ON NON-CONVEXITY IN PLASTICITY

- S.J. KIM and J.T. ODEN

Department of Aerospace Engineering and Engineering Mechanics
The University of Texas at Austin

1. INTRODUCTION

The role of the assumption of convexity of the yield surface in classical plasticity has been central to the development of many aspects of this theory for several decades. Convexity, for example, is basic to many notions of material stability that have been put forth in recent years; it is the basis of the standard flow rules of classical plasticity, and it is a key property of theories based on various stress potentials that have been advanced by several notable authors.

Nevertheless, the necessity of convexity in a more general theoretical setting for a plasticity theory has been challenged by several authors on both theoretical and experimental grounds. ILYUSHIN [1960], for example, argued that the assumption of convexity of the yield surface can be omitted if one drops the hypothesis that elastic properties of a material be invariant during plastic deformation and PANAGIOTOPoulos [1982] used these arguments in a note aimed at applications of non-convex optimization theory to certain plasticity problems. So-called anisotropic hardening of certain elasto-plastic materials can be characterized by a distorted yield surface which, according to results of some experimentalists, need not be convex. In this regard, the experimental results of TOXAWA [1978] are also worth mentioning. These tests pertained to mixed biaxial testing for
70/30 brass and S10C stainless steel with $\varepsilon_{\text{pre}} = 0$ (no prehistory effects of loading) and $\varepsilon_{\text{pre}} = 0.3$ or $\varepsilon_{\text{pre}} = 0.2$ (each material specimen had suffered permanent plastic strain of 30 percent before testing). In the case of $\varepsilon_{\text{pre}} = 0$, yield loci of the each material show that the yield surface of the material follows closely the Von-Mises yield criteria which is an isotropic hardening representation, as shown in Figure 1. In the case of substantial prestrain, however, experimental results show a significant departure from this isotropic law. In the cases of .2% plastic strain for $\varepsilon_{\text{pre}} = 30$ percent for brass and $\varepsilon_{\text{pre}} = 20\%$ for stainless steel, measured portions of the yield surfaces opposite to the loading direction are straight lines. Furthermore, if we examine the data points in the case of brass, a noticeably non-convex portion of the yield surface can be obtained depending on how these data points are converted (see Fig. 1(c)).

The so-called mixed hardening rules can produce non-convexity in the yield surface. In micro-mechanics, dislocation theory suggests two kinds of hardening: isotropic hardening, which may arise from the cutting of the dislocation forest and/or three dimensional, statistically developed symmetric dislocation tangles, etc. and anisotropic hardening, which may be produced from the dislocation pile-ups and dislocation rings encircling dispersions, etc. The natural extension of these micro-mechanical observations to a phenomenological theory might involve a combination of these hardening mechanisms. This may be achieved by using the concept of "degree of isotropy" as did Weng [1980] in his study of crystal plasticity. For example, to combine these representations, a yield stress measure $\tau_{\theta}$, in the direction oriented an angle $\theta$ from the loading direction, can be introduced in the form,
Figure 1. a) Yield surfaces after and before uniaxial tension in $x$ direction for 70/30 Brass, b) yield surface after uniaxial tension in $x$ direction for S10C steel (TOXAMA [1979]) and c) non-convex fitting of the test results in (a).
\[ \tau_\theta = \tau_\theta + \alpha \tau_I + (1-\alpha)f(\theta) \tau_A \] (1)

where \( \tau_\theta \) is an initial measure of yield and \( \alpha \) is the degree of isotropy, \( \tau_I \) is the magnitude of isotropic hardening and \( \tau_A \) is the magnitude of anisotropic hardening, i.e.,

\[ \alpha = 0 \Rightarrow \text{pure isotropic} \]
\[ \alpha = 1 \Rightarrow \text{pure anisotropic} \]

and \( f(\theta) \) may be acquired from experimental results. The graphs in Fig. 2(a) were produced by using the above idea. Notice, again, that a departure from convexity is observed for certain choices of parameters.

Another possibility for non-convexity of a yield surface arises when the material experiences different directions of loading histories as seen in Figure 2(b). For arguments about the absence of convexity in so-called \textit{instable materials}, see the papers of GREEN and NAGHDI [1965], PALMER, MAIER and DRUCKER [1967], NAGHDI and TRAPP [1975] and the references therein.

In the present paper, we explore the consequences of non-convexity in a general thermodynamic setting by making use of the relatively new mathematical apparatus for handling non-convex non-differentiable functions developed by CLARKE [1973, 1975, 1977] and ROCKAFELLAR [1979, 1980]. In sections 2 and 3, we lay down preliminary mathematical concepts pertaining to non-convex analysis, while relegating better-known ideas from convex analysis to an Appendix. We also review certain thermodynamic ideas. The major results begin in Section 4 where we introduce materials of type \( \mathbb{N} \), which we designate a general class of materials which is very general and includes very general viscoplastic materials and which are partially characterized by the existence of a generalized "elastoplastic" potential, whose generalized sub-differentials contain the rates of plastic deformation and internal state variable, and this prospect makes it possible...
An original isotropic yield surface (convex)

(a) Due to a strong prestrain

(b) Due to two different loading histories.

Figure 2. Generation of non-convex yield surfaces.
2. MATHEMATICAL PRELIMINARIES AND
NON-CONVEX ANALYSIS

A review and summary of many of the concepts of convex analysis are
given in the Appendix. Here we shall record extensions of these concepts
to non-convex problems following the ideas of CLARKE [1973] and ROCKAFEL-
LAR [1980]. Unless noted otherwise, $V$ denotes a topological vector space,
$V^*$ the topological dual of $V$, and $\langle \cdot , \cdot \rangle$ duality pairing on $V^* \times V$.

Contingent and Tangent Cones

Let $K$ be a nonempty subset of a topological vector space $V$. Then
the contingent cone to $K$ at a point $u \in K$ is defined as the set

$$C_K(u) = \lim \sup_{\varepsilon \to 0^+} \frac{1}{\varepsilon}(K-u) \quad (2.1)$$

Likewise, the tangent cone to $K$ at $u$ is defined as the set

$$T_K(u) = \lim \inf_{u' \to u, \varepsilon \to 0^+} \frac{1}{\varepsilon} [K - u'] \quad (2.2)$$

To interpret the notation used in (2.1) and (2.2), we use the concept
of a limit superior (inferior) of a multifunction defined on a topological
vector space given in the Appendix. Let $\Gamma$ be a set-valued function from
$[0, \infty) \times K$ into $V$ such that

$$\Gamma(\theta, u) \equiv \frac{1}{\theta}(K - u) = \{v \in V \mid v = \frac{1}{\theta}(w-u), w \in K, \theta \in [0, \infty)\} \quad (2.3)$$

Then
\[ C_K(u) = \limsup_{\theta \to 0^+} \Gamma(\theta, u) \]
\[ = \bigcap_{A \in N(0)} \bigcup_{\lambda > 0} \bigcup_{\theta \in (0, \lambda)} [(\Gamma(\theta, u) + A)] \tag{2.4} \]

and

\[ T_K(u) = \liminf_{\theta \to 0^+} \Gamma(\theta, u') \]
\[ = \bigcap_{A \in N(0)} \bigcup_{B \in N(u)} \bigcup_{u' \in K \cap B} \bigcup_{\lambda > 0} \bigcup_{\theta \in (0, \lambda)} [(\Gamma(\theta, u') + A)] \tag{2.5} \]

where \( N(0) \) and \( N(u) \) denoted collections of neighborhoods of 0 and \( u \), respectively.

