



A THERMODYNAMICALLY CONSISTENT FRAMEWORK FOR THEORIES OF ELASTOPLASTICITY COUPLED WITH DAMAGE

N. R. HANSEN

Sandia National Laboratories, Albuquerque, NM 87115, U.S.A.

and

H. L. SCHREYER

Department of Mechanical Engineering, University of New Mexico, Albuquerque,
NM 87131, U.S.A.

(Received 7 December 1992; in revised form 28 July 1993)

Abstract—A unified framework for coupled elastoplastic and damage theories is developed. A rigorous thermodynamic procedure is followed that is sufficiently general to include anisotropic plasticity and anisotropic damage formulations. The concept of effective stress is the critical mechanism for coupling these theories. Yield and damage functions, constructed of homogeneous functions of degree one, are shown to satisfy thermodynamic restrictions. The principle of maximum entropy provides the evolutionary relations, the loading and unloading conditions, and the convexity of the undamaging elastic domain. The plastic and damage variables evolve normal to their respective surfaces which for plasticity corresponds to an associative flow for plastic strain. This general framework is shown to be sufficiently general to encompass several popular theories for plasticity and damage. Limitations of some existing damage theories are discussed. The performance of two specific coupled formulations are illustrated by replicating the experimental behavior of an aluminum alloy.

1. INTRODUCTION

The ever increasing need to advance the performance and the understanding of material response has increased the emphasis on the proper modeling of inelastic constitutive behavior. The two most popular classes of inelastic material constitutive theories are elastoplasticity and continuum damage mechanics (CDM). These theories have traditionally been used to represent completely different physical phenomena. The theory of plasticity attempts to replicate the dislocation or “slip” of the material at the micro-scale or sub-scale. In contrast, CDM is concerned with the evolution and effective continuum representation of a material with distributed microdefects (microcracks and microvoids). Theories of CDM are generally either based on a micromechanical or a phenomenological approach. The micromechanical technique employs basic mechanics principles, such as well-posed boundary value problems and fracture mechanics on the microscopic scale, to describe the macro behavior. In contrast, the phenomenological theories involve a set of internal variables motivated by experimental observations and then the principles of irreversible thermodynamics are employed. Krajcinovic (1989) has compiled a rather comprehensive review of CDM. In general, materials can exhibit both the damage and the dislocation (plasticity) behavior. The existence of both phenomena has motivated several researchers to couple these theories to form a general description (Simo and Ju, 1987; Chow and Wang, 1987a; Yazdani and Schreyer, 1990; Ju, 1989; Stevens and Liu, 1992; and Yazdani and Karnawat, 1992).

In this paper a thermodynamic framework for a coupled elastoplastic and damage model is developed. An outline of the remaining content of this paper is as follows. In Section 2, some basic thermodynamic concepts and relations are reviewed, prior to the development of the thermodynamic framework. Then the thermodynamic variables and conjugate relations are introduced, followed by the postulated form of the Helmholtz free energy. The concepts of effective stress and effective strain are presented which leads to the introduction of a yield function based in effective-stress space. Both the yield and the

damage functions are required to be constructed of homogeneous functions of degree one. The principle of maximum entropy is used to determine the evolutionary relation for the variables. This principle also guarantees that the undamaging elastic domain is convex. The second law of thermodynamics is shown to be satisfied by all constitutive relations that are compatible with this framework.

In Section 3, some specific models are cast into the general framework. For plasticity the general class of J_2 plasticity is considered. Various forms of hardening are allowed. For the damage portion of the formulation, three popular methods of phenomenologically based damage representation are considered; in particular a scalar, a second-order tensor, and a fourth-order tensor are used to represent the degraded state of the material. The limitations of each of these damage models are discussed. The utility of the developed framework becomes most evident by considering the coupled models. Since the individual models were developed within the established requirements, the coupled formulation is constructed by simply combining models.

Finally in Section 4, two of the coupled formulations introduced in Section 3 are applied to an aluminum alloy. The anisotropic formulations correlate well with the experimental results from three independent researchers for uniaxial tension. The scalar isotropic model is shown to be deficient in representing the damage induced change in the Poisson's ratio. The concept of using a coupled model to represent the temperature dependent ductile-to-brittle transition is also introduced. The simulations replicate the general trends observed in many materials.

2. GENERAL THERMODYNAMIC FRAMEWORK

In this section, a general framework for coupled elastoplasticity and damage formulations is developed following a rigorous thermodynamics approach. This framework follows an irreversible thermodynamic approach using internal variables. The development is preceded by a review of some basic continuum thermodynamic relations.

2.1. Continuum thermodynamics

The thermodynamic relations presented in this section follow the widely accepted approach of internal variable representations given by Coleman and Gurtin (1967) and further elaborated by Lubliner (1972). The restrictive assumptions used in this work are (1) a purely mechanical theory (no internal heat generation sources and heat fluxes) and (2) infinitesimal deformations. The latter restriction allows for an additive decomposition of the strain tensor, $\boldsymbol{\varepsilon}$, into elastic and plastic components, that is $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}^e + \boldsymbol{\varepsilon}^p$. Many of the presented results are either directly applicable or may be generalized to finite deformations.

The internal energy per unit mass, u , at a local continuum point, x , depends on a set of the internal thermodynamic state variables at x . In functional form, the internal energy potential is:

$$u = u[x, s(x), \boldsymbol{\varepsilon}(x), v_i(x)], \quad (1)$$

where s is the entropy per unit mass and v_i is a set of mechanical variables or substates used to model the irreversible or dissipative processes. The explicit notation of dependence on location x is dropped but this dependency is implied throughout the remaining text. The second law of thermodynamics is expressed by a form of the Clausius-Duhem inequality:

$$\rho\theta\dot{s} \geq \rho\dot{u}(s, \boldsymbol{\varepsilon}, v_i) - \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} \quad (2)$$

where ρ is the density, θ is the absolute temperature, and $\boldsymbol{\sigma}$ is the Cauchy stress tensor. The substitution of eqn (1) into (2), yields:

$$\rho \left(\theta - \frac{\partial u}{\partial s} \right) \dot{s} + \left(\boldsymbol{\sigma} - \rho \frac{\partial u}{\partial \boldsymbol{\varepsilon}} \right) : \dot{\boldsymbol{\varepsilon}} - \rho \frac{\partial u}{\partial v_i} \dot{v}_i \geq 0. \quad (3)$$

We exploit the fact that this inequality must hold for all admissible processes. Since \dot{s} and $\dot{\boldsymbol{\varepsilon}}$ are arbitrary, their coefficients must vanish, resulting in two consequences. The absolute temperature, θ , is the thermodynamic variable or force conjugate to the entropy and the stress tensor, $\boldsymbol{\sigma}$, is the thermodynamic force conjugate to the strain tensor, that is :

$$\theta = \frac{\partial u}{\partial s}, \quad \boldsymbol{\sigma} = \rho \frac{\partial u}{\partial \boldsymbol{\varepsilon}}. \quad (4)$$

A thermodynamic variable conjugate to an extensive parameter, such as strain, is often called a thermodynamic force and the time rate of change of the extensive parameter (strain rate) is termed a flux. The final term in eqn (3) is often defined as the dissipation rate, due to the association with the dissipative variables, v_i . The dissipation rate is defined to be the following :

$$\Upsilon \equiv -\rho \frac{\partial u}{\partial v_i} \dot{v}_i. \quad (5)$$

Then the second law reduces to :

$$\Upsilon \geq 0. \quad (6)$$

The Helmholtz free energy is a thermodynamic potential given by a ‘‘contact’’ or Legendre transformation (Callen, 1985) of the internal energy using the conjugate pair (s , θ). The Helmholtz function is :

$$\Psi = u(s, \boldsymbol{\varepsilon}, v_i) - \theta s. \quad (7)$$

By taking the total derivative, the following functional dependency is illustrated :

$$d\Psi = du - d(\theta s) = \frac{\partial u}{\partial s} ds + \frac{\partial u}{\partial \boldsymbol{\varepsilon}} d\boldsymbol{\varepsilon} + \frac{\partial u}{\partial v_i} dv_i - \theta ds - s d\theta = \frac{\partial u}{\partial \boldsymbol{\varepsilon}} d\boldsymbol{\varepsilon} + \frac{\partial u}{\partial v_i} dv_i - s d\theta. \quad (8)$$

Therefore,

$$\Psi = \Psi(\theta, \boldsymbol{\varepsilon}, v_i). \quad (9)$$

Hence, the internal energy is a thermodynamic potential for entropy and the mechanical variables and the Helmholtz function is a potential for temperature and the mechanical variables. Natural choices for isentropic and isothermal processes are the internal and the Helmholtz potentials, respectively.

For purely mechanical theories the first law of thermodynamics or balance of energy yields :

$$\begin{aligned} \rho \dot{u} &= \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} \\ \rho \theta \dot{s} &= \Upsilon. \end{aligned} \quad (10)$$

2.2. Coupled elastoplasticity and damage formulations

The coupled elastoplastic and damage framework is developed in this section. First the internal variables and potentials used to describe the processes are introduced. The concept of transformation or mapping to effective-stress and effective-strain spaces is

introduced. General forms for damage functions and yield functions based in effective-stress space are constructed using homogeneous functions of degree one. Then with the use of the Lagrange minimization method and the principle of maximum entropy, the evolutionary relations are derived. In addition, the convexity of the undamaged elastic domain is shown and the consistency conditions are determined. The satisfaction of the first and second laws is demonstrated. Finally, some specific forms of homogeneous functions of degree one are presented.

2.2.1. *Thermodynamic variables and potentials.* An isothermal process is assumed. Since the plasticity and the damage processes are irreversible, they are by definition dissipative. The excess energy is dissipated in the form of heat. The local generation of heat, which for heterogeneous deformation fields results in a flow of heat, is a direct violation of the stated isothermal assumption. A general formulation, including heat flow and temperature dependent properties, is more complicated. The isothermal assumption is a reasonable approximation when the amount of heat generated is relatively small, or when a process occurs rapidly and the material parameters are relatively insensitive to changes in temperature.

