $F : V \to \overline{\mathbb{R}}$ is equivalent to the problem of seeking minima in the *effective domain* of F

$$\text{dom}(F) = \{v \in V \mid F(v) < +\infty\} \tag{2.4}$$

The constraints in any minimization problem are thus implicit in the requirement that $u \in \text{dom}(F)$. Of course, if $\text{dom}(F) = \varnothing$ (i.e., $F(v) \equiv +\infty \quad v \in V$), then the minimization problem is meaningless. Also, if for some $u \in \text{dom}(F)$, $F(u) = -\infty$, then there cannot be a solution to the minimization problem. There is, at least, some hope of solving a minimization problem if we place some mild restrictions on F. For this purpose, we shall say that $F : V \to \overline{\mathbb{R}}$ is *proper* if and only if

(i) it nowhere takes the value $-\infty$ and

(ii) $F(v) \not\equiv +\infty$; i.e. there exists at least one point $u \in V$ at which $F(u) < +\infty$.

We emphasize that an optimization problem may not have a solution even though $F(v) \neq -\infty$ for any v. Consider the functional $F : \mathbb{R} \to \overline{\mathbb{R}}$ defined by

$$F(x) = \begin{cases} +\infty & \text{if } x \leq 0 \\ 1/x & \text{if } x > 0 \end{cases}$$

Then $\inf F(x) = 0$, but there is no point in $\text{dom}(F)$ at which $F(x) = 0$.

3. Convex Sets, Hyperplanes, and Cones

Let V be a linear vector space over $\mathbb{R}$. If u and v are two points in V, then the set of points

$$\{w \in V \mid w = \theta u + (1-\theta) v, \ \theta \in \mathbb{R}, \ 0 \leq \theta \leq 1\} \tag{3.1}$$

is called the *line segment* between u and v, and u and v are called the *endpoints* of this line segment.

Convex Set. A subset $K \subset V$ is said to be *convex* if and only if it contains the line segment between any two of its elements.

It is easily verified that $K \subset V$ is convex if and only if the convex combination of any finite subset of elements $u_1, u_2, \ldots, u_n \in K$ is in K; i.e. if and only if

$$\sum_{i=1}^{n} \lambda_i u_i \in K, \lambda_i > 0, \sum_{i=1}^{n} \lambda_i = 1$$

The entire space V is convex and, by convention, so is the empty set $K = \varnothing$. The intersection of convex sets is convex, but the union of convex sets is, in general, not convex. □

If K is any subset of V, convex or not, the intersection of all convex sets containing K is convex. It is called the *convex hull* of K and denoted $\text{co}(K)$.

The notion of a convex set is the first of a large collection of geometrical abstractions which are very useful in optimization theory. For instance, the set P of points $v \in V$ such that

$$l(v) = \lambda, \ \lambda \in \mathbb{R}$$

where l is a nonzero linear functional on V, is called an *affine hyperplane* and the sets

$$\{v \in V \mid l(v) < \lambda\}, \ \{v \in V \mid l(v) > \lambda\}$$

are *open half spaces* bounded by P. The sets

$$\{v \in V \mid l(v) \leqslant \lambda\}, \ \{v \in V \mid l(v) \geqslant \lambda\}$$

are *closed half spaces* bounded by P.

As another important example of geometrical notions in analysis, we mention the *geometrical form of the Hahn-Banach theorem*. Recall that the *analytical form* of this theorem establishes the following fact:

Let j denote a sublinear functional defined on a real topological vector space V (i.e., $j: V \to \mathbb{R}$ is such that

$$j(\lambda v) = \lambda j(v), \ \lambda \in \mathbb{R}, \ \lambda \geqslant 0, \text{ and } j(u+v) \leqslant j(u) + j(v)$$

and let l denote a linear functional defined on a linear subspace M of V, such that

$$l(v) \leqslant j(v) \quad \forall v \in M.$$

Then there exists a linear functional L defined on all of V, such that

(i) L is an extension of l (i.e., $L(v) = l(v) \ \ \forall v \in M$) and

(ii) $L(v) \leqslant j(v) \ \ \forall v \in V$.