To visualize \( C_K(u) \) and \( T_K(u) \), we note that for \( K \subset \mathbb{R}^N \),

\[ C_K(u) = \{ v \in \mathbb{R}^N \mid \exists \theta_k \to 0^+, v_k \to v, \text{ such that } u + \theta_k v_k \in K \} \tag{2.6} \]

\[ T_K(u) = \{ v \in \mathbb{R}^N \mid \forall \theta_k \to 0^+, u_k \to u, \text{ with } u_k + \theta_k v_k \in K \} \]

A two-dimensional case is illustrated in Fig. 3. Suppose \( u \) terminates at a cusp in a non-convex set \( K \), as shown. The entire plane can be represented as the union of four cones with vertex \( u \): BOD, DOC, COA, and AOB, with 0 the terminix of \( u \). Clearly, for any point \( v \) inside the cone BOD \( \cup \) DOC \( \cup \) COA, it is always possible to find a sequence of positive numbers \( \{ \theta_k \} \) such that \( u + \theta_k v_k \in K \) of any sequence \( v_k \to v \). Outside of this cone (interior to AOB), it is impossible to find \( \{ \theta_k \} \) for which...
Figure 3. Interpretation of contingent cone $C_K(u)$ and tangent cone $T_K(u)$ for a non-convex set $K$. 
\[ u + \theta_k \frac{v}{\theta_k} \in K. \] Hence,

\[ C_K(u) = BOD \cup DOC \cup AOB \]

Similarly, pick a sequence \( \frac{v}{\theta_k} + u \) where \( \frac{v}{\theta_k} \) is a sequence of vectors tracing out the arc EO on K. The legitimate vectors \( v \) with sequences \( \frac{v}{\theta_k} \rightarrow V \) such that \( \frac{v}{\theta_k} + \theta_k \frac{v}{\theta_k} \in K \) as \( \theta_k \rightarrow 0 \) will be those in the half-space BOD \cup DOC. Similarly, for \( \frac{v}{\theta_k} \) approaching \( u \) along FO, we must choose \( v \) in DOC \cup COA. All other sequences \( \frac{v}{\theta_k} \in K, \frac{v}{\theta_k} + u \) yield acceptable \( v \) in either of these half spaces. Thus, \( T_K(u) \) must represent the intersection:

\[ T_K(u) = DOC \]

If \( K \) is convex, then

\[ T_K(u) = C_K(u) \]

**Normal Cone**

For \( K \subset V, K \neq \emptyset \), the normal cone to \( K \) at \( u \) is a subset of the dual space \( V^* \) defined by

\[ N_K(u) = \{ u^* \in V^* | \langle u^*, v \rangle \leq 0 \ \forall v \in T_K(u) \} \]

(2.8)

In two dimensions, \( N_K(u) \) consists of the vectors through \( u \) which make obtuse angles with the vectors in \( T_K(u) \), as shown in Fig. 4.

**Clarke-Rockafellar Derivatives.**

Let \( F \) be any extended real-valued function defined on \( V \) and let \( F \) be finite at a point \( u \in V \). Then various types of subderivatives of \( F \) at \( u \) can be defined as follows.
Figure 4. Normal and tangent cones at points of a set $K$. 
• **Upper Subderivative.** The upper subderivative of \( F \) at \( u \) in direction \( v \) is defined as

\[
D^+F(u;v) = \limsup_{\substack{\theta \to 0^+}} \frac{1}{\theta} \left[ F(u' + \theta v') - \alpha \right]
\]

The notation \( \limsup \) is defined in the Appendix, and by the notation

\[(u',\alpha) \rightarrow (u,F(u))\]

we signify the convergence of a sequence \((u',\alpha) \in \text{epi } F\) to a point on the graph of \( F \):

\[
(u',\alpha) \rightarrow (u,F(u)) \iff (u',\alpha) \rightarrow (u,F(u)),
\]

\[\alpha \geq F(u')\]

If \( F \) is l.s.c. (lower semicontinuous) at \( u \), then (2.8) reduces to

\[
D^+F(u;v) = \limsup_{\substack{\theta \to 0^+}} \inf \frac{1}{\theta} \left[ F(u' + \theta v') - F(u') \right]
\]

• **Lower Subderivative.** The lower subderivative of \( F \) at \( u \) in direction \( v \) is defined by

\[
D^\ast F(u;v) = \liminf_{\substack{\theta \to 0^+}} \sup \frac{1}{\theta} \left[ F(u' + \theta v') - \alpha \right]
\]

with

\[(u',\alpha) \rightarrow (u,F(u)) \iff (u',\alpha) \rightarrow (u,F(u)),
\]

\[\alpha \leq F(u')\]

If \( F \) is u.s.c. at \( u \),
\[ D^+ F(u;v) = \liminf_{(u',F(u')) \to (u,F(u)), v' \to v} \frac{1}{\theta} F(u' + \theta v') - F(u') \] (2.12)

The derivatives \( D^+ F(u;v), D^- F(u;v) \) are referred to here as Clarke-Rockafellar derivatives (or C/R-derivatives for compactness).

To help understand the meaning of the C/R-derivatives, we consider the two examples shown in Figs. 5 and 6.

**Example 1.** Let \( f_1 \) and \( f_2 \) denote the real-valued functions,

\[
f_1(x) = \begin{cases} 
  x^2 - 4x + 4 & x \leq 1 \\
  -x^2/2 + 2x - 1/2 & x > 1
\end{cases}
\]

and

\[
f_2(x) = \begin{cases} 
  x^2 - 4x + 1 & x \leq 1 \\
  -x^2/2 + 2x + 3/2 & x > 1
\end{cases}
\]

As we observe in Fig. 5, \( f_1(x) \) is continuous at \( x = 1 \) but nondifferentiable in classical sense and \( f_2(x) \) is discontinuous but lower semicontinuous at \( x = 1 \). Since these functions are defined on \( \mathbb{R} \), the definition of C/R-derivatives reduces to

\[
D^+ f(x;y) = \limsup_{(x',f(x')) \to (x,f(x))} \frac{1}{\theta}[F(x' + \theta y) - F(x')]
\]

and

\[
D^- f(x;y) = \liminf_{(x',f(x')) \to (x,f(x))} \frac{1}{\theta}[F(x' + \theta y) - F(x')]
\]

Around \( x = 1 \), we have two subsequences of derivatives with \( y=1 \) which con-
Figure 5. The upper (lower) subderivatives of (a) continuous but nondifferentiable and (b) discontinuous at $x = 1$. 
verges to 1 and -2 for both \( f_1 \) and \( f_2 \) and we easily conclude that

\[
\begin{align*}
D^+f_1(1;1) &= 1 \quad (= \"\sup\{1,-2\}\") \\
D^+f_2(1;1) &= 1 \\
D^+f_1(1;1) &= -2 \\
D^+f_2(1;1) &= -2
\end{align*}
\]

We remark that the convergence \((x',a) \mapsto (x,f(x))\) in the definition of C/R derivatives (2.8) and the definition of tangent cone (2.2) lead to the fact that the epigraph of the function \( y \to D^+f(x;y) \) is the tangent cone \( T_{\text{epi}f}(x,f(x)) \) which is shown in Fig. 5(a).