To describe the irreversibility associated with the plastic and the damaging processes, a set of variables is introduced. For plasticity, let $\boldsymbol{\varepsilon}^P$ be the plastic-strain tensor and introduce two second-order tensor variables in strain space that will be used to describe the plastic hardening phenomena. Let $\boldsymbol{\varepsilon}^S$ be the hardening variable that describes the shift in the center of the yield surface (a kinematic type hardening) and $\boldsymbol{\varepsilon}^H$ be the hardening variable that describes the shape and size of the yield surface (such as isotropic hardening). For the damage process, introduce a generalized damage tensor, \mathbf{D} , that is some measure of degradation of the material integrity. At this time the rank of the tensors associated with damage will remain unspecified to allow the following framework to be applicable to a large class of damage theories. Let \mathbf{D}^S be the damaging variable which describes the shift in the center of the damage surface and that is of the same tensor order as \mathbf{D} . Furthermore, let \mathbf{D}^H be the damaging variable that describes the shape and size of the damage surface. The final assumption is that of rate independence. With these definitions and assumptions, the functional form of the Helmholtz free energy is:

$$\Psi = \Psi(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^P, \boldsymbol{\varepsilon}^S, \boldsymbol{\varepsilon}^H, \mathbf{D}, \mathbf{D}^S, \mathbf{D}^H). \quad (11)$$

The conjugate thermodynamic forces are defined by:

$$\begin{aligned} \boldsymbol{\sigma}^P &\equiv -\rho \frac{\partial \Psi}{\partial \boldsymbol{\varepsilon}^P} & \mathbf{Y} &\equiv -\rho \frac{\partial \Psi}{\partial \mathbf{D}} \\ \boldsymbol{\sigma}^S &\equiv -\rho \frac{\partial \Psi}{\partial \boldsymbol{\varepsilon}^S} & \mathbf{Y}^S &\equiv -\rho \frac{\partial \Psi}{\partial \mathbf{D}^S} \\ \boldsymbol{\sigma}^H &\equiv -\rho \frac{\partial \Psi}{\partial \boldsymbol{\varepsilon}^H} & \mathbf{Y}^H &\equiv -\rho \frac{\partial \Psi}{\partial \mathbf{D}^H}. \end{aligned} \quad (12)$$

The internal energy takes the following form according to eqn (7):

$$\rho u(s, \boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^P, \boldsymbol{\varepsilon}^S, \boldsymbol{\varepsilon}^H, \mathbf{D}, \mathbf{D}^S, \mathbf{D}^H) = \rho \Psi(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^P, \boldsymbol{\varepsilon}^S, \boldsymbol{\varepsilon}^H, \mathbf{D}, \mathbf{D}^S, \mathbf{D}^H) + \rho \theta s. \quad (13)$$

The Helmholtz free energy is postulated to be separable as follows:

$$\rho \Psi(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^P, \boldsymbol{\varepsilon}^S, \boldsymbol{\varepsilon}^H, \mathbf{D}, \mathbf{D}^S, \mathbf{D}^H) \equiv W(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^P, \mathbf{D}) + H(\boldsymbol{\varepsilon}^S, \boldsymbol{\varepsilon}^H, \mathbf{D}^S, \mathbf{D}^H) + G_D(\mathbf{D}), \quad (14)$$

where the stored or elastic energy function is defined to be:

$$W(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^p, \mathbf{D}) \equiv \frac{1}{2}(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p) : \mathbf{E}(\mathbf{D}) : (\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p). \quad (15)$$

The exact form of the damage stiffness tensor, $\mathbf{E}(\mathbf{D})$, depends on the specific damage representation theory employed, several of which are presented in later sections. The fourth-order stiffness tensor is initially denoted by \mathbf{E}^o in the undamaged state. The term $H(\boldsymbol{\varepsilon}^S, \boldsymbol{\varepsilon}^H, \mathbf{D}^S, \mathbf{D}^H)$ is the contribution of the hardening and damaging variables to the Helmholtz potential. This is an unspecified function that will be resolved later by postulating constitutive relations for these variables.

If the constitutive relations for the hardening and damaging variables are strict functions of their conjugate pair, for example $\boldsymbol{\sigma}^S = \boldsymbol{\sigma}^S(\boldsymbol{\varepsilon}^S)$, then from eqns (12) and (14):

$$\begin{aligned} -H(\boldsymbol{\varepsilon}^S, \boldsymbol{\varepsilon}^H, \mathbf{D}^S, \mathbf{D}^H) = & \int \boldsymbol{\sigma}^S(\boldsymbol{\varepsilon}^S) : d\boldsymbol{\varepsilon}^S + \int \boldsymbol{\sigma}^H(\boldsymbol{\varepsilon}^H) : d\boldsymbol{\varepsilon}^H + \int \mathbf{Y}^S(\mathbf{D}^S) : d\mathbf{D}^S \\ & + \int \mathbf{Y}^H(\mathbf{D}^H) : d\mathbf{D}^H. \quad (16) \end{aligned}$$

The final contribution to the Helmholtz potential is the surface energy term, $G_D(\mathbf{D})$. Damage, or the generation and propagation of microdefects in the material, causes microcracks and microspheres to grow. A fundamental thermodynamic principle, used in Griffith crack theory, for example, states that the increase in the size of material surfaces corresponds to an increase in the material surface energy; thus the motivation for this term in the Helmholtz potential. For theories employing a second-order tensor as the damage measure, it seems reasonable to assume the surface energy term is proportional to a second-invariant of the damage tensor,

$$G_D(\mathbf{D}) \equiv \gamma_D \mathbf{D} : \mathbf{D}, \quad (17)$$

where γ_D is a constant. Other forms can likewise be postulated for other types of damage measures but the proper form can only be determined by the application of micromechanical concepts with experimental confirmation. The inclusion of the surface energy term in the Helmholtz potential implies that some of the energy required to damage the material is converted to surface energy and the remaining portion of that energy is dissipated. This is contrary to a hypothesis in which the damaging process totally converts the mechanical energy to surface energy, an approach that is not followed here.

The form of the Helmholtz function given in eqn (14) assumes an additive decomposition into a stored elastic energy term and additional terms related to the hardening and damaging variables. The approach may be too restrictive but it is used by others because the formulation is still fairly general but not so abstract that the derivations become overly complicated. An attractive feature of this form is that the conjugate variables $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}^p$ yield classical relations. From eqns (4), (14), and (15):

$$\boldsymbol{\sigma} = \mathbf{E}(\mathbf{D}) : (\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p) \quad \boldsymbol{\sigma}^p = \boldsymbol{\sigma}. \quad (18)$$

In contrast, if $H = H(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^S, \boldsymbol{\varepsilon}^H, \mathbf{D}^S, \mathbf{D}^H)$, then the stress takes the form given by the following:

$$\boldsymbol{\sigma} = \mathbf{E}(\mathbf{D}) : (\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p) + \frac{\partial H(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^S, \boldsymbol{\varepsilon}^H, \mathbf{D}^S, \mathbf{D}^H)}{\partial \boldsymbol{\varepsilon}}, \quad (19)$$

which is more complicated and the resulting constitutive relation for stress is at variance with the classical form. Alternate forms of the Helmholtz potential will yield similar consequences according to the conjugate relations presented in eqn (12). Finally, from eqn (5) and with the choice of the variables associated with the irreversible processes as $v_i = \{\boldsymbol{\varepsilon}^p, \boldsymbol{\varepsilon}^S, \boldsymbol{\varepsilon}^H, \mathbf{D}, \mathbf{D}^S, \mathbf{D}^H\}$, the dissipation rate is:

$$\Upsilon = -\rho \left(\frac{\partial \Psi}{\partial \boldsymbol{\varepsilon}^P} : \dot{\boldsymbol{\varepsilon}}^P + \frac{\partial \Psi}{\partial \boldsymbol{\varepsilon}^S} : \dot{\boldsymbol{\varepsilon}}^S + \frac{\partial \Psi}{\partial \boldsymbol{\varepsilon}^H} : \dot{\boldsymbol{\varepsilon}}^H + \frac{\partial \Psi}{\partial \mathbf{D}} : \dot{\mathbf{D}} + \frac{\partial \Psi}{\partial \mathbf{D}^S} : \dot{\mathbf{D}}^S + \frac{\partial \Psi}{\partial \mathbf{D}^H} : \dot{\mathbf{D}}^H \right). \quad (20)$$

By the conjugate force relations (12), the dissipation rate can be rewritten as :

$$\Upsilon = \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}^P + \boldsymbol{\sigma}^S : \dot{\boldsymbol{\varepsilon}}^S + \boldsymbol{\sigma}^H : \dot{\boldsymbol{\varepsilon}}^H + \mathbf{Y} : \dot{\mathbf{D}} + \mathbf{Y}^S : \dot{\mathbf{D}}^S + \mathbf{Y}^H : \dot{\mathbf{D}}^H. \quad (21)$$

2.2.2. *Effective stress and effective strain.* As the material becomes damaged, stress at the subscale becomes magnified due to the diminished material integrity. This subscale stress, called the effective stress, was first introduced by Kachanov (1958), and is the foundation for the field of continuum damage mechanics. It is in the effective continuum where the plasticity process evolves. Hence, the effective stress is the essential mechanism by which theories of elastoplasticity are coupled with damage theories, and therefore, a proper effective-stress relation is critical. The effective-stress tensor, $\hat{\boldsymbol{\sigma}}$, can in general be represented by a projection of the Cauchy stress tensor :

$$\hat{\boldsymbol{\sigma}} = \mathbf{M}(\mathbf{D}) : \boldsymbol{\sigma}, \quad (22)$$

where $\mathbf{M}(\mathbf{D})$ is the effective-stress operator (a fourth-order tensor) that is a function of the damage state. A specific form of this operator depends on the damage representation theory that can either be derived from a phenomenological or a micromechanical approach. Some examples are presented in later sections. The stress-space variables associated with the plastic hardening are assumed to be mapped into effective-stress space by the same operator :

$$\hat{\boldsymbol{\sigma}}^S \equiv \mathbf{M}(\mathbf{D}) : \boldsymbol{\sigma}^S, \quad \hat{\boldsymbol{\sigma}}^H \equiv \mathbf{M}(\mathbf{D}) : \boldsymbol{\sigma}^H. \quad (23)$$

The mapping of the stress to the effective stress is required for later derivations, but the mapping of $\boldsymbol{\sigma}^S$ and $\boldsymbol{\sigma}^H$ into effective-stress space is not required for the thermodynamic derivations to hold. In fact, equivalent results are obtained if these variables are not transformed according to eqn (23). The mapping was assumed for a purely conceptual reason.