To construct a geometrical version of this result, we note that any nonempty convex set K containing the origin 0 can be completely characterized by a special sublinear functional m_K called the Minkowski functional of K in V and defined by

$$m_K(v) = \inf\{\lambda \mid \lambda^{-1} v \in K, \ \lambda > 0\}$$

One can show, in fact, that

$$0 \leqslant m_K(v) < +\infty \ \text{ and } m_K \text{ is sublinear}$$

and, moreover, that

$$K = \{v \in V \mid m_K(v) \leqslant 1\}; \ \text{int}\, K = \{v \in V \mid m_K(v) < 1\}$$

With this characterization of convex K in mind, consider an affine subspace M of V which does not intersect K and suppose K is open. Since $0 \in M$, by an appropriate scaling we can characterize M by the hyperplane $\{v \in V \mid l(v) = 1\}$ where l is a real linear functional defined on a linear subspace of V. But if $v \in \text{int}K$, then $m_K(v) \geqslant 1$; hence

$$l(v) = 1 \leqslant m_K(v) \quad v \in M$$

By the Hahn-Banach theorem, there exists an extension L of l to all of V, such that

$$L(v) = l(v) \leq m_K(v), \quad v \in M$$

and

$$L(v) \leq m_K(v) \text{ for all } v \in V$$

But, by appropriate scaling, L defines the affine hyperplane $P = \{v \in V \mid L(v) = 1\}$ which contains the hyperplane $M = \{v \in W \mid l(v) = 1\}$ (W being the subspace on which l is defined) and which, by construction, does not intersect int K. Finally, we note that these same arguments apply to any nonempty convex set $K \subset V$ since we can always assume that $0 \in K$ after an appropriate translation.

We summarize these observations in the following basic theorem:

Theorem 3.1 (The Geometrical Form of the Hahn-Banach Theorem). Let V be a real topological vector space. Let K be an open nonempty convex subset of V and M a nonempty affine subspace of V which does not intersect K. Then there exists a closed affine hyperplane P which contains M and does not intersect K. □

The importance of these ideas for our immediate purposes lies in an important corollary to the above theorem on the separation of convex sets.

Corollary 3.1.1. (i) If K and M are disjoint, nonempty, convex subsets of a real topological vector space V and if K is open, then there exists an affine hyperplane P which separates K and M (i.e., P is such that K and M lie in opposite open half spaces defined by P).

(ii) If V is locally convex (i.e. if there exists a fundamental system of convex neighborhoods of the origin of V) and if K is convex and closed and M is convex and compact, K and M disjoint, then there exists an affine hyperplane P which strictly separates K and M (i.e. P separates K and M and the points of P are not in K or M). □

From these results it also follows that for any convex subset K of a real topological vector space V which contains interior points, and any point u which is not an interior point of K, there exists a closed affine hyperplane P which contains u and which is such that K is completely contained in one of the closed half spaces determined by P. If $u \in \bar{K}$, then we say that P is a *supporting hyperplane* of K and that u is a *supporting point* of K (see Fig. 3.1).

One consequence of these results is that in locally convex Hausdorff topological spaces, we are guaranteed the existence of nonzero continuous linear functionals on V. In particular, if u and w are distinct points of V, the Hahn-Banach theorem (or Corollary 3.1.1) allows us to separate them by a closed affine hyperplane, $P: l(v) = \lambda$. The nonzero linear form l is continuous because P is closed and $l(u) \neq l(w)$. Thus, there exist nonzero continuous linear functionals on V. The space V^* of all continuous linear functionals on V is the topological dual of V. Thus, if $v^* \in V^*$, its value $v^*(v)$ at a point v in V is a real number which depends linearly and continuously on v. Conversely, for fixed v, $v^*(v)$ defines a linear functional on V^*. To emphasize this symmetry in the roles of V and V^*, we use the notation

$$v^*(v) = \langle v^*, v \rangle$$

Thus, $\langle \cdot,\cdot \rangle : V^* \times V \to \mathbf{R}$. Every affine hyperplane on such a space is, thus, representable in the form

$$\{w \in V \mid \langle v^*, w \rangle = \lambda, \; v^* \in V^*, \; \lambda \in \mathbf{R}\} \tag{3.2}$$

and every affine continuous functional $l: V \to \mathbf{R}$ is of the form

$$l(v) = \langle v^*, v \rangle + \lambda \tag{3.3}$$

with $\lambda \in \mathbf{R}$, $v^* \in V^*$.