**Example 2.** A better example of C/R derivatives can be constructed in \( \mathbb{R}^2 \). Consider the lower semicontinuous function \( f = f(x,y) \) shown in Fig. 6. The numbers indicated in the figure are intended to mean the following:

- The slope of line GH at G in the direction \( v = (1,0) \) is +0.5
- The slope of line EF at F in the direction \( v = (1,0) \) is -0.5
- The slope of line CD at C in the direction \( v = (1,0) \) is -0.8
- The slope of line AB at B in the direction \( v = (1,0) \) is +0.2

Let us calculate C/R derivative at the origin \( u = (0,0) \) with direction \( v = (1,0) \). Recall that

\[
D^+f((0,0);(1,0)) = \limsup_{(u',F(u')) \to ((0,0),f(0,0))} \inf_{v' \to v \to 0^+} \frac{1}{\delta} (f(u'+\delta v')-f(u'))
\]

Along the direction of \( v = (1,0) \), we choose two sequences approaching \( u = (0,0) \) from either positive or negative x-axis: \( \liminf Df(u',v') \)

\( v' \to v \) taken from the positive side of x-axis will be a sequence of slopes along
Figure 6. The upper (lower) subderivative of the lower semicontinuous function $F(u)$. 
EF and \( \liminf_{v' \to v} Df(u',v') \) taken from the negative side of x-axis will be the slopes along DC, where \( Df(u',v') = \frac{1}{\epsilon}[f(u' + \epsilon v') - f(u')] \). Next we take \( \limsup_{(u',f(u')) \to ((0,0),f(0,0))} \) of the sequences of slopes and then finally the slope at \( F \) along the curve EF as \( D^+f((0,0):(1,0)) \). Similarly, we can get

\[ D^+f((0,0):(1,0)) = \text{the slope at } B \text{ along } AB. \]

Quantitatively

\[ D^+f((0,0):(1,0)) = -0.5 \]

and

\[ D^+f((0,0):(1,0)) = +0.2 \]

**Generalized Subdifferentials**

The subdifferential \( \partial F(u) \) of a function \( F \) at \( u \) is a well known concept in convex analysis (see the Appendix for details). By using the subderivatives defined in (2.8) and (2.10), we define the generalized subdifferential of \( F:V \to \mathbb{R} \), at a point \( u \) where \( F(u) \) is finite, as the set

\[ \partial F(u) = \{ u^* \in V^* | \langle u^*, v \rangle \leq D^+F(u;v) \forall v \in V \} \quad (2.13) \]

Now we list two useful theorems due to ROCKAFELLAR [1980].

**Theorem 2.1.** Let \( F \) be any extended real-valued function on \( V \), and let \( u \) be any point at which \( F \) is finite. Then \( \partial F \) is a weak*–closed convex subset of \( V^* \) and

\[ \partial F(u) = \{ u^* \in V^* | (u^*, -1) \in \operatorname{epi} F (u,F(u)) \} \quad (2.14) \]

If \( D^+F(u;0) = -\infty \), then \( \partial F(u) \) is empty, but otherwise \( \partial F(u) \) is nonempty.
\[ D^+F(u;v) = \sup \{ \langle u^*, v \rangle \mid u^* \in \overline{\partial F}(u) \} \quad \text{for all } v \in V \quad (2.15) \]

**Theorem 2.2.** If \( F \) is a convex function on \( V \), then \( \overline{\partial F}(u) \) agrees with the subgradient set in the sense of convex analysis:

\[ \overline{\partial F}(u) = \partial F(u) = \left\{ u^* \in V^* \mid \langle u^*, v \rangle \leq F'(u;v), \forall v \in V \right\} = \left\{ u^* \in V^* \mid \langle v-u, u^* \rangle \leq F(v)-F(u), \forall v \in V \right\} \quad (2.16) \]

Here \( F'(u;v) = \lim_{t \to 0^+} \frac{F(u+tv) - F(u)}{t} \) is called the one-sided directional derivative which exists for all \( v \) when \( F \) is convex (although it may be infinitive). \( \square \)

**Remark.** If \( F(u) \) is a characteristic function with respect to a set \( K \), i.e., if

\[ F(u) = \psi_K(u) = \begin{cases} 0 & \text{if } u \notin K \\ +\infty & \text{if } u \in K \end{cases} \]

then

\[ \overline{\partial F}(u) = \left\{ u^* \in V^* \mid \langle u^*, v \rangle \leq D^+\psi_K(u;v), \forall v \in V \right\} \quad (2.17) \]

\[ = N_K(u) \]

This fact can be more easily visualized in the case of convex \( F \), i.e.,

\[ \partial \psi_K(u) = \left\{ u^* \in V^* \mid \langle u^*, v-u \rangle \leq \psi_K(v) - \psi_K(u), \forall v \in K \right\} = \left\{ u^* \in V^* \mid \langle u^*, v-u \rangle \leq 0, \forall v \in K \right\} = N_K(u) \]

since \( \psi_K(v) = \psi_K(u) = 0 \).
3. MECHANICAL PRELIMINARIES

3.1 Classical Convex Plasticity

Classical plasticity theory rests on the assumption of the existence of a convex function \( F : M = \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow [0, \infty) \) of the stress tensor \( \sigma \), called the yield function of the material, which has the property that plastic flow at a particle \( X \) of the material is signaled whenever \( F(\sigma(X)) = 0 \); otherwise the deformation at \( X \) is elastic:

\[
F(\sigma(X)) \begin{cases} < 0 \Rightarrow \text{elastic deformation} \\ = 0 \Rightarrow \text{plastic flow} \end{cases} \tag{3.1}
\]

The only stress states admissible in such theories are those for which \( F(\sigma) \leq 0 \) or, equivalently, those stresses which belong to the convex set

\[
K = \{ \sigma \in M \mid F(\sigma) \leq 0 \} \tag{3.2}
\]

The infinitesimal strain tensor \( \varepsilon \) is representable as the sum of an elastic part \( \varepsilon^e \) and a plastic strain \( \varepsilon^p \), and its time rate-of-change is denoted \( \dot{\varepsilon} \):

\[
\dot{\varepsilon} = \dot{\varepsilon}^e + \dot{\varepsilon}^p \tag{3.3}
\]

It is meaningful to assume the existence of a plastic stress potential \( \phi : M + \mathbb{R} \) which is convex and l.s.c. and which has the property that

\[
\varepsilon^p \in \partial \phi(\sigma) \tag{3.4}
\]
In particular, the indicator function $\Psi_K$ of the set $K$ may define a specific stress potential:

$$\Psi_K(q) = \begin{cases} 0 & \text{if } F(q) \leq 0 \\ +\infty & \text{if otherwise} \end{cases}$$

(3.5)

Note that $\Psi_K$ is l.s.c. on $\mathbb{M}$ and that $\Psi_K$ is convex if $F$ is convex.

From (3.4) and the definition of the subdifferential,

$$\varepsilon \in \partial \Psi_K(\sigma)$$

(3.6)

for some particular stress $\sigma$ implies that

$$\langle \varepsilon^p, \sigma - \sigma \rangle \geq 0 \quad \forall \sigma \in K$$

(3.7)

This result, of course, is the classical normality condition which establishes that the strain rate is normal to the yield surface or lies in the normal cone of the yield surface at corners. (see Fig. 7).