If the inverse of the effective-stress operator, \mathbf{M}^{-1} , exists, then the dissipation rate as given in eqn (21) can be rewritten using the effective-stress operator as :

$$\Upsilon = \boldsymbol{\sigma} : (\mathbf{M}^T : \mathbf{M}^{-T}) : \dot{\boldsymbol{\varepsilon}}^P + \boldsymbol{\sigma}^S : (\mathbf{M}^T : \mathbf{M}^{-T}) : \dot{\boldsymbol{\varepsilon}}^S + \boldsymbol{\sigma}^H : (\mathbf{M}^T : \mathbf{M}^{-T}) : \dot{\boldsymbol{\varepsilon}}^H + \mathbf{Y} : \dot{\mathbf{D}} + \mathbf{Y}^S : \dot{\mathbf{D}}^S + \mathbf{Y}^H : \dot{\mathbf{D}}^H. \quad (24)$$

Motivated by this form of the dissipation rate, the rate forms of the plasticity variables based in strain space are assumed to be mapped into an effective space by the inverse of the effective-stress operator, that is :

$$\hat{\boldsymbol{\varepsilon}}^P \equiv \mathbf{M}^{-T}(\mathbf{D}) : \dot{\boldsymbol{\varepsilon}}^P, \quad \hat{\boldsymbol{\varepsilon}}^S \equiv \mathbf{M}^{-T}(\mathbf{D}) : \dot{\boldsymbol{\varepsilon}}^S, \quad \hat{\boldsymbol{\varepsilon}}^H \equiv \mathbf{M}^{-T}(\mathbf{D}) : \dot{\boldsymbol{\varepsilon}}^H \quad (25)$$

and then the dissipation rate becomes :

$$\Upsilon = \hat{\boldsymbol{\sigma}} : \hat{\boldsymbol{\varepsilon}}^P + \hat{\boldsymbol{\sigma}}^S : \hat{\boldsymbol{\varepsilon}}^S + \hat{\boldsymbol{\sigma}}^H : \hat{\boldsymbol{\varepsilon}}^H + \mathbf{Y} : \dot{\mathbf{D}} + \mathbf{Y}^S : \dot{\mathbf{D}}^S + \mathbf{Y}^H : \dot{\mathbf{D}}^H. \quad (26)$$

The definition of the effective strains follows naturally from eqn (25). The effective-plastic strain is given by :

$$\tilde{\boldsymbol{\varepsilon}}^P = \int (\dot{\boldsymbol{\varepsilon}}^P) dt = \int (\mathbf{M}^{-T} : \dot{\boldsymbol{\varepsilon}}^P) dt \quad (27)$$

and similar relations hold for $\tilde{\boldsymbol{\varepsilon}}^S$ and $\tilde{\boldsymbol{\varepsilon}}^H$. If an operator, \mathbf{N}^P , corresponding to effective-plastic strain, is defined such that :

$$\tilde{\boldsymbol{\varepsilon}}^P \equiv \mathbf{N}^P : \boldsymbol{\varepsilon}^P \quad (28)$$

then the relation between the effective-space operators is given by :

$$\mathbf{N}^P : \boldsymbol{\varepsilon}^P = \int (\mathbf{M}^{-T} : \dot{\boldsymbol{\varepsilon}}^P) dt. \quad (29)$$

This definition for the effective strain is at variance with the definitions used in other phenomenological theories for damage representation as presented in Section 3.2. As will be shown, the present approach, which follows as a natural consequence of the dissipation rate, produces a thermodynamically rigorous framework without the need for additional assumptions. For practical applications the effective strains are not required in an explicit form. Once the increment in effective-plastic strain is determined, then it is immediately transformed, according to eqn (25), to a plastic-strain increment in actual space and the total plastic strain is updated.

A consequence of the effective-strain relations developed above is that thermodynamic variables that are conjugate in actual-stress and actual-strain space are no longer conjugate in effective-stress and effective-strain space. This is shown by the following :

$$\begin{aligned} \boldsymbol{\sigma}^H &\equiv -\rho \frac{\partial \Psi}{\partial \boldsymbol{\varepsilon}^H} = -\rho \frac{\partial \Psi}{\partial \hat{\boldsymbol{\varepsilon}}^H} : \frac{\partial \hat{\boldsymbol{\varepsilon}}^H}{\partial \boldsymbol{\varepsilon}^H} = -\rho \mathbf{N}^H : \frac{\partial \Psi}{\partial \hat{\boldsymbol{\varepsilon}}^H} \\ \hat{\boldsymbol{\sigma}}^H &\equiv \mathbf{M} : \boldsymbol{\sigma}^H = -\rho \mathbf{M} : \mathbf{N}^H : \frac{\partial \Psi}{\partial \hat{\boldsymbol{\varepsilon}}^H} \neq -\rho \frac{\partial \Psi}{\partial \hat{\boldsymbol{\varepsilon}}^H}. \end{aligned} \quad (30)$$

2.2.3. *Effective-stress-space yield function and damage function.* A generalized yield function, Φ_p , that separates the elastic and elastoplastic domains in effective-stress space is assumed to be of the form :

$$\Phi_p(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\sigma}}^S, \hat{\boldsymbol{\sigma}}^H) \equiv \phi_{p1}(\hat{\boldsymbol{\sigma}} - \hat{\boldsymbol{\sigma}}^S) - [\phi_{p2}(\hat{\boldsymbol{\sigma}}^H) + \sigma_y] \leq 0, \quad (31)$$

where σ_y is a positive scalar material parameter used to describe the onset of plastic behavior (an initial yield stress). The scalar-valued tensor functions, $\phi_{p1}(\mathbf{X})$ and $\phi_{p2}(\mathbf{X})$, are required to be homogeneous and of degree one. A function $F(x, y)$ is said to be homogeneous of degree n provided it satisfies the condition :

$$F(\alpha x, \alpha y) = \alpha^n F(x, y). \quad (32)$$

A fundamental property of such functions is given by Euler's theorem, Davis (1960), which states that if $\phi_{pi}(\mathbf{X})$ is a homogeneous function of degree one then the following is satisfied :

$$\left(\frac{\partial \phi_{pi}(\mathbf{X})}{\partial \mathbf{X}} \right) : \mathbf{X} = \phi_{pi}(\mathbf{X}) \quad \forall \mathbf{X}. \quad (33)$$

Homogeneous functions of degree one will hereafter be referred to as HODO functions. By exploiting the chain rule, the following two properties can be derived :

$$\left(\frac{\partial \phi_{p1}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^S)}{\partial \boldsymbol{\sigma}} \right) : (\boldsymbol{\sigma} - \boldsymbol{\sigma}^S) = \phi_{p1}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^S)$$

$$\frac{\partial \phi_{p1}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^S)}{\partial \boldsymbol{\sigma}} = - \left(\frac{\partial \phi_{p1}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^S)}{\partial \boldsymbol{\sigma}^S} \right). \quad (34)$$

In a similar manner, consider a damage criterion that takes the following form :

$$\Phi_D(\mathbf{Y}, \mathbf{Y}^S, \mathbf{Y}^H) \equiv \phi_{D1}(\mathbf{Y} - \mathbf{Y}^S) - [\phi_{D2}(\mathbf{Y}^H) + \omega_0] \leq 0 \quad (35)$$

where ω_0 is a positive scalar material parameter used to describe the onset of damage behavior (a damage energy threshold). The scalar-valued tensor functions, $\phi_{D1}(\mathbf{X})$ and $\phi_{D2}(\mathbf{X})$, are likewise required to be homogeneous of degree one, HODO, as defined above. The use of a tensor, \mathbf{Y}^H , equal in tensor order to the damage variable is required for a general anisotropic description of the shape of the damage surface. Many anisotropic damage theories have been proposed that employ only a scalar variable to describe the shape of the damage surface. Among others, these include Simo and Ju (1987), Ju (1989), Chow and Wang (1987a, b), and Yazdani and Schreyer (1990). A scalar can only describe an isotropic surface or equal damage evolution in all directions which is inconsistent with the use of an anisotropic description of damage.

The above presentation implies two independent dissipation criteria, one for plasticity and the other for damage. This is consistent with the formulations of Simo and Ju (1987) and Ju (1989), among several others. Some researchers have employed one surface for both plasticity and damage, among them Stevens and Liu (1992). In addition, Yazdani and Karnawat (1992) have combined the two surfaces with a pressure dependence. The presented form of independent surfaces allows for the greatest flexibility in simulating the complete spectrum of material behavior.

2.2.4. *Principle of maximum entropy.* There appears to be some confusion of terminology in the literature as it relates to extremum principles for constitutive relations (Simo and Hughes, 1988 ; Lubliner, 1984 ; Ziegler, 1963). In the field of irreversible thermodynamics two principles are widely accepted ; namely at a stable equilibrium state 1) the entropy production rate is minimized while 2) the entropy is a maximum for a given total energy (Callen, 1985). Recall from eqn (10)₂ that the entropy production is proportional to the dissipation rate for purely mechanical processes so these terms are used interchangeably. The entropy production is defined as the derivative with respect to time at a state point. For incremental inelastic theories, such as plasticity, a variable with superimposed (·) really represents a finite increment from one stable state to the next stable state. Such a change might be more appropriately denoted by $\Delta(\cdot)$, but the common notation in the literature is as presented. However, the above thermodynamic principles still apply to these two states. Since the entropy is maximum at each of the two states the change in the entropy between the two states is also a maximum. Here lies the confusion in the literature. The entropy production at each state is a minimum, but the change in entropy in going between the states is a maximum. All inelastic theories presented in this work are incremental theories where each quantity denoted by (·) really represents a finite change between two stable states. The (·) should not be confused with instantaneous time derivatives.