Cones. Another important geometrical concept is the notion of a cone in a linear vector space V. A subset C of a linear vector space is a *cone* if and only if $\lambda C \subset C$ for $\lambda > 0$; C is a

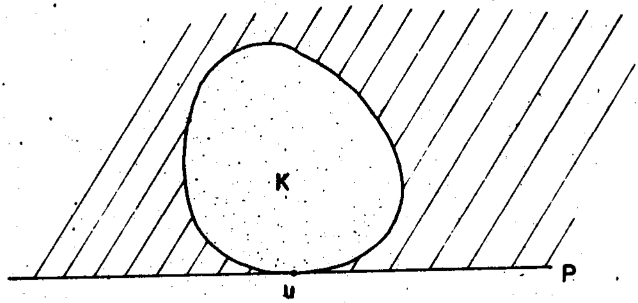

Figure 3.1 Supporting point u in a supporting hyperplane P of a convex set K.

cone with vertex at the origin if and only if $u \in C$ implies $\lambda u \in C$ for $\lambda \geqslant 0$. A cone with vertex $u_0 \in V$ is defined as the translation $u_0 + C$ of a cone C with vertex at the origin. We note that any linear subspace of V and any translated linear subspace of V is a cone.

A cone C of a linear vector space V or $\mathbb{R}$ is said to be *proper* if and only if

(i) C is a cone with vertex at the origin

(ii) C is convex

For such cones, we shall also assume that $C \cap -C = \{0\}$.

The importance of a proper cone in V is that it permits us to introduce an ordering on elements of the linear space V. In particular, if C is a proper cone in V, we introduce a relation R on V defined by

$$uRv \iff u-v \in C$$

The relation R is reflexive ($u-u = 0 \in C$), antisymmetric (uRv and $vRu \Rightarrow v = u$), and transitive (since C is convex, $u_1, u_2 \in C \Rightarrow u_1 + u_2 \in C$; hence, uRv and $vRw \Rightarrow ((u-v) + (v-w)) \in C \Rightarrow uRw$). Thus, R is a total ordering on V and we use the conventional notation $\geqslant$ to describe R; i.e.

$$u \geqslant v \iff u-v \in C \tag{3.4}$$

We refer to C as the *positive cone* in V corresponding to the ordering $\geqslant$. Likewise, the cone $C^- = -C$ is the *negative cone* in V and we write

$$u \leqslant v \iff u-v \in C^- \tag{3.5}$$

It is evident that the specification of a positive cone in V makes it possible to add to the purely linear structure of V inequalities such as $u \geqslant 0$, $u \leqslant 0$, etc. In particular, the partial ordering (3.4) is said to be *compatible with the structure of a linear space* since

$$\begin{aligned} u \geqslant 0 &\Rightarrow \lambda u \geqslant 0 \quad \forall \lambda > 0 \\ u \geqslant v &\Rightarrow u+w \geqslant v+w \quad \forall w \in V \end{aligned} \tag{3.6}$$

Conversely, when an ordering $\geqslant$ is defined on V satisfying (3.6), V becomes an *ordered vector space* and the set $C = \{v \mid v \geqslant 0\}$ is the positive cone in V corresponding to the ordering $\geqslant$; it is a proper cone in V.

If V is normed linear space with positive cone C, a corresponding positive cone C^* can be defined in the dual space V^* by

$$C^* = \{u^* \in V^* \mid \langle u^*, v\rangle \geqslant 0 \quad \forall v \in C\} \tag{3.7}$$

It is not difficult to show that C^* is closed even if C is not. If C is closed, then C and C^* are related as follows:

Theorem 3.2. Let a positive cone C in a normed linear space V be closed. If $v \in V$ is such that

$$\langle u^*, v\rangle \geqslant 0 \quad \forall u^* \in C^*$$

then

$$v \geqslant 0$$

Proof: Assume $u \notin C$, C closed, but $\langle u^*, u\rangle \geqslant 0$, $u^* \in C^*$. Then, by Corollary 3.1.1 (ii), there is a hyperplane strictly separating u and C (since C is convex); i.e. there is a $u^* \in V^*$ such that $\langle u^*, u\rangle < \langle u^*, v\rangle$ for all $v \in C$. Since C is a cone, $\langle u^*, v\rangle$ cannot be negative, because then $\langle u^*, \lambda v\rangle < \langle u^*, u\rangle$ for some $\lambda > 0$. Therefore, $u^* \in C^*$. Finally, since $\inf\{\langle u^*, v\rangle \mid v \in C\} = 0$, we must have $\langle u^*, u\rangle < 0$. □

The set $C^* \subset V^*$ is a positive cone in V^*; it is a proper cone and, therefore, defines an ordering $\leqslant$ on V^* according to

$$u^* \geqslant v^* \iff u^* - v^* \in C^* \tag{3.8}$$

4. Convex Functionals

Let $F: V \to \overline{\mathbb{R}}$ denote a functional defined on a linear vector space V. We recall that the graph of F is the subset of $V \times \overline{\mathbb{R}}$ defined by

$$\text{graph } F = \{(v,\lambda) \in V \times \overline{\mathbb{R}} \mid \lambda = F(v)\} \tag{4.1}$$

the set of points above the graph of F is called the *epigraph* of F and is denoted epi F:

$$\text{epi } F = \{(v,\lambda) \in V \times \overline{\mathbb{R}} \mid \lambda \geqslant F(v)\} \tag{4.2}$$

Convex Functional. A functional $F: V \to \overline{\mathbb{R}}$ is said to be *convex* if and only if its epigraph is convex. □

The fact that this definition coincides with our usual notion of a convex function is established in the following theorem:

Theorem 4.1. Let K be a nonempty convex subset of a linear vector space V and let $F: K \to \overline{\mathbb{R}}$ be a functional defined on K. Then F is convex if and only if, for every u and v in K,

$$F(\theta u + (1-\theta) v) \leqslant \theta F(u) + (1-\theta) F(v) \tag{4.3}$$

for all $\theta \in \overline{\mathbb{R}}$ satisfying $0 \leqslant \theta \leqslant 1$, whenever the right-hand side is defined.

Proof: Suppose (4.3) holds and (u,λ) and (v,μ) are in epi F. Then $F(u) \leqslant \lambda \leqslant +\infty$ and $F(v) \leqslant \mu < +\infty$, by definition. Therefore, for any $\theta \in [0,1]$,

$$F(\theta u + (1-\theta) v) \leqslant \theta F(u) + (1-\theta) F(v) \leqslant \theta\lambda + (1-\theta)\mu$$

But this means that $\theta(u,\lambda) + (1-\theta)(v,\mu) \in \text{epi } F$; i.e. epi F is convex.

Conversely, suppose that epi F is convex. For $u, v \in \text{dom}(F)$, let $\lambda \geqslant F(u)$ and $\mu \geqslant F(v)$. By the convexity of epi F, $\theta(u,\lambda) + (1-\theta)(v,\mu) \in \text{epi} F$ for any $\theta \in [0,1]$. Hence,

$$F(\theta u + (1-\theta) v) \leqslant \theta\lambda + (1-\theta)\mu$$

If $F(u)$ and $F(v)$ are finite, we may take $\lambda = F(u)$ and $\mu = F(v)$ to obtain (4.3). If either $F(u)$ or $F(v) = -\infty$, we allow λ and μ to tend to $-\infty$ to obtain (4.3). □

If the strict inequality ($<$) in (4.3) holds, we say that F is *strictly convex*. If $-F$ is convex, we say that F is *concave*.

The following properties of convex functionals are easily verified:

(i) If $F: V \to \overline{\mathbb{R}}$ is convex, then λF is convex, $\lambda \in \mathbb{R}$, provided $\lambda \geqslant 0$.

(ii) If $F: V \to \overline{\mathbb{R}}$ is convex and $G: V \to \overline{\mathbb{R}}$ is convex, then $F + G$ is convex.

(iii) Let $F = \{F_i\}_{i \in I}$ be a family of convex functions from V into $\overline{\mathbb{R}}$. Then their *pointwise supremum* F, where $F(v) = \sup_{i \in I} F_i(v) \ \forall\, v \in V$, is convex.