3.2 Continuum Thermodynamics of Elasto-Plastic Materials.

Our purpose in this subsection is to introduce notation and to review some aspects of continuum thermodynamics that are to be used later.

We begin by considering the motion of a material body $B$ relative to a fixed reference configuration $C_0 \subseteq \mathbb{R}^N$ (N typically 3), which is defined by the map $\kappa_0 : B \rightarrow \mathbb{R}^N, \ X = \kappa_0(X)$. The spatial position $x$ of a particle $X$ at time $t$ is then given by a relation of the type

$$x = \chi(X,t)$$

(3.8)
Figure 7. A convex yield surface and the normality condition.
with $X \in \mathcal{K}_0(B)$, $t \geq 0$, and $X$ a continuous invertible map from $C_0$ into $\mathbb{R}^N$. The deformation gradient tensor $F$ at $X$ at time $t$ is defined by

$$F = \nabla_{X^X} X = \frac{\partial X}{\partial \tilde{X}}$$

(3.9)

The motion of the body from the reference configuration $C_0$ to its current configuration $C_t$ can be decomposed into steps $C_0 \rightarrow C_p \rightarrow C_t$ or $C_0 \rightarrow C_e \rightarrow C_t$, where $C_p$ and $C_e$ denote the plastic intermediate configuration and the elastic intermediate configuration, respectively as in Fig. 8. Following the notation in NEMAT-NASSER [1979], consider the deformation of a vector element of the material $d\tilde{X}$ in $C_0$. The relative displacement of its ends is given by

$$d\tilde{x} - d\tilde{X} = (F - I) d\tilde{X} = d\tilde{u}$$

(3.10)

The relative plastic displacement is given by

$$d\tilde{p} - d\tilde{x} = (\tilde{F}^p - I) d\tilde{X} = d\tilde{u}^p$$

(3.11)

and the relative elastic displacement by

$$d\tilde{x} - d\tilde{p} = (F^e - I) d\tilde{p}$$

$$= d\tilde{n} - d\tilde{X}$$

$$= (\tilde{F}^e - I) d\tilde{X} = d\tilde{u}^e$$

(3.12)

Thus

$$\tilde{du} = \tilde{du}^e + \tilde{du}^p$$

(3.13)

gives
Figure 8. Neumüller's decomposition of configurations.
\[
\mathbf{F} = \mathbf{F}^e + \mathbf{F}^p - \mathbf{I} \quad (3.14)
\]

where \( \mathbf{p} \) and \( \mathbf{n} \) are position vectors in \( \mathbb{C}_p \) and \( \mathbb{C}_e \), respectively.

Therefore, the velocity gradient tensor \( \mathbf{L} \) is given by

\[
\mathbf{L} = \frac{\partial \mathbf{X}}{\partial \mathbf{X}} = \mathbf{F} \mathbf{F}^{-1} = \mathbf{L}^e + \mathbf{L}^p \quad (3.15)
\]

where

\[
\mathbf{L}^e = \dot{\mathbf{F}} \mathbf{F}^{-1} \quad \text{and} \quad \mathbf{L}^p = \dot{\mathbf{F}} \mathbf{F}^{-1} \quad (3.16)
\]

Here superimposed dots \( (\cdot) \) indicate time-rates.

The symmetric part of \( \mathbf{L} \) is the deformation rate tensor \( \mathbf{D} \) and is also representable as the sum of two parts,

\[
\mathbf{D} = \frac{1}{2}(\mathbf{L} + \mathbf{L}^T) = \frac{1}{2}(\mathbf{L}^e + \mathbf{L}^T + \mathbf{L}^p + \mathbf{L}^{pT}) \quad (3.17)
\]

\[
= \mathbf{D}^e + \mathbf{D}^p
\]

The thermomechanical behavior of the body is governed by the principles of conservation of mass, energy, balance of linear and angular momentum, and the law of entropy production (the second law). Local forms of these principles are listed as follows:

Conservation of Mass

\[
\rho \det \mathbf{F} = \rho_0 \quad (3.18)
\]

Balance of Linear and Angular Momentum

\[
\text{div} \mathbf{\sigma} + \rho \mathbf{b} = \rho \dot{\mathbf{X}} \quad (3.19)
\]

\[
\mathbf{\sigma} = \mathbf{\sigma}^T \quad (3.20)
\]
Conservation of Energy

\[ \rho \dot{E} = \text{tr}(\sigma L) - \text{div} \, q + \rho r \]  

(3.21)

Clausius-Duhem Inequality

\[ \rho \dot{\eta} - \text{div} \frac{q}{\theta} - \frac{r}{\theta} \geq 0 \]  

(3.22)

Here \( \rho \) is the mass density, \( \rho_0 \) the mass density in the reference configuration, \( \sigma \) is the Cauchy stress tensor, \( b \) the body force per unit mass, \( \varepsilon \) the specific internal energy, \( q \) the heat flux vector, \( r \) the heat supply per unit mass per unit time, \( \eta \) the specific entropy, and \( \theta \) the absolute temperature. The Helmholtz free energy is defined by

\[ \phi = \varepsilon - \theta \eta \]  

(3.23)

so that (3.22) can also be written in the form,

\[ -\rho \dot{\phi} - \rho \dot{\eta} + \text{tr}[\sigma(L^e + L^p)] - \frac{1}{\theta^2} \cdot \text{grad} \, \theta \geq 0 \]  

(3.24)

It is well known that the Clausius-Duhem inequality can impose conditions on the forms of constitutive equations for the material of which the body is composed. For example, consider a class of materials characterized by a set of five constitutive equations of the form (cf. Coleman and Gurtin [1967]):

\[
\begin{align*}
\phi &= \phi(F^e, \theta, g, \alpha) \\
\sigma &= \sum(F^e, \theta, g, \alpha) \\
\eta &= N(F^e, \theta, g, \alpha) \\
q &= Q(F^e, \theta, g, \alpha) \\
\dot{\alpha} &= H(F^e, \theta, g, \alpha)
\end{align*}
\]  

(3.25)
where $g = \text{grad } \theta$, $\alpha$ is a tensor referred to as an internal state variable which is sometimes introduced to model changes in the microstructure of the material or to depict history effects via the evolution equation indicated above.

Thermodynamic restrictions on the forms of the constitutive functionals in (3.25) can be determined using the well-known strategy of Coleman and Noll [1963]: we assume that the map

$$(F^e, \theta, g, \alpha) \rightarrow \Phi(F^e, \theta, g, \alpha)$$

is $C^1$ in each argument. The rate-of-change of the free energy satisfies

$$\dot{\Phi} = \frac{\partial \Phi}{\partial F^e} : \dot{F}^e + \frac{\partial \Phi}{\partial \theta} \dot{\theta} + \frac{\partial \Phi}{\partial g} : \dot{g} + \frac{\partial \Phi}{\partial \alpha} : \dot{\alpha}$$

wherein

$$\frac{\partial \Phi}{\partial F^e} : \dot{F}^e = \frac{\partial \Phi}{\partial F_{ij}} \dot{F}_{ij}, \quad [A : B = \text{tr}(AB)]$$

$$\frac{\partial \Phi}{\partial g} : \dot{g} = \frac{\partial \Phi}{\partial g_1} \dot{g}_1, \quad \text{etc.}$$

and we have used obvious indicial notation and the summation convention.