Consider the actual state, \mathbf{A} , where $\mathbf{A} \equiv \{\boldsymbol{\sigma}, \boldsymbol{\sigma}^S, \boldsymbol{\sigma}^H, \mathbf{Y}, \mathbf{Y}^S, \mathbf{Y}^H\}$ which corresponds to the state, $\hat{\mathbf{A}}$, at the subscale or in effective-stress space, where $\hat{\mathbf{A}} \equiv \{\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\sigma}}^S, \hat{\boldsymbol{\sigma}}^H, \mathbf{Y}, \mathbf{Y}^S, \mathbf{Y}^H\}$. Now consider all admissible states for $\hat{\mathbf{A}}$, denoted by $\hat{\mathbf{X}} \equiv \{\mathbf{R}, \mathbf{S}, \mathbf{T}, \mathbf{U}, \mathbf{V}, \mathbf{W}\}$. The object is to determine the state $\hat{\mathbf{X}}$ that maximizes the dissipation or entropy production (between two stable states) subject to the constraints of both the yield and the damage conditions. The problem is essentially one of constrained optimization. This can be solved using the Lagrange minimization method. First, introduce a Lagrangian functional :

$$\begin{aligned} \mathbf{L}(\hat{\mathbf{X}}) &\equiv -\Upsilon(\hat{\mathbf{X}}) + \dot{\lambda}_P \Phi_P(\hat{\mathbf{X}}_P) + \dot{\lambda}_D \Phi_D(\hat{\mathbf{X}}_D) \\ &= -(\mathbf{R} : \dot{\hat{\boldsymbol{\epsilon}}}^P + \mathbf{S} : \dot{\hat{\boldsymbol{\epsilon}}}^S + \mathbf{T} : \dot{\hat{\boldsymbol{\epsilon}}}^H + \mathbf{U} : \dot{\mathbf{D}} + \mathbf{V} : \dot{\mathbf{D}}^S + \mathbf{W} : \dot{\mathbf{D}}^H) + \dot{\lambda}_P \Phi_P(\hat{\mathbf{X}}_P) + \dot{\lambda}_D \Phi_D(\hat{\mathbf{X}}_D), \end{aligned} \quad (36)$$

where the Lagrange multipliers, $\dot{\lambda}_P$ and $\dot{\lambda}_D$, are associated with the yield and damage constraints, respectively. The subscripts P and D distinguish the subsets of variables associated with plasticity and damage, respectively. The Lagrange multiplier method determines the minimum of a functional. Since the maximum is desired, a negative sign is included on the dissipation term. The state that minimizes the functional is obtained by the Kuhn-Tucker optimality condition (Strang, 1986). This intermediate result, which is not presented, is a function of the admissible states, \mathbf{X} . The principle of maximum dissipation is employed to determine which admissible state, \mathbf{X} , is the actual state, \mathbf{A} . This principle is credited to von Mises for plasticity (Hill, 1950), and is a basic governing postulate in thermodynamics. It states: that amongst all admissible states, $\hat{\mathbf{X}}$, the actual state, $\hat{\mathbf{A}}$, is the state attained that maximizes the entropy. Hence, state $\hat{\mathbf{A}}$, in conjunction with the results of maximization, yield the following evolution relations:

$$\begin{aligned} \dot{\hat{\boldsymbol{\epsilon}}}^P &= \dot{\lambda}_P \frac{\partial \Phi_P(\hat{\mathbf{A}}_P)}{\partial \hat{\boldsymbol{\sigma}}} = \dot{\lambda}_P \frac{\partial \phi_{P1}(\hat{\boldsymbol{\sigma}} - \hat{\boldsymbol{\sigma}}^S)}{\partial \hat{\boldsymbol{\sigma}}} \\ \dot{\hat{\boldsymbol{\epsilon}}}^S &= \dot{\lambda}_P \frac{\partial \Phi_P(\hat{\mathbf{A}}_P)}{\partial \hat{\boldsymbol{\sigma}}^S} = \dot{\lambda}_P \frac{\partial \phi_{P1}(\hat{\boldsymbol{\sigma}} - \hat{\boldsymbol{\sigma}}^S)}{\partial \hat{\boldsymbol{\sigma}}^S} \\ \dot{\hat{\boldsymbol{\epsilon}}}^H &= \dot{\lambda}_P \frac{\partial \Phi_P(\hat{\mathbf{A}}_P)}{\partial \hat{\boldsymbol{\sigma}}^H} = -\dot{\lambda}_P \frac{\partial \phi_{P2}(\hat{\boldsymbol{\sigma}}^H)}{\partial \hat{\boldsymbol{\sigma}}^H} \\ \dot{\mathbf{D}} &= \dot{\lambda}_D \frac{\partial \Phi_D(\hat{\mathbf{A}}_D)}{\partial \mathbf{Y}} = \dot{\lambda}_D \frac{\partial \phi_{D1}(\mathbf{Y} - \mathbf{Y}^S)}{\partial \mathbf{Y}} \\ \dot{\mathbf{D}}^S &= \dot{\lambda}_D \frac{\partial \Phi_D(\hat{\mathbf{A}}_D)}{\partial \mathbf{Y}^S} = \dot{\lambda}_D \frac{\partial \phi_{D1}(\mathbf{Y} - \mathbf{Y}^S)}{\partial \mathbf{Y}^S} \\ \dot{\mathbf{D}}^H &= \dot{\lambda}_D \frac{\partial \Phi_D(\hat{\mathbf{A}}_D)}{\partial \mathbf{Y}^H} = \dot{\lambda}_D \frac{\partial \phi_{D2}(\mathbf{Y}^H)}{\partial \mathbf{Y}^H}. \end{aligned} \quad (37)$$

In addition, the unloading and loading conditions follow directly from the Kuhn-Tucker optimality conditions. These are as follows:

$$\begin{aligned} \dot{\lambda}_P &\geq 0 & \dot{\lambda}_D &\geq 0 \\ \Phi_P(\hat{\mathbf{A}}_P) &\leq 0 & \Phi_D(\hat{\mathbf{A}}_D) &\leq 0 \\ \dot{\lambda}_P \Phi_P(\hat{\mathbf{A}}_P) &= 0 & \dot{\lambda}_D \Phi_D(\hat{\mathbf{A}}_D) &= 0. \end{aligned} \quad (38)$$

All variables evolve in the direction normal to the respective yield and damage surfaces. Plastic-strain evolution of this form is typically called associative flow. The plasticity variables evolve in effective-strain space and are associative in the effective-stress space. This associativity is preserved in the actual-stress space, as is illustrated by the following:

$$\dot{\hat{\boldsymbol{\epsilon}}}^P = \mathbf{M}^T : \dot{\hat{\boldsymbol{\epsilon}}}^P = \dot{\lambda}_P \mathbf{M}^T : \frac{\partial \Phi_P}{\partial \hat{\boldsymbol{\sigma}}} = \dot{\lambda}_P \mathbf{M}^T : \frac{\partial \Phi_P}{\partial \boldsymbol{\sigma}} : \left(\frac{\partial \boldsymbol{\sigma}}{\partial \hat{\boldsymbol{\sigma}}} \right) = \dot{\lambda}_P \mathbf{M}^T : \frac{\partial \Phi_P}{\partial \boldsymbol{\sigma}} : \mathbf{M}^{-1} = \dot{\lambda}_P \frac{\partial \Phi_P}{\partial \boldsymbol{\sigma}}. \quad (39)$$

With the property derived in eqn (34), the second and fifth relations of eqn (37) reduce to:

$$\dot{\hat{\boldsymbol{\epsilon}}}^S = -\dot{\hat{\boldsymbol{\epsilon}}}^P \quad \dot{\hat{\mathbf{D}}}^S = -\dot{\hat{\mathbf{D}}}. \quad (40)$$

An additional property that can be derived using the principle of maximum entropy is the convexity of the undamaged elastic domain. First consider a definition of a smooth convex function as given by Simo and Hughes (1988). A smooth function, f , is convex *if and only if* the following holds:

$$f(\mathbf{r}) - f(\mathbf{s}) \geq (\mathbf{r} - \mathbf{s}) \cdot \nabla f(\mathbf{s}). \quad (41)$$

Suppose the state $\hat{\mathbf{A}}$ is on both the yield and damage surfaces. By definition, this implies that:

$$\Phi_P(\hat{\mathbf{A}}_P) = \Phi_D(\hat{\mathbf{A}}_D) = 0. \quad (42)$$

Then for all admissible states:

$$\Phi_P(\hat{\mathbf{X}}_P) \leq 0, \quad \Phi_D(\hat{\mathbf{X}}_D) \leq 0. \quad (43)$$

With $\hat{\mathbf{X}}$ and $\hat{\mathbf{A}}$ associated with \mathbf{r} and \mathbf{s} , respectively, the plastic condition for convexity is:

$$\Phi_P(\hat{\mathbf{X}}_P) - \Phi_P(\hat{\mathbf{A}}_P) \geq (\hat{\mathbf{X}}_P - \hat{\mathbf{A}}_P) : \nabla \Phi_P(\hat{\mathbf{A}}_P). \quad (44)$$

By eqns (42) and (43), the left side of the inequality can be simplified and the plastic convexity condition becomes:

$$0 \geq \Phi_P(\hat{\mathbf{X}}_P) \geq (\hat{\mathbf{X}}_P - \hat{\mathbf{A}}_P) : \nabla \Phi_P(\hat{\mathbf{A}}_P) \Rightarrow 0 \geq (\hat{\mathbf{X}}_P - \hat{\mathbf{A}}_P) : \nabla \Phi_P(\hat{\mathbf{A}}_P). \quad (45)$$

By similar arguments the damage convexity condition is:

$$0 \geq (\hat{\mathbf{X}}_D - \hat{\mathbf{A}}_D) : \nabla \Phi_D(\hat{\mathbf{A}}_D). \quad (46)$$

The principle of maximum entropy can be stated in the following form:

$$\Upsilon(\hat{\mathbf{A}}) \geq \Upsilon(\hat{\mathbf{X}}). \quad (47)$$

The dissipation rate and the evolutionary relations of eqns (26) and (37), respectively, allow eqn (47) to be rewritten as:

$$0 \geq \dot{\lambda}_P \left((\mathbf{R} - \hat{\boldsymbol{\sigma}}) : \frac{\partial \Phi_P(\hat{\mathbf{A}}_P)}{\partial \hat{\boldsymbol{\sigma}}} + (\mathbf{S} - \hat{\boldsymbol{\sigma}}^S) : \frac{\partial \Phi_P(\hat{\mathbf{A}}_P)}{\partial \hat{\boldsymbol{\sigma}}^S} + (\mathbf{T} - \hat{\boldsymbol{\sigma}}^H) : \frac{\partial \Phi_P(\hat{\mathbf{A}}_P)}{\partial \hat{\boldsymbol{\sigma}}^H} \right) \\ + \dot{\lambda}_D \left((\mathbf{U} - \mathbf{Y}) : \frac{\partial \Phi_D(\hat{\mathbf{A}}_D)}{\partial \mathbf{Y}} + (\mathbf{V} - \mathbf{Y}^S) : \frac{\partial \Phi_D(\hat{\mathbf{A}}_D)}{\partial \mathbf{Y}^S} + (\mathbf{W} - \mathbf{Y}^H) : \frac{\partial \Phi_D(\hat{\mathbf{A}}_D)}{\partial \mathbf{Y}^H} \right). \quad (48)$$

Since this inequality must hold for arbitrary processes, the two terms can be uncoupled without loss of generality. In addition, employ the notation of $\hat{\mathbf{X}}$ and $\hat{\mathbf{A}}$, so eqn (48) becomes:

$$0 \geq \dot{\lambda}_P (\hat{\mathbf{X}}_P - \hat{\mathbf{A}}_P) : \nabla \Phi_P(\hat{\mathbf{A}}_P) \\ 0 \geq \dot{\lambda}_D (\hat{\mathbf{X}}_D - \hat{\mathbf{A}}_D) : \nabla \Phi_D(\hat{\mathbf{A}}_D). \quad (49)$$

Since both of the consistency parameters are non-negative, eqns (49) reduce exactly to the convexity condition of eqns (45) and (46). Hence, the yield and the damage functions are convex or the convexity of the undamaging elastic domain is shown.