(iv) If, in a neighborhood $N(u_0)$ of a point $u_0 \in V$, a convex function F is bounded above by a positive finite constant, then F is continuous at u_0.

Note that (ii) holds with the convention that $(F+G)(v) = +\infty$ if $F(v) = -G(v) = +\infty$. Note that (i) infers that the set $C(V)$ of all convex functions on V is a convex cone with vertex at the origin. For introductions to these related ideas, see LUENBERGER [56] or EKELAND and TEMAM [26].

5. Limits Superior and Inferior

Because of their importance in subsequent discussions, we will review briefly the concepts of limit superior and limit inferior of a sequence of real (or extended real) numbers.

We recall that the elementary notion of the convergence of a sequence $\{x_n\}$ of real numbers to a limit x^* describes the situation in which, for any given ϵ-neighborhood of x^*, there is a positive integer N such that each x_n is in this neighborhood for all $n \geqslant N$. Thus, if $x^* = \lim_{n\to\infty} x_n$, then, for any given $\epsilon > 0$, all but a finite number of terms of the sequence are within ϵ of x^*. This also means that if $\{x_n\}$ converges to x^* there are infinitely many terms within a distance ϵ of x^*.

The converse of this last statement is false. For given $\epsilon > 0$, a sequence $\{x_n\}$ may have infinitely many terms such that $|x_n - x^*| < \epsilon$, but $\{x_n\}$ may not necessarily converge to x^* (the key issue is the existence of an N such that $|x_n - x^*| < \epsilon$ *for all* $n \geqslant N$). For instance, the divergent sequence $x_n = (-1)^n$ has an infinite number of entries in any neighborhood of 1 and -1. In such cases, we say that x^* is a *cluster point* of the sequence $\{x_n\}$; i.e., x^* is a cluster point of $\{x_n\}$ if and only if, for given $\epsilon > 0$ and given N, there exist $n > N$ such that $|x_n - x^*| < \epsilon$. It is clear that much weaker requirements are needed for a sequence to have cluster points than a limit; every convergent sequence has a cluster point x^* which is precisely the limit of the sequence, but a divergent sequence may have more than one cluster point.

With these observations, we have all of the ingredients necessary to formulate a generalization of the concept of convergence of sequences which reduces to the usual notions whenever a sequence actually has a limit. The idea is simple: if $\{x_n\}$ is a sequence, possibly divergent, which has cluster points, we refer to the largest cluster point of $\{x_n\}$ as the *limit superior* of $\{x_n\}$, denoted lim sup x_n, (or $\limsup_{n\to\infty} x_n$ or $\overline{\lim}\, x_n$) and the smallest cluster point as the *limit inferior* of $\{x_n\}$, denoted lim inf x_n (or $\liminf_{n\to\infty} x_n$ or $\underline{\lim}\, x_n$). More specifically, if $\{x_n\}$ is a sequence of real numbers, its limit superior is defined by

$$\limsup x_n = \inf_N \sup_{n \geqslant N} x_n \tag{5.1}$$

and its limit inferior by

$$\liminf x_n = \sup_N \inf_{n \geqslant N} x_n \tag{5.2}$$

Alternatively, x^* is the limit superior of $\{x_n\}$ if and only if

(i) For every $\epsilon > 0$, there exist natural numbers N such that

$$x_n < x^* + \epsilon \tag{5.3a}$$

for all $n \geqslant N$, and

(ii) For given $\epsilon > 0$ and N, there exist $n \geqslant N$ such that

$$x_n > x^* - \epsilon \tag{5.3b}$$

Analogous properties hold for the limit inferior of $\{x_n\}$.

There are other ways to interpret the limits superior and inferior of sequences which help clarify these concepts. We record some of these in the following theorem.

Theorem 5.1. Let $\{x_n\}$ be a sequence of real numbers which is bounded. Then the following are equivalent:

(i) $x^* = \limsup x_n$

(ii) If $y_m = \sup\{x_n \mid n \geqslant m\}$, then $x^* = \inf\{y_m \mid m \geqslant 1\}$.

(iii) If $y_m = \sup\{x_n \mid n \geqslant m\}$, then $x^* = \lim_{m\to\infty} y_m$.