Inequality (3.24) thus yields

$$(\sigma - \rho \frac{\partial \Phi}{\partial \theta}) : \cdot F^e + L^p : \sigma - A : \alpha$$

$$-(\rho n + \frac{\partial \Phi}{\partial \theta} \dot{\theta} - \rho \frac{\partial \Phi}{\partial g} : \dot{g} - \frac{1}{\theta} \frac{\partial \Phi}{\partial \alpha} : \dot{\alpha}) \cdot \text{grad } \theta \geq 0$$

(3.26)
where

\[
A = \frac{\partial \Phi}{\partial \alpha}
\]

(3.27)

From standard arguments, it follows that

\[
\xi_F^{-}\xi = \rho \frac{\partial \Phi}{\partial F}
\]

\[
N = - \frac{\partial \Phi}{\partial \Theta}
\]

(3.28)

\[
O = \frac{\partial \Phi}{\partial g}
\]

and we return the inequality

\[
L^D : \xi - A : H - \frac{1}{\Theta} Q \cdot g > 0
\]

In these results, it is understood that \( \xi, A, \) and \( N \) now depend upon \((F, F, \Theta, \alpha)\) and that \( H \) and \( Q \) are functions of \((F, F, \Theta, \alpha)\).

4. MATERIALS OF TYPE \( N \)

In this section, we introduce the concept of property \( T \) and a class of hypothetical materials which exhibit this property.

**Property T.** Let \( S \) denote an \( M \)-dimensional linear space \((S \cong \mathbb{R}^M)\) if a functional \( \Phi : M \rightarrow C^{[-\infty, \infty]} \) is said to have property \( T \) if and only if there exists a vector \( \alpha \), in the dual space \( S^* \) of \( S \), such that

\[\Phi(T) = \alpha \Phi(T) \] for \( T \in S \). Then we are assured that there exists a vector \( \alpha \) in the dual space \( S^* \) of \( S \) such that...
We wish to consider a hypothetical class of materials which can be partially characterized using property $T$. Let $\sum \subseteq \mathbb{R}^{3x3}$ denote the space of stress values at a point $x \in \Omega$ at time $t$ and let $\mathcal{A} \subseteq \mathbb{R}^{3x3}$ denote the space of values of the thermodynamic force conjugate to the time rate of the internal state variable $\dot{\alpha}$. We denote the space of stress-force pairs as

$$ W = \sum \mathcal{A} $$

Materials of Type $T$. A material is said to be of Type $T$ if and only if it is characterized by constitutive equations of the form (3.25) and there exists a potential function $\psi: W \to [0, \infty)$ which has a nonempty generalized subdifferential. Then, for any $(\tilde{\sigma}, \tilde{\alpha}) \in W$, there exists $(\overline{D}^\psi \tilde{\alpha}, \dot{\alpha}) \in \partial \psi(\tilde{\sigma}, \tilde{\alpha})$ such that

$$ (\overline{D}^\psi \tilde{\alpha}, \dot{\alpha}) \in \partial \psi(\tilde{\sigma}, \tilde{\alpha}) $$ (4.1)

From the results in section 2, the relation (4.1) implies the inequality
\[ \langle d^P, \sigma^* \rangle_\Sigma + \langle \dot{\alpha}, \Lambda^* \rangle_A \leq d^\psi((\sigma, \Lambda); (\tau^*, \Lambda^*)) \]

\[ \forall (\tau^*, \Lambda^*) \in \Psi \]

(\langle \cdot, \cdot \rangle_\Sigma and \langle \cdot, \cdot \rangle_A denote duality pairings on \Sigma^* \times \Sigma and \Lambda^* \times \Lambda, respectively) and geometrically, we have the normality (Hera, N-) condition.

\[ ((d^P, \dot{\alpha}), -1) \in N_{\text{epi} \psi} \left[ (\tau, \Lambda), \psi(\tau, \Lambda) \right] \]

Furthermore, if potential \( \psi \) is differentiable at \((\tau, \Lambda)\) then \( N_{\text{epi} \psi} \) at \((\tau, \Lambda)\) has a single element. So,

\[ (d^P, \dot{\alpha}) = \frac{\partial \psi}{\partial (\tau, \Lambda)} \]

or

\[ d^P = \frac{\partial \psi}{\partial \tau} \]

and

\[ \dot{\alpha} = \frac{\partial \psi}{\partial \Lambda} \]

(4.4)

We remark that if \( \psi \) is convex and l.s.c., the materials of type \( T \) reduce to the generalized plastic materials introduced by HALPHEN and NGUYEN [1975].
5. EXAMPLE

In this section, we will introduce a potential which is not necessarily convex and examine the properties of the potential.

**An Elastoplastic Potential**

We first introduce a function \( F : \mathcal{W} \rightarrow \mathbb{R} \) and a set \( C = \{ (\sigma, A) \in \mathcal{W} \mid F(\sigma, A) \leq 0 \} \). Then a potential of the elastoplasticity can be introduced as follows:

\[
\psi(\sigma, A) = \begin{cases} 
- \varepsilon \cdot \frac{1}{F(\sigma, A)} & \text{if } (\sigma, A) \in C \\
+\infty & \text{if } (\sigma, A) \notin C
\end{cases}
\] (5.1)

where \( \varepsilon \) is an arbitrary positive number.

Next, we endow the property \( T \) to this potential. These potentials characterize families:

\[
(D^P, \dot{\sigma}) \in \partial \psi(\sigma, A)
\] (5.2)

\[
\iff \langle D^P, \sigma^* \rangle + \langle \dot{\sigma}, A \rangle A \leq D\psi((\sigma, A), (\sigma^*, A))
\] (5.3)

and we investigate how this endowment can be justified. First we look at some properties of this potential \( \psi(\sigma, A) \).

**Lemma 5.1.** If a scalar valued function \( F(x) \) is convex over \( \mathbb{R}^N \) and \( F(x) < 0 \) for all \( x \in C \), then the inverse of this function \( G(x) = 1/F(x) \) is concave over \( C \).

**Proof.** \( F(x) \) is convex, i.e.,
\[
F(\lambda x + (1-\lambda)y) \leq \lambda F(x) + (1-\lambda)F(y)
\]
\[
\forall x, y \in C \quad \text{and} \quad \forall \lambda \in [0, 1] 
\]

Now,
\[
R(x, y, \lambda) = G[\lambda x + (1-x)y] - \lambda G(x) - (1-\lambda)G(y)
\]
\[
= \frac{1}{D} \left\{ F(x)P(y) - \lambda F(y)P(x,y,\lambda) - (1-\lambda)F(x)P(x,y,\lambda) \right\}
\]

where
\[
P(x, y, \lambda) = F[\lambda x + (1-\lambda)y] \quad \text{and} \quad D = P(x, y, \lambda)F(x)P(y) < 0
\]

Since \(-\lambda F(y)/D \leq 0\) and \(-(1-\lambda)F(x)/D \leq 0\), from (5.4), we have
\[
R(x, y, \lambda) \geq \frac{1}{D} \left\{ F(x)P(y) - \lambda^2 F(x)P(y) + (1-2\lambda + \lambda^2)F(x)P(y) \right\}
\]
\[
+ (\lambda - \lambda^2)(F^2(x) + F^2(y)) \right\}
\]
\[
= - \frac{(\lambda - \lambda^2)}{D} \left\{ F(x) - F(y) \right\}^2
\]
\[
\geq 0 \quad \text{since} \quad -(\lambda - \lambda^2)/D \geq 0
\]

\textbf{Theorem 5.1} If the set \(C\) is convex, then the potential \(\psi_\varepsilon(\sigma, \tilde{\lambda})\) is convex. Otherwise, \(\psi_\varepsilon\) is nonconvex.