The parameters $\dot{\lambda}_P$ and $\dot{\lambda}_D$ are determined by the consistency requirements, which are:

$$\dot{\lambda}_P \dot{\Phi}_P(\hat{\mathbf{A}}_P) = 0, \quad \dot{\lambda}_D \dot{\Phi}_D(\hat{\mathbf{A}}_D) = 0. \quad (50)$$

If $\dot{\lambda}_P = 0$, which implies that no plastic processes are occurring, then the state is inside the yield surface, $\dot{\Phi}_P(\hat{\mathbf{A}}_P) < 0$, but if $\dot{\lambda}_P > 0$ or plastic flow is occurring then by the consistency condition :

$$\dot{\Phi}_P(\hat{\mathbf{A}}_P) = 0 = \frac{\partial \Phi_P}{\partial \hat{\boldsymbol{\sigma}}} : \dot{\boldsymbol{\sigma}} + \frac{\partial \Phi_P}{\partial \hat{\boldsymbol{\sigma}}^S} : \dot{\boldsymbol{\sigma}}^S + \frac{\partial \Phi_P}{\partial \hat{\boldsymbol{\sigma}}^H} : \dot{\boldsymbol{\sigma}}^H. \quad (51)$$

If the plasticity hardening variables are assumed to be strict functions of their conjugate pairs, $\hat{\boldsymbol{\sigma}}^S = \hat{\boldsymbol{\sigma}}^S(\hat{\boldsymbol{\varepsilon}}^S)$ and $\hat{\boldsymbol{\sigma}}^H = \hat{\boldsymbol{\sigma}}^H(\hat{\boldsymbol{\varepsilon}}^H)$, then it can be shown that the consistency condition for plasticity reduces to the following :

$$\begin{aligned} \dot{\lambda}_P \left(\frac{\partial \Phi_P}{\partial \hat{\boldsymbol{\sigma}}^H} : \frac{\partial \hat{\boldsymbol{\sigma}}^H}{\partial \hat{\boldsymbol{\varepsilon}}^H} : \frac{\partial \Phi_P}{\partial \hat{\boldsymbol{\sigma}}^H} : + \frac{\partial \Phi_P}{\partial \hat{\boldsymbol{\sigma}}} : \frac{\partial \hat{\boldsymbol{\sigma}}^S}{\partial \hat{\boldsymbol{\varepsilon}}^S} : \frac{\partial \Phi_P}{\partial \hat{\boldsymbol{\sigma}}} - \frac{\partial \Phi_P}{\partial \hat{\boldsymbol{\sigma}}} : \mathbf{M} : \mathbf{E} : \mathbf{M} : \frac{\partial \Phi_P}{\partial \hat{\boldsymbol{\sigma}}} \right) \\ + \dot{\lambda}_D \left(\frac{\partial \Phi_P}{\partial \hat{\boldsymbol{\sigma}}} : \mathbf{M} : \frac{\partial^2 \Psi}{\partial \boldsymbol{\varepsilon} \partial \mathbf{D}} : \frac{\partial \Phi_D}{\partial \mathbf{Y}} \right) + \frac{\partial \Phi_P}{\partial \hat{\boldsymbol{\sigma}}} : \mathbf{M} : \mathbf{E} : \dot{\boldsymbol{\varepsilon}} = 0. \end{aligned} \quad (52)$$

By a completely analogous procedure for the damage consistency condition along with $\mathbf{Y}^S = \mathbf{Y}^S(\mathbf{D}^S)$ and $\mathbf{Y}^H = \mathbf{Y}^H(\mathbf{D}^H)$ it can be shown that :

$$\begin{aligned} \dot{\lambda}_D \left(\frac{\partial \Phi_D}{\partial \mathbf{Y}} : \frac{\partial \mathbf{Y}^S}{\partial \mathbf{D}^S} : \frac{\partial \Phi_D}{\partial \mathbf{Y}} + \frac{\partial \Phi_D}{\partial \mathbf{Y}^H} : \frac{\partial \mathbf{Y}^H}{\partial \mathbf{D}^H} : \frac{\partial \Phi_D}{\partial \mathbf{Y}^H} - \frac{\partial \Phi_D}{\partial \mathbf{Y}} : \frac{\partial^2 \Psi}{\partial \mathbf{D}^2} : \frac{\partial \Phi_D}{\partial \mathbf{Y}} \right) \\ + \dot{\lambda}_P \left(\frac{\partial \Phi_D}{\partial \mathbf{Y}} : \frac{\partial^2 \Psi}{\partial \mathbf{D} \partial \boldsymbol{\varepsilon}^P} : \mathbf{M} : \frac{\partial \Phi_P}{\partial \hat{\boldsymbol{\sigma}}} \right) - \frac{\partial \Phi_D}{\partial \mathbf{Y}} : \frac{\partial^2 \Psi}{\partial \boldsymbol{\varepsilon} \partial \mathbf{D}} : \dot{\boldsymbol{\varepsilon}} = 0. \end{aligned} \quad (53)$$

The plasticity consistency condition, eqn (52), is a function of the damage consistency parameter, $\dot{\lambda}_D$, and conversely the damage consistency condition is a function of the plasticity consistency parameter. This coupling of the consistency relations is a consequence of a coupled formulation. There are two linear equations in two unknowns so that these consistency parameters can be solved explicitly as a function of the strain increment.

Finally, the tangent modulus is derived using these relations for the consistency parameters. The tangent modulus maps a strain increment into a stress increment according to the following :

$$\dot{\boldsymbol{\sigma}} = \mathbf{C} : \dot{\boldsymbol{\varepsilon}} \quad (54)$$

where \mathbf{C} is the fourth-order tangent modulus tensor. Begin with the functional form for the stress, that is $\boldsymbol{\sigma} = \boldsymbol{\sigma}(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^P, \mathbf{D})$. By taking the time derivative we obtain :

$$\dot{\boldsymbol{\sigma}} = \frac{\partial \boldsymbol{\sigma}}{\partial \boldsymbol{\varepsilon}} : \dot{\boldsymbol{\varepsilon}} + \frac{\partial \boldsymbol{\sigma}}{\partial \boldsymbol{\varepsilon}^P} : \dot{\boldsymbol{\varepsilon}}^P + \frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{D}} : \dot{\mathbf{D}} = \mathbf{E} : (\dot{\boldsymbol{\varepsilon}} - \dot{\boldsymbol{\varepsilon}}^P) + \frac{\partial^2 \Psi}{\partial \boldsymbol{\varepsilon} \partial \mathbf{D}} : \dot{\mathbf{D}}. \quad (55)$$

With the substitution of the evolution relations, the result is :

$$\dot{\boldsymbol{\sigma}} = \mathbf{E} : \dot{\boldsymbol{\varepsilon}} - \dot{\lambda}_P \mathbf{E} : \mathbf{M} : \frac{\partial \Phi_P}{\partial \hat{\boldsymbol{\sigma}}} + \dot{\lambda}_D \frac{\partial^2 \Psi}{\partial \boldsymbol{\varepsilon} \partial \mathbf{D}} : \frac{\partial \Phi_D}{\partial \mathbf{Y}}. \quad (56)$$

The explicit relations for the consistency parameters are then used. Then all terms on the right hand side are linear in $\dot{\boldsymbol{\varepsilon}}$ and hence, the expression for \mathbf{C} is completely determined according to eqn (54). For the general case this expression is rather lengthy, and for this

reason the explicit expression for the tangent modulus is not presented but it can be easily determined for any particular formulation.

2.2.5. *Second law.* With the use of the evolution relations, eqns (37) and (40), the dissipation rate can be written in the form

$$\Upsilon = \dot{\lambda}_P \left((\hat{\boldsymbol{\sigma}} - \hat{\boldsymbol{\sigma}}^S) : \frac{\partial \phi_{P1}}{\partial \hat{\boldsymbol{\sigma}}} - \hat{\boldsymbol{\sigma}}^H : \frac{\partial \phi_{P2}}{\partial \hat{\boldsymbol{\sigma}}^H} \right) + \dot{\lambda}_D \left((\mathbf{Y} - \mathbf{Y}^S) : \frac{\partial \phi_{D1}}{\partial \mathbf{Y}} - \mathbf{Y}^H : \frac{\partial \phi_{D2}}{\partial \mathbf{Y}^H} \right). \quad (57)$$

Next, employ the HODO properties of eqns (33) and (34) along with the postulated form of the yield and damage functions, eqns (31) and (35), respectively, to obtain :

$$\Upsilon = \dot{\lambda}_P [\Phi_P(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\sigma}}^S, \hat{\boldsymbol{\sigma}}^H) + \sigma_y] + \dot{\lambda}_D [\Phi_D(\mathbf{Y}, \mathbf{Y}^S, \mathbf{Y}^H) + \omega_0]. \quad (58)$$

The first term in each set of square parentheses is zero by the loading and unloading conditions of eqn (38). Hence the dissipation is given by the following :

$$\Upsilon = \dot{\lambda}_P \sigma_y + \dot{\lambda}_D \omega_0. \quad (59)$$

Since $\dot{\lambda}_P \geq 0$ and $\dot{\lambda}_D \geq 0$ by the loading/unloading conditions and σ_y, ω_0 are defined to be positive parameters, the dissipation is guaranteed to be nonnegative. Therefore, the second law of thermodynamics is always satisfied according to eqn (6).