\textbf{Proof.}

i) The set \(C\) is convex
\[
\rightarrow F(\sigma, \tilde{\lambda}) \text{ is a convex function}
\]
\[
\rightarrow 1/F \text{ is a concave function (from lemma 5.1)}
\]
\[
\rightarrow -1/F \text{ is convex}
\]

ii) Take a point \(M\) along the line LN in the Fig. 9 with nonconvex set \(C\). \(F\) is infinity at \(M\) but values of \(f\) at \(L\) and \(N\) are finite. So it is not convex.
Nonconvex set $C$

Figure 9. Nonconvex set $C$
Next, we will establish the fact that the proposed potential converges to the characteristic function in an appropriate sense. To do this, we introduce the notion of infimal convergence of WIJSMAN's [1964, 1966].

**Definition 5.1**

i) For any function $g$ defined on $\mathbb{R}$ and for any $r < 0$, define the function $r^g$ on $\mathbb{R}^n$ by

$$r^g(x) = \inf_{y \in B_r(x)} g(y)$$

(5.6)

where $B_r(x)$ is the closed ball of radius $r$ with center $x$.

ii) Let $\{g_k; k = 1, 2, \ldots\}$ be a sequence of functions on $\mathbb{R}^n$. The sequence is said to converge infimally to a function $g$ if, for every $x \in \mathbb{R}^n$,

$$\lim_{r \to 0} \lim_{k \to \infty} r^{g_k}(x) = \lim_{r \to 0} \lim_{k \to \infty} r^{g_k}(x) = g(x)$$

and we write

$$g_k \to g_{\inf}$$

**Theorem 5.2.** When $\varepsilon \to 0$, then $\psi_{\varepsilon}(v) \to \psi(v)_{\inf}$, where $v \in \{\alpha, \beta\}$, for $C \neq \phi$.

**Proof.** To prove this theorem, it suffices to show that

$$\lim_{r \to 0} \lim_{\varepsilon \to 0} r^{\phi_{\varepsilon}}(v) > 0 \text{ if } v \notin C$$

(5.7)

$$\lim_{r \to 0} \lim_{\varepsilon \to 0} r^{\phi_{\varepsilon}}(v) < 0 \text{ if } v \notin C$$

(5.8)

and

$$\lim_{r \to 0} \lim_{\varepsilon \to 0} r^{\phi_{\varepsilon}}(v) = +\infty \text{ if } v \notin C$$

(5.9)
It is easy to prove (5.7) since

$$\phi_\varepsilon(v) \geq 0, \quad \forall v \in W$$

from definition (5.1). So

$$\lim_{r \to 0} \lim_{\varepsilon \to 0} \inf \phi_\varepsilon(v) > 0$$

Next, for any $r > 0$, there is

$$w \in \mathring{C} \cap B_r(v), \quad v \in C$$

By the definition (5.6),

$$\phi_\varepsilon(v) < \phi_\varepsilon(w) \quad \forall \varepsilon$$

Thus

$$\lim_{\varepsilon \to 0} \sup_{r \to 0} \phi_\varepsilon(v) < \lim_{\varepsilon \to 0} \phi_\varepsilon(w)$$

$$= \lim_{\varepsilon \to 0} \frac{1}{-F(w)}$$

$$\leq \lim_{\varepsilon \to 0} \varepsilon M = 0$$

where $M$ is finite positive real number. Consequently,

$$\lim_{r \to 0} \lim_{\varepsilon \to 0} \sup_{r \to 0} \phi_\varepsilon(v) < 0$$

Finally, we assume the contrary to (5.9) i.e.,

$$\lim_{r \to 0} \lim_{\varepsilon \to 0} \inf_{r \to 0} \phi_\varepsilon(v) > 4\omega, \quad v \in C$$

But the openness of the set $\mathring{C} = \{v \in \mathbb{W} | v \notin C\}$ guarantees the existence of infinitely many $r$ which satisfy $0 < r < r_1$ where $r_1$ = distance $(C, v)$ for $v \notin \mathring{C}$. In the limit as $r \to 0$, $r$ has
to reach the region \( 0 < r < r_1 \), where \( r \phi_E(v) = +\infty \). So none of the sequence can assume a finite value. Therefore

\[
\lim_{r \to 0} \lim \inf_{\varepsilon \to 0} r \phi_E(v) = +\infty, \quad v \notin C
\]

\[ \square \]

Remark. In fact, the potential (5.1) is an interior penalty function for the yield constraint \( F(\sigma, A) \leq 0 \) (see ODEN and KIM [1982]).

For details on convex interior penalty functions see Mine and Fukushima [1978].

\[ \square \]

If \( \psi_C \) is convex (i.e., \( C \) is convex), then from (2.15) we have the inequality

\[
\langle g^* - g, D_p \rangle + \langle A^* - A, \dot{\alpha} \rangle \leq \phi_E(v) - \phi_C(v)
\]

\[ \forall (g^*, A^*) \in V \] (5.10)

Obviously, \( (\sigma, A) \) should belong to \( C \). Then the inequality (5.10) need be true only for all \( (g^*, A^*) \in C \), i.e.

\[
\langle g^* - g, D_p \rangle + \langle A^* - A, \dot{\alpha} \rangle \leq \psi_E(\sigma^*, A^*) - \psi_C(\sigma, A)
\]

\[ \forall (\sigma^*, A^*) \in V \] (5.11)

Furthermore, if we let \( \varepsilon \to 0 \), then \( \psi_E \inf \psi_C \) and \( \psi_C(\sigma^*, A^*) = 0 = \psi_C(\sigma^*, A^*) \). So, we recover the result of HALPHEN and NGUYEN [1975], i.e.,

\[
\langle g^* - g, D_p \rangle + \langle A^* - A, \dot{\alpha} \rangle \leq 0, \quad \forall (\sigma^*, A^*) \in C
\]

(5.12)

By using the potential in the form of interior penalty for the set \( C \), we can easily visualize the result (2.16) from (2.13).
a) the half cylinder $\Psi_C$ and $\psi_C$

b) the cross section A-A

Figure 10. The potential functions $\psi_\epsilon$ and $\Psi_C$
The indicator function \( \psi_C \) is nothing but the half cylinder, the cross-section of which is the set \( C \) shown in Fig. 10.

From the definition (2.13), \( (D^p, \alpha, -1) \) is in \( N_{\text{epi}} \phi((\sigma, A), \Phi(\sigma, A)) \).

Upon allowing \( \varepsilon \rightarrow 0 \), this element of the normal cone for \( \phi \) will coincide with an element of the normal cone for \( \psi_C \) and the projection of this vector to the space including set \( C \) (i.e., \( V \)) gives the result (2.16), i.e.,

\[
\bar{\psi}_C(\sigma, A) = N_C(\sigma, A) \\
= \{(D^p, \alpha) \in \omega^* \langle D^p, \sigma^* \rangle | \langle \alpha, \alpha^* \rangle \leq D^p \psi((\sigma, A); (\sigma^*, \alpha^*)) \} \\
(\sigma^*, \alpha^*) \in \omega \}
\]

(5.13)

We will call the relation (5.13) as the generalized normality rule for the materials of Type T.