2.2.6. *First law.* One of the equations that follows from the first law given in eqn (10) provides an explicit relation for the rate of entropy production, which is always positive :

$$\dot{s} = \frac{(\dot{\lambda}_P \sigma_y + \dot{\lambda}_D \omega_0)}{\rho \theta}. \quad (60)$$

Other relations involving internal energy or the Helmholtz function can likewise be derived.

2.2.7. *Specific HODO functions.* Both the damage and the yield criteria are constructed using scalar-valued HODO functions. This requirement is integral to satisfying the laws of thermodynamics. Any function that satisfies the HODO requirement has been shown to be sufficient. General HODO functions, a scalar form and three forms that are functions of second-order tensors, are presented below. It is not claimed that those presented are all inclusive but it will be shown that many popular theories can be constructed with these forms. It should be noted that any linear combination of HODO functions is also a HODO function. HODO forms of increasingly higher powers of the independent variable and increasing tensor order can be constructed in a similar fashion.

Scalar form

Only one form of a HODO function exists for a scalar argument, x . This function along with its associated derivative is :

$$\phi(x) \equiv Kx \quad \frac{\partial \phi(x)}{\partial x} = K, \quad (61)$$

where K is a scalar constant.

Linear form for second-order tensors

A scalar-valued function of a second-order tensor, \mathbf{X} , is a linear HODO function if:

$$\phi(\mathbf{X}) \equiv \mathbf{X} : \mathbf{C}, \quad (62)$$

where \mathbf{C} is a second-order tensor. Specifically, if \mathbf{C} is the second-order identity then the function $\phi(x)$ is the first-invariant of \mathbf{X} , or the $\text{tr}(\mathbf{X})$. The derivative with respect to \mathbf{X} is:

$$\frac{\partial \phi(\mathbf{X})}{\partial \mathbf{X}} = \mathbf{C}. \quad (63)$$

Quadratic form for second-order tensors

A quadratic HODO function of a second-order tensor and its derivative with respect to the independent variable, \mathbf{X} , are

$$\phi(\mathbf{X}) \equiv \sqrt{\mathbf{X} : \mathbf{A} : \mathbf{X}}, \quad \frac{\partial \phi(\mathbf{X})}{\partial \mathbf{X}} = \frac{\mathbf{A} : \mathbf{X}}{\sqrt{\mathbf{X} : \mathbf{A} : \mathbf{X}}}, \quad (64)$$

where \mathbf{A} is a positive semi-definite operator (a fourth-order tensor) that is independent of \mathbf{X} . \mathbf{A} is required to possess the symmetries $A_{ijkl} = A_{jikl} = A_{ijlk} = A_{klij}$. If \mathbf{A} is the fourth-order identity then the HODO function is the square-root of a second-invariant, $\phi(\mathbf{X}) = \sqrt{\text{tr}(\mathbf{X}^2)}$.

Cubic form for second-order tensors

A cubic HODO function of a second-order tensor and its derivative are:

$$\phi(\mathbf{X}) \equiv \sqrt[3]{(\mathbf{X} : \mathbf{B} : \mathbf{X}) : \mathbf{X}}, \quad \frac{\partial \phi(\mathbf{X})}{\partial \mathbf{X}} = \frac{(\mathbf{X} : \mathbf{B} : \mathbf{X})}{[(\mathbf{X} : \mathbf{B} : \mathbf{X}) : \mathbf{X}]^{2/3}}, \quad (65)$$

where \mathbf{B} is a positive semi-definite operator (a sixth-order tensor) that is independent of \mathbf{X} and possesses the appropriate symmetries. If \mathbf{X} is symmetric and $B_{ijklmn} = \delta_{il}\delta_{jm}\delta_{kn}$, where δ_{ab} is the Kronecker delta, then the function is the cube-root of a third-invariant, $\phi(\mathbf{X}) = \sqrt[3]{\text{tr}(\mathbf{X}^3)}$.

2.2.8. *Summary of the thermodynamic framework.* The development of the thermodynamic framework for coupled elastoplasticity and damage formulations is now complete. The remaining sections deal with specific applications of this framework. The assumptions and limitations used in the derivations are summarized below:

1. Only isothermal processes are considered which precludes the inclusion of heat conduction and temperature dependent material properties.
2. Although infinitesimal deformations are assumed most of the results can be generalized to the case of finite deformations.
3. Strain rate dependence is not included.
4. The postulated form of the Helmholtz free energy is not the most general and other forms may be more appropriate for some materials. The form used is relatively simple and yet yields classical relations. The consequences of other postulated forms follow from the conjugate relationships.
5. The concept of an effective stress that is associated with a damaged material is the critical element that couples the theories of elastoplasticity with damage and leads naturally to the use of the effective-stress plasticity formulation.
6. The use of HODO functions in the construction of the yield and damage functions is not overly restrictive. As is shown, many of the popular theories can be constructed from these functions.

7. The principle of maximum entropy is used to determine the evolutionary relations. This assumption implies that these evolutionary relations involve tensors normal to the yield and damage surfaces which for plasticity is equivalent to an associated flow rule. Convexity is also implied by this principle.

None of these assumptions are too restrictive. If a formulation is needed that violates one of these restrictions then a similar derivation could be followed with modified assumption(s). The details and rigor presented in the derivations allows for extensions to alternate applications. By developing a general framework for coupled formulations, specific constitutive theories can easily be constructed by following the stated requirements.

3. SPECIFIC COUPLED ELASTOPLASTICITY AND DAMAGE MODELS

In the preceding section a unified framework for elastoplasticity coupled with damage was developed following a rigorous thermodynamic approach. In order to construct specific models, without repeating the rigorous thermodynamic arguments, the formulation needs only to adhere to the requirements summarized below and the thermodynamic restrictions will be satisfied.

1. Based on either a phenomenological or a micromechanical approach, develop an effective-stress relationship, that is, determine $\mathbf{M}(\mathbf{D})$.
2. Construct a yield function using effective stresses and a linear combination of HODO functions. The evolution relations for $\tilde{\boldsymbol{\varepsilon}}^p$ and the hardening variables, $\tilde{\boldsymbol{\varepsilon}}^S$ and $\tilde{\boldsymbol{\varepsilon}}^H$ are determined by eqn (37).
3. Based on experimental observations, construct the constitutive relations for the hardening variables, $\hat{\boldsymbol{\sigma}}^S(\tilde{\boldsymbol{\varepsilon}}^S)$ and $\hat{\boldsymbol{\sigma}}^H(\tilde{\boldsymbol{\varepsilon}}^H)$.
4. With the use of the appropriate thermodynamic damage variable, determine the damage energy release rate, \mathbf{Y} , from eqn (12).
5. From a linear combination of HODO functions, construct a damage function using \mathbf{Y} , \mathbf{Y}^S , and \mathbf{Y}^H . The corresponding evolution relations for the damage, \mathbf{D} , and the damaging variables, \mathbf{D}^S and \mathbf{D}^H are determined by eqn (37).
6. Postulate forms of the constitutive relations for the damaging variables, $\mathbf{Y}^S(\mathbf{D}^S)$ and $\mathbf{Y}^H(\mathbf{D}^H)$.

The motivation for any of the postulated relations should be based on micromechanical considerations or physical observations. The utility of the unified framework developed in Section 2, is now illustrated by considering some specific plasticity, damage, and coupled theories.

3.1. Specific elastoplasticity models

Consider now a specific class of plasticity models, called J_2 plasticity. This encompasses the largest class of models in use today, including von Mises plasticity. For J_2 plasticity the deviatoric projection operator is used in place of the operator \mathbf{A} of eqn (64). Without the shift stress term, $\phi_{p1}(\hat{\boldsymbol{\sigma}})$ becomes:

$$\phi_{p1}(\hat{\boldsymbol{\sigma}}) = \sqrt{\hat{\boldsymbol{\sigma}} : \mathbf{P}^d : \hat{\boldsymbol{\sigma}}} \equiv \sqrt{\hat{\boldsymbol{\sigma}} : \mathbf{P}^d : \mathbf{P}^d : \hat{\boldsymbol{\sigma}}} = \sqrt{\hat{\boldsymbol{\sigma}}^d : \hat{\boldsymbol{\sigma}}^d}, \quad (66)$$

where $\hat{\boldsymbol{\sigma}}^d$ is the deviatoric stress tensor. The function $\phi_{p1}(\hat{\boldsymbol{\sigma}})$ is essentially the square root of the J_2 invariant of the function argument, hence the name “ J_2 plasticity”. If the plasticity hardening variable associated with the shift of the stress, or the kinematic hardening, is included then:

$$\phi_{p1}(\hat{\boldsymbol{\sigma}} - \hat{\boldsymbol{\sigma}}^S) \equiv \sqrt{(\hat{\boldsymbol{\sigma}} - \hat{\boldsymbol{\sigma}}^S) : \mathbf{P}^d : (\hat{\boldsymbol{\sigma}} - \hat{\boldsymbol{\sigma}}^S)} = \sqrt{(\hat{\boldsymbol{\sigma}} - \hat{\boldsymbol{\sigma}}^S)^d : (\hat{\boldsymbol{\sigma}} - \hat{\boldsymbol{\sigma}}^S)^d}. \quad (67)$$

A Prager–Drucker pressure dependence can be included by adding a linear HODO function such as $K_{PD}(\hat{\boldsymbol{\sigma}} : \mathbf{i})$ where K_{PD} is a constant and \mathbf{i} is the second-order identity tensor. The evolution of the effective-plastic-strain tensor that corresponds to eqn (67) is:

$$\dot{\tilde{\boldsymbol{\epsilon}}}^P = \dot{\lambda}_p \frac{(\hat{\boldsymbol{\sigma}} - \hat{\boldsymbol{\sigma}}^S)^d}{\sqrt{(\hat{\boldsymbol{\sigma}} - \hat{\boldsymbol{\sigma}}^S)^d : (\hat{\boldsymbol{\sigma}} - \hat{\boldsymbol{\sigma}}^S)^d}}. \quad (68)$$