When we have a plastic potential of the form (5.1), there exists plastic deformation for any finite value of \( \varepsilon \), no matter how small, since we have non zero elements \( (D^p, \alpha) \) even though \( F(\sigma, A) < 0 \).

But this property may, in fact, by physically reasonable for many engineering materials. Generally, in single crystalline models, we have four distinct stages in the stress-strain curve as seen in Fig. 11. After an elastic stage, one often observes an "easy-glide" stage (Stage I in the figure) during which all the free (mobile) dislocations move and a large amount of plastic strain is realized. Stage II is called the work-hardening stage, and unit dislocations are generated from, say, Frank-Read sources and interactions between dislocations making barriers like Lomer-Cottrell locks. In stage III, breakdowns of these barriers may occur, giving very low hardening with some plastic strain. For more detailed description of these
Figure 11. Stress-strain curve of single crystal material.
behavior, see DIETER [1976] or WILKO [1983].

In polycrystalline materials, the easy glide stage is not observed as frequently as in the single crystal since the movements of dislocations are stopped more readily by a larger density of built-in obstacles due to the nature of the grain boundaries. Therefore, the usual stress strain curve, i.e., a linear elastic region and a strain hardening region, is obtained. In fact, the easy glide stage is included in the elastic region since the flow of dislocations and plastic strain are ignorable. But, even though dislocations move very short distance, if the irrecoverable dislocation movement is interpreted as plastic deformation, some changes in plastic strain may actually occur in the "elastic" region, specifically in the unloading process. This point is discussed in some papers, e.g., HUTCHINSON [1970] and NEMAT-NASSER [1982].

For materials with a significant elastic response regions characterized by such potentials, it may be convenient to regard $\varepsilon$ as an interior parameter associated with the constraint set $C$ and to choose $\varepsilon$ small so that negligible plastic strain rates and rate of internal state variable, $(D^P, \dot{\alpha})$, occur for elastic deformations. As indicated in the reduction of our theory to the cases covered by (5.10) to (5.12), taking $\varepsilon \to 0$ provides for the concept of a convex yield function and essence of classical plasticity within the framework of our theory. However, if this theory were used to develop computational methods for the solution of classical plasticity problems, one may still wish to choose a small value of $\varepsilon > 0$ to handle yielding by means of an interior penalty method.

ACKNOWLEDGEMENT. Research supported by the U.S. Army Research Office under contract No. DAAG29-83-K-0036 is gratefully acknowledged.
REFERENCES


ROCKAFELLAR, R.T. [1979], La théorie de sous-gradients et ses applications à l'optimisation, Les Presses de l'Université de Montréal, Montréal, Canada.


WILKOV, M. [1983], "Dislocations and Mechanical Properties," Lecture Notes, The University of Texas at Austin, Austin, Texas.
APPENDIX

ELEMENTS OF CONVEX ANALYSIS

We shall provide here a brief summary of some of the concepts of convex optimization theory which are prerequisite to the ideas discussed in the body of the paper. For more detailed accounts, the books of EKELAND and TEMAM [1976] or ROCKAFELLAR [1970, 1979] or the recent text of ODEN [1984] can be consulted.

We begin by introducing the following notations:

\[ \mathbb{R} = \text{the extended real numbers; if } \mathbb{R} \text{ is the real number system, } \mathbb{R} = \mathbb{R} \cup \{\pm \infty\} \]

\[ U, V = \text{topological vector spaces} \]

\[ U^*, V^* = \text{topological dual spaces of } E \text{ and } S \text{ respectively} \]

\[ \langle \cdot, \cdot \rangle_U, \langle \cdot, \cdot \rangle_V = \text{duality pairing on } V^* \times V \text{ and } U^* \times U, \text{ respectively; i.e. if } v^* \in V^* \text{ and } v \in V, \text{ then } v^*(v) = \langle v^*, v \rangle_V, \]

e etc.

It is worthwhile to recall the definition of the limit-superior (\( \lim \sup \)) and the limit-inferior (\( \lim \inf \)) of sequences of real numbers, extended real-valued functions of sequences, and sequences of sets in a topological vector space \( V \).

\[ \lim \sup, \lim \inf, \]

- For \( \{a_n\} \) a sequence of real numbers

\[
\begin{align*}
\limsup a_n &= \inf_{n \to \infty} \sup_{N \in \mathbb{N}, n \geq N} a_n \\
\liminf a_n &= \sup_{n \to \infty} \inf_{N \in \mathbb{N}, n \geq N} a_n
\end{align*}
\]  

\( (A.1) \)
Figure A.1. Limit superior, limit inferior of discontinuous function $f: \mathbb{R} \to \mathbb{R}$ at $x_0$. 
• For $f : \mathbb{R} \to \mathbb{R}$

\[
\begin{align*}
\limsup_{x'} f(x') &= \inf_{\delta > 0} \sup_{0 < |x' - x| < \delta} f(x') \\
\liminf_{x'} f(x') &= \sup_{\delta > 0} \inf_{0 < |x' - x| < \delta} f(x')
\end{align*}
\]  

(A.2)

• For $\{A_n\}$ a sequence of subsets of the underlying set of topological space $V$,

\[
\begin{align*}
\limsup_{n \to \infty} A_n &= \bigcap_{m=1}^{\infty} \left( \bigcup_{n=m}^{\infty} A_n \right) \\
\liminf_{n \to \infty} A_n &= \bigcup_{m=1}^{\infty} \left( \bigcap_{n=m}^{\infty} A_m \right)
\end{align*}
\]  

(A.3)

For example, if $\{x_n\}$ is a sequence of real numbers which converges to $x$ and $f : \mathbb{R} \to \mathbb{R}$, $\limsup f(x_n)$ is the supremum of all cluster points of $x_n \to x$ of $f$ at $x$, as indicated in Fig. A.1 (with an analogous interpretation for $\liminf$).

The concept can also be applied to multifunctions from one topological vector space to another. Indeed, if $\Gamma : U \to V$ (with $\Gamma(u)$ a subset of $V$ for each vector $u \in U$), then

\[
\limsup_{u' \to u} \Gamma(u') = \bigcap_{A \in N(0)} \bigcup_{B \in N(u)} \left[ \bigcup_{u' \in B} (\Gamma(u') + A) \right]
\]

(A.4)

and

\[
\limsup_{u' \to u} \Gamma(u') = \bigcap_{A \in N(0)} \bigcup_{B \in N(u)} \left[ \bigcap_{u' \in B} (\Gamma(u') + A) \right]
\]

(A.5)

where $N(0)$ and $N(u)$ are collections of neighborhoods of $0,u$, respectively.
**Lim sup inf/Lim inf sup**

In addition to the notion of limit superior and limit inferior, it is convenient to introduce the concepts of lim sup inf and lim inf sup introduced by ROCKAFELLAR [1980].

Let $F$ be an extended real-valued function from $U \times V$ into $\mathbb{R}$, let $u' \to u$ in $U$ and $v' \to v$ in $V$. Then we define

$$\text{Lim sup inf } F(u',v') \triangleq \sup_{u' \to u} \inf_{v' \to v} \sup_{B \subseteq N(v)} \inf_{A \subseteq N(u)} \inf_{u' \in A} \sup_{v' \in B} F(u',v') \quad (A.6)$$

Likewise,

$$\text{Lim inf sup } F(u',v') \triangleq \inf_{u' \to u} \sup_{v' \to v} \inf_{B \subseteq N(v)} \sup_{A \subseteq N(u)} \inf_{u' \in A} \sup_{v' \in B} F(u',v') \quad (A.7)$$

Similarly, Lim sup sup and Lim inf inf can be defined in an analogous way.

The meaning of these operations can be more easily understood in the case of a real-valued function $F$ defined on $\mathbb{R}^2$, such as the discontinuous at the origin shown in Fig. A.2. To compute Lim inf inf $F(x',y')$, $x' \to 0$, $y' \to 0$ for example, we compute lim inf $F(x',y')$ for a fixed $y'$. This gives the function of $y'$ which has as its graph the curve $\overline{AC} \cup \overline{EH}$. The lim inf of this function is the point $E$, denoted $F_1$ in the figure. Similarly lim sup $F(x',y')$ for fixed $y'$, is the curve, $\overline{HD} \cup \overline{AB}$ and lim inf of this curve is the point $D$, denoted $F_2$ in the figure. In summary, for this example,

$$\text{Lim inf inf } F(x',y') = F_1 \quad x' \to 0 \quad y' \to 0$$

$$\text{Lim sup inf } F(x',y') = F_2 \quad x' \to 0 \quad y' \to 0$$
Figure A.2. A function $F$ discontinuous at the origin $O$: solid lines at surfaces of discontinuity indicate values assumed by $F$. 
\[ \liminf_{x' \to 0} \sup_{y' \to 0} F(x',y') = F_3 \]
\[ \limsup_{x' \to 0} \sup_{y' \to 0} F(x',y') = F_4 \]

**Convex Set.** A set \( C \subseteq V \) is convex if and only if
\[ \theta u + (1-\theta)v \in C \quad \forall u,v \in C \text{ and } \forall \theta \in [0,1] \]  
(A.8)
i.e. \( C \) is convex iff the line segment connecting any two points \( u \) and \( v \) of \( C \) lies totally in \( C \).

**Epigraph.** The epigraph of an extended real-valued function \( F:V \to \mathbb{R} \) is defined as the set
\[ \text{epi } F = \{(v,\lambda) \in V \times \mathbb{R} \mid \lambda > F(v)\} \]  
(A.9)

**Convex Function.** \( F:V \to \mathbb{R} \) is convex iff \( \text{epi } F \) is a convex set (see Fig. A3). This is equivalent to \( F \) satisfying the condition,
\[ F(\theta u + (1-\theta)v) \leq \theta F(u) + (1-\theta)F(v) \quad \forall u,v \in V, \forall \theta \in [0,1] \]  
(A.10)

**Lipschitzian Function.** A function \( F:V \to \mathbb{R} \) is Lipschitzian around \( v \in V \) iff there exists a neighborhood \( N(v) \) of \( v \) on which \( F \) is finite and such that
\[ |F(u) - F(v)| \leq C P(u-v) \quad \forall u,v \in N(v) \]  
(A.11)
where \( C \) is a positive constant and \( P \) is a continuous seminorm on \( V \).

**Lower Semicontinuity.** A function \( F:V \to \mathbb{R} \) is lower semicontinuous (l.s.c.) at a point \( u \in V \) iff
\[ \liminf_{u' \to u} F(u') \geq F(u) \]  
(A.12)
Figure A.3. A convex function.
Thus, \( F \) is l.s.c. at \( u \) if \( F(u) \) is less than or equal to the values of all cluster points of \( F \) at \( u \). For example, the function shown in Fig.A4a is l.s.c. at \( x_0 \in \mathbb{R} \) whereas that in Fig.A4b is not. It is easily shown that \( F \) is l.s.c. on all of \( V \) iff \( \text{epi } F \) is closed (in the topology of \( V \times \mathbb{R} \)). The concept of upper semicontinuity (u.s.c.) is defined in an analogous manner.

**Differentiable Function.** A function \( F:V \to \overline{\mathbb{R}} \) is differentiable (or Gâteaux differentiable) at a point \( u \in V \) iff a unique continuous linear functional \( DF(u) \in V^* \) exists such that

\[
\lim_{\theta \to 0^+} \frac{1}{\theta} (F(u + \theta v) - F(u)) = \left\langle DF(u), v \right\rangle \quad \forall v \in V
\] (A.13)

If \( F:V \to \overline{\mathbb{R}} \) is differentiable on \( V \), then it is easily shown that \( F \) is convex iff

\[
F(v) - F(u) \geq \left\langle DF(u), v-u \right\rangle \quad \forall u,v \in V
\] (A.14)

**Subdifferentials and Subgradients.** Let \( F:V \to \overline{\mathbb{R}} \) be a proper function. The subdifferential of \( F \) at \( u \in V \) is the (possibly empty) subset

\[
\partial F(u) \subseteq V^*
\]

defined by

\[
\partial F(u) = \{ u^* \in V^* \mid F(v) - F(u) \geq \left\langle u^*, v-u \right\rangle \forall v \in V \} \] (A.15)

The elements \( u^* \in \partial F(u) \) are called subgradients of \( F \) at \( u \). Subdifferentials and subgradients have the following properties:
Figure A4. Semicontinuity (and absence of it) of a discontinuous function.
• If $F$ is differentiable at $u$, then its subdifferential consists of only the gradient of $F$,

$$\partial F(u) = \{DF(u)\}$$

• $u^* \in \partial F(u)$ iff

$$\inf_{v \in V} \{F(v) - \langle u^*, v \rangle\} = F(u) - \langle u^*, u \rangle_V$$

Equivalently, the conjugate or polar function $F^*$ of $F$ is defined by

$$F^*: V^* \rightarrow \overline{\mathbb{R}}$$

$$F^*(u^*) = \sup_{v \in V} \{\langle u^*, v \rangle - F(v)\}$$

(A.16)

Thus, $u^* \in \partial F(u)$ whenever

$$F^*(u^*) + F(u) = \langle u^*, v \rangle_V$$

(A.17)

• Any convex function $F$ continuous at $u$ has a non-empty compact subdifferential $\partial F(u)$.

Geometrically, the subdifferential of a continuous convex function $F$, not differentiable at a point $u$, can be regarded as the set of slopes contained in the cone defined by tangents to $F$ at $u$, as illustrated in Fig. A.5.

Indicator Function. Let $K$ be a non-empty subset of $V$. Then the indicator function $\psi_K: V \rightarrow \overline{\mathbb{R}}$ of the set $K$ is defined by

$$\psi_K(u) = \begin{cases} 
0 & \text{if } u \in K \\
\infty & \text{if } u \notin K
\end{cases}$$

(A.18)

If $K$ is convex, $\psi_K$ is a convex function. In this case,
Figure A5. Subdifferential of a non-differentiable function $F$ at a point $u$. 
\[ \Phi_K(u) = \{ u^* \in V \mid \psi_K(v) - \psi_K(u) \geq \langle u^*, v-u \rangle_v \} \] (A.19)

or

\[ \Phi_K(u) = \{ u^* \in V^* \mid u^*, v-u \leq 0 \ \forall \ v \in K \} \] (A.20)