The yield criterion of eqn (31) is formulated such that an anisotropic change in the shape of the yield surface is allowed as evident by dependence of the function $\phi_{p_2}(\hat{\boldsymbol{\sigma}}^H)$ on the second-order tensor $\hat{\boldsymbol{\sigma}}^H$. Although some have developed anisotropic plasticity theories (Hill, 1950), many still use an isotropic change in the yield surface or hardening. For isotropic hardening the second-order hardening tensors, $\hat{\boldsymbol{\sigma}}^H$ and $\tilde{\boldsymbol{\epsilon}}^H$, degenerate to the following isotropic forms :

$$\hat{\boldsymbol{\sigma}}^H = \hat{\sigma}^H \mathbf{i}, \quad \tilde{\boldsymbol{\epsilon}}^H = \tilde{\epsilon}^H \mathbf{i}, \quad (69)$$

where $\hat{\sigma}^H$ and $\tilde{\epsilon}^H$ are the isotropic hardening variables (scalars). Either the substitution of an isotropic second-order tensor into eqn (64) or the substitution of the scalar variables into eqn (61) results in a scalar form of the hardening HODO function, such as $\phi_{p_2}(\hat{\boldsymbol{\sigma}}^H) = K_H \hat{\sigma}^H$ where K_H is a constant. Instead, the constant is absorbed in $\hat{\sigma}^H(\tilde{\epsilon}^H)$, which is still to be defined, and the hardening HODO function reduces to :

$$\phi_{p_2}(\hat{\boldsymbol{\sigma}}^H) \equiv \hat{\sigma}^H. \quad (70)$$

From eqn (37), the evolution relation for the isotropic hardening variable is then :

$$\dot{\tilde{\epsilon}}^H = -\dot{\lambda}_p. \quad (71)$$

Since λ_p is a monotonically increasing positive parameter, this implies that $\tilde{\epsilon}^H$ is a monotonically decreasing negative parameter. Recall that the constitutive relation for the hardening variables is postulated in lieu of postulating a specific form for the contribution to the Helmholtz free energy in eqn (14). A simple linear form for the isotropic hardening constitutive relationship is assumed :

$$\hat{\sigma}^H(\tilde{\epsilon}^H) \equiv -K_I \tilde{\epsilon}^H \quad (72)$$

where K_I is a material parameter. The negative sign is required to make $\hat{\sigma}^H$ be a positive variable. Other forms that might be postulated could include trigonometric, exponential, and/or polynomial functions of $\tilde{\epsilon}^H$. A nonlinear constitutive relation for the isotropic hardening variables that resembles a Ramberg-Osgood form is given by :

$$\hat{\sigma}^H(\tilde{\epsilon}^H) \equiv K_I [-\tilde{\epsilon}^H]^{n_p} \quad (73)$$

where n_p is an additional material parameter. Whatever the specific form of the hardening relation, experimental evidence should be used as the motivation.

Similar arguments apply to the choice of a constitutive relationship for the kinematic hardening variables. A general linear relationship between $\tilde{\boldsymbol{\epsilon}}^S$ and $\hat{\boldsymbol{\sigma}}^S$ is given by :

$$\hat{\boldsymbol{\sigma}}^S(\tilde{\boldsymbol{\epsilon}}^S) \equiv -\mathbf{G} : \tilde{\boldsymbol{\epsilon}}^S \quad (74)$$

where \mathbf{G} is a fourth-order tensor to be specified. If \mathbf{G} is the fourth-order identity, then :

$$\hat{\boldsymbol{\sigma}}^S(\tilde{\boldsymbol{\epsilon}}^S) \equiv -K_k \tilde{\boldsymbol{\epsilon}}^S = K_k \tilde{\boldsymbol{\epsilon}}^P \quad (75)$$

where K_k is a scalar material parameter. This is the common Prager kinematic hardening rule.

3.2. Specific damage models

The thermodynamic framework in Section 2 was developed without regard to a specific damage representation theory. The only requirement was that an effective-stress operator be provided. The origin of this operator can either be phenomenological or micromechanical. Presently, most of the micromechanical models are too complicated for practical applications so phenomenological models are emphasized in this section for illustrating the utility of the thermodynamic framework. Most of the damage representation theories of phenomenological origins employ one of two fundamental hypotheses, these being the principle of equivalent elastic energy or the principle of equivalent strains. These hypotheses are reviewed prior to the presentation of specific models.

3.2.1. *Principle of equivalent elastic energy.* The first class of damage representation theories employs the principle of equivalent elastic energy as was introduced by Cordebois and Sidoroff (1979), and has since been employed by other researchers including Chow and Wang (1987a, b). To begin, introduce a form of the stored elastic energy, using the stress tensor:

$$W(\boldsymbol{\sigma}, \mathbf{D}) = \frac{1}{2} \boldsymbol{\sigma} : \mathbf{E}^{-1} : \boldsymbol{\sigma}. \quad (76)$$

The principle of equivalent elastic energy (PEEE) states that (PEEE postulate #1) “the elastic energy of the damaged material is the same in form as that of an undamaged material except that the stress (strain) is replaced by the effective stress (effective strain)”. The stored elastic energy using the effective stress and the undamaged stiffness is given by:

$$W(\hat{\boldsymbol{\sigma}}, \mathbf{0}) = \frac{1}{2} \hat{\boldsymbol{\sigma}} : (\mathbf{E}^o)^{-1} : \hat{\boldsymbol{\sigma}} = \frac{1}{2} \boldsymbol{\sigma} : \mathbf{M}^T : (\mathbf{E}^o)^{-1} : \mathbf{M} : \boldsymbol{\sigma}. \quad (77)$$

The PEEE implies that the two preceding forms for the stored energy are equivalent which results in the following form of the damage stiffness tensor:

$$\begin{aligned} \mathbf{E}^{-1} &= \mathbf{M}^T : (\mathbf{E}^o)^{-1} : \mathbf{M} \\ \mathbf{E} &= \mathbf{M}^{-1} : \mathbf{E}^o : \mathbf{M}^{-T}. \end{aligned} \quad (78)$$

Implicit in this hypothesis is that the same procedure for deriving the average elastic properties (stiffness tensor) is applicable for determining the effective stresses. Recall the effective stress is the mechanism for coupling plasticity with damage, and the effective stress operator, \mathbf{M} , should be derived based on a proper effective-stress tensor as it relates to plasticity. In contrast the stiffness tensor, which corresponds to the elastic process, may or may not follow the same derivation. Hence the implicit assumption may not be appropriate. Whatever the method of derivation, a stiffness tensor as a function of a damage measure is the only required entity in addition to the effective-stress operator as a function of damage for incorporation in the thermodynamically consistent framework. Additional postulates and relations, such as a relation for the effective strain, may lead to contradictions. A common approach is presented below to illustrate these inconsistencies.

Typically, one of two equivalent assumptions is employed to determine an effective-elastic strain. Either the constitutive relation between the effective stress and effective-elastic strain is assumed to be (PEEE postulate #2a):

$$\hat{\boldsymbol{\sigma}} \equiv \mathbf{E}^o : \tilde{\boldsymbol{\epsilon}}^e \quad (79)$$

or an equivalent assumption is (PEEE postulate #2b) to apply the strain form of the PEEE and use the form of the damaged stiffness given in eqn (78) such as:

$$W(\boldsymbol{\epsilon}^e, \mathbf{D}) = \frac{1}{2} \boldsymbol{\epsilon}^e : \mathbf{E} : \boldsymbol{\epsilon}^e = \frac{1}{2} \boldsymbol{\epsilon}^e : \mathbf{M}^{-1} : \mathbf{E}^o : \mathbf{M}^{-T} : \boldsymbol{\epsilon}^e = \frac{1}{2} \tilde{\boldsymbol{\epsilon}}^e : \mathbf{E}^o : \tilde{\boldsymbol{\epsilon}}^e = W(\tilde{\boldsymbol{\epsilon}}^e, \mathbf{0}). \quad (80)$$

From either postulate, the relation for the effective-elastic strain is given by:

$$\tilde{\epsilon}^e = \mathbf{M}^{-T} : \epsilon^e. \tag{81}$$

Damage theories employing the PEEE require one final postulate for application to coupled formulations. This postulate is that all strain-space variables are assumed (PEEE postulate # 3) to be mapped by the same operator as given in eqn (81), for example,

$$\tilde{\epsilon}^P = \mathbf{M}^{-T} : \epsilon^P. \tag{82}$$

The effective-strain relations of eqns (81) and (82) are incompatible with the transformation relations for effective strain given in eqn (25). Formulations that employ the principle of equivalent elastic energy may still be incorporated in the thermodynamic framework with the restriction that only the PEEE postulate # 1 is allowed because the additional postulates yield a contradiction in the effective strains.

3.2.2. *Principle of strain equivalence.* The second class of damage representation theories employ the principle of strain equivalence, as postulated by Lemaitre (1971) and subsequently used by Simo and Ju (1987) and Ju (1989). This hypothesis states “The strain associated with a damaged state under the applied stress is equivalent to the strain associated with its undamaged state under the effective stress”. In essence, the effective material behavior is represented in effective-stress and actual-strain space. With this principle, an equivalent energy state does not exist and effective strains are not employed. Then the damaged stiffness and the effective-stress relations take the following forms :

$$\mathbf{E}' = \mathbf{M}^{-1} : \mathbf{E}^o \quad \hat{\sigma}' = \mathbf{E}^o : \epsilon^e. \tag{83}$$

The difference between these principles is further illustrated in Fig. 1. The undamaged portion of the response is from points ① to ②. This is followed by a regime in which the damage continually evolves up to the maximum damage state, point ③. Point ⑦ denotes the effective continuum state using the equivalent strain approach, where the effective stress is denoted by $\hat{\sigma}'$. The effective-stress and effective-strain state, denoted as point ⑤, is determined by equating stored energy. The stored energy associated with the actual state variables, given by the area ①–③–⑥, is equated to the area in the effective space using the undamaged modulus, which is given by the area ①–⑤–④. Note that the effective stress for the principle of equivalent strain is considerably higher than that for the equivalent energy approach. Since the effective stress is utilized in the yield function, the choice between

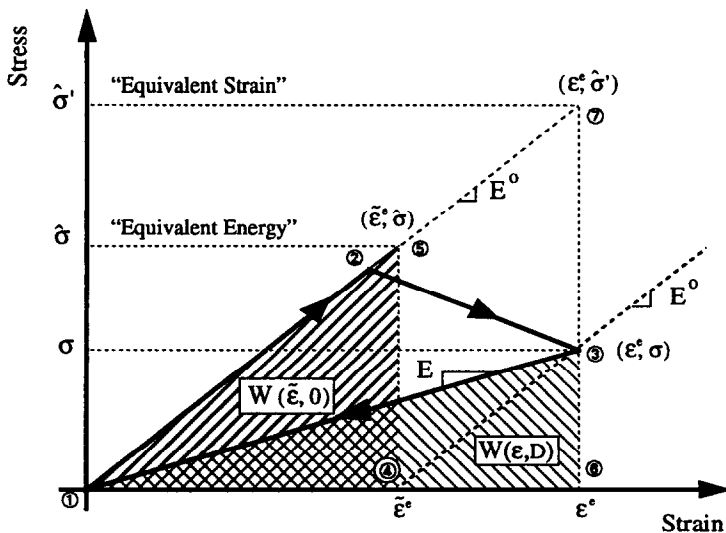


Fig. 1. Difference between equivalent energy and equivalent strain hypotheses.

the principle of equivalent energy and principle of equivalent strain has a marked difference on the behavior of the model.

In summary, two popular phenomenological hypotheses employed by various damage representation theories have been evaluated relative to their applicability to the derived thermodynamic framework. Additional postulates to determine the effective strains were shown to be at variance with the effective strains given in Section 2.2.2. based on the form of the dissipation rate. Therefore, for the present thermodynamic framework, the relations for the effective strains are a consequence of the form of the dissipation rate and may not be postulated or derived separately.

In the next subsections, three specific damage formulations are presented and evaluated using the general framework. The principle of equivalent elastic energy with the restrictions as discussed in Section 3.2.1 is employed for the specific formulations. A similar set, with similar results, could be developed using the principle of equivalent strains, but are not presented. All three damage formulations are found in the literature and have been developed using a phenomenological approach to damage representation. First, a simple scalar isotropic formulation is considered. Then a second-order anisotropic damage theory is evaluated. Finally, a class of damage models where the damage measure is directly associated with the damaged stiffness tensor is formulated; hence, this is a fourth-order anisotropic damage theory. Due to their phenomenological origins, each of these formulations is shown to possess deficiencies.

3.2.3. Scalar isotropic damage. Isotropic scalar damage theories are the most common models found in the literature. The scalar variable, d , represents the amount of volumetric material damage. The effective stress is postulated to be of the form :

$$\tilde{\sigma} \equiv \frac{\sigma}{1-d}. \quad (84)$$

With the use of the principle of equivalent elastic energy, the damaged stiffness relation is :

$$\mathbf{E} = (1-d)^2 \mathbf{E}^e. \quad (85)$$

The thermodynamic damage variable is selected as d and all other damage related variables are scalars. For easy comparison with other models, neglect the surface energy contribution to the Helmholtz function and the damaging shift variable. The damage energy release rate, y , or the variable conjugate to the thermodynamic damage is :

$$y = (1-d)\boldsymbol{\varepsilon}^e : \mathbf{E}^e : \boldsymbol{\varepsilon}^e. \quad (86)$$

Recall that the damage function is constructed using HODO functions. For scalar damage variables, the damage function degenerates to the following :

$$\Phi_D(y, y^H) = y - (y^H + \omega_0) \leq 0. \quad (87)$$

The resulting evolution relations and a postulated constitutive relation for $y^H(d^H)$ are :

$$\dot{d} = \dot{\lambda}_D \quad \dot{d}^H = -\dot{\lambda}_D \quad y^H \equiv -\kappa_a d^H, \quad (88)$$

where κ_a is a scalar material parameter. Only two material parameters, ω_0 and κ_a are required for this formulation. The inability to model the anisotropy associated with damage is the major weakness of isotropic theories. As shown by Ju (1990) an isotropic damage model implies that the Poisson's ratio does not change, a feature at variance with some experimental data. In addition, compressive and tensile loads cannot be differentiated. Hence, equal damage states would be predicted for uniaxial tension and compression load

paths of the same magnitude. A hydrostatic pressure dependence in the damage function could be incorporated to differentiate between tension and compression in some sense but this would not be sufficiently general to properly represent all load paths. These deficiencies render scalar theories ineffective for the modeling of some materials.

3.2.4. *Second-order damage measure.* In this section, an anisotropic damaged-elasticity theory initially proposed by Cordebois and Sidoroff (1979) is considered. This has since been used in one form or another by several investigators including Chow and Wang (1987a, b). First, introduce a second-order symmetric damage tensor, \mathbf{D} . A symmetric second-order tensor as the damage measure implies that the most complicated anisotropy that can be represented is orthotropic, a rectangular symmetry with axis aligned with the principal basis of the damage tensor. This orthotropy is not fixed in the material but the orthotropy rotates with the principal basis of the damage tensor. From a micromechanical derivation, Kachanov (1987) showed that even for high crack densities with interacting cracks, the effective elastic properties remained orthotropic with good accuracy. Therefore, this limitation on the type of anisotropy does not appear to be too restrictive.

Cordebois and Sidoroff (1979) postulated the effective stress to be of the form :

$$\hat{\sigma} = (\mathbf{i} - \mathbf{D})^{-1/2} \cdot \boldsymbol{\sigma} \cdot (\mathbf{i} - \mathbf{D})^{-1/2} \equiv \mathbf{M}(\mathbf{D}) : \boldsymbol{\sigma} \tag{89}$$

which also defines the fourth-order effective-stress operator $\mathbf{M}(\mathbf{D})$ as :

$$M_{ijkl} = (\delta_{ik} - D_{ik})^{1/2} (\delta_{jl} - D_{jl})^{1/2}. \tag{90}$$

By employing the principle of elastic energy equivalence the damaged stiffness tensor is :

$$E_{ijkl} = (\delta_{im} - D_{im})^{1/2} (\delta_{jn} - D_{jn})^{1/2} (\delta_{ko} - D_{ko})^{1/2} (\delta_{lp} - D_{lp})^{1/2} E_{mnop}^o, \tag{91}$$

where the symmetry of the damage tensor causes \mathbf{E} to possess all the symmetries of \mathbf{E}^o .

The logical choice for the thermodynamic damage variable is the second-order damage measure, \mathbf{D} . The damage energy release rate associated with this selection is :

$$\mathbf{Y} = \frac{1}{2} \left\{ \begin{array}{l} \Omega_{ip}^{-1} \varepsilon_{qb}^e \Omega_{bj} E_{ijkl}^o \Omega_{km} \varepsilon_{mn}^e \Omega_{nl} \\ + \Omega_{ia} \varepsilon_{ab}^e \Omega_{bp}^{-1} E_{iqkl}^o \Omega_{km} \varepsilon_{mn}^e \Omega_{nl} \end{array} \right\} (\mathbf{e}_p \otimes \mathbf{e}_q) - 2\gamma_D \mathbf{D}, \tag{92}$$

where

$$\boldsymbol{\Omega} \equiv (\mathbf{i} - \mathbf{D})^{1/2}. \tag{93}$$

Unfortunately, a thorough evaluation reveals that \mathbf{Y} is *not* a symmetric second-order tensor unless the damage and the strain tensors have the same principal bases. A damage tensor which is nonsymmetric is a direct violation of the assumed symmetry for \mathbf{D} . A simple change in the selection of the thermodynamic variable will remedy this dilemma. Instead of \mathbf{D} , let $\boldsymbol{\Omega}$ be the thermodynamic damage variable. Physically, $\boldsymbol{\Omega}$ is more of a measure of the material integrity than a measure of the damage, but the selection as a thermodynamic damage measure is allowable. The following consequences hold :

$$\mathbf{D} \Big|_{t=0} = \mathbf{0} \stackrel{\mathbf{D} = (\mathbf{i} - \boldsymbol{\Omega}^2)}{\Rightarrow} \boldsymbol{\Omega} \Big|_{t=0} = \mathbf{i}. \tag{94}$$

By taking advantage of the symmetry of the individual tensors and noting that $\hat{\sigma}_{ij} = E_{ijkl}^o \Omega_{km} \varepsilon_{mn}^e \Omega_{nl}$, the damage energy release rate, \mathbf{Y} , associated with $\boldsymbol{\Omega}$ is given by :

$$\mathbf{Y} = \frac{-1}{2} (\hat{\boldsymbol{\sigma}} \cdot \boldsymbol{\Omega} \cdot \boldsymbol{\varepsilon}^e + \boldsymbol{\varepsilon}^e \cdot \boldsymbol{\Omega} \cdot \hat{\boldsymbol{\sigma}}) + 2\gamma_D (\mathbf{D} \cdot \boldsymbol{\Omega} + \boldsymbol{\Omega} \cdot \mathbf{D}). \quad (95)$$

The credibility of the damage function is based on how well it correlates with experimental observations. Consider some observed phenomena in brittle materials. Experimentally, it is observed in uniaxial tension that the microcracks develop perpendicular to the loading direction or the damage is in the direction of loading. Therefore, in uniaxial tension the damage evolves in the direction of tensile stress and tensile strain. Conversely, in an unconfined uniaxial compressive stress state, the microcracks develop in the direction parallel to the loading axis. This then corresponds to damage perpendicular to the direction of loading. No stresses are present in this off-axis direction in uniaxial compression but the Poisson effect produces tensile strains in the off-axis direction. Based on these observations it appears reasonable, to postulate that damage evolves in the direction of tensile strains. Define a fourth-order projection operator, \mathbf{P}^+ , such that only the tensile strain components of $\boldsymbol{\varepsilon}$ are extracted :

$$\boldsymbol{\varepsilon}^+ \equiv \mathbf{P}^+ : \boldsymbol{\varepsilon}. \quad (96)$$

This projection operator can be constructed by following a procedure introduced by Ju (1989). The proposed projection operator is constructed in a series of steps as outlined below. First consider the spectral decomposition of the strain tensor.

$$\boldsymbol{\varepsilon} = \sum_{i=1}^3 \varepsilon_i (\mathbf{p}_i \otimes \mathbf{p}_i), \quad (97)$$

where ε_i is the i th principal strain, the unit vector \mathbf{p}_i is the corresponding i th principal strain direction (or eigenvalue and eigenvector of $\boldsymbol{\varepsilon}$). The positive (tensile) spectral tensor is defined as :

$$\mathbf{Q}^+ \equiv \sum_{i=1}^3 h(\varepsilon_i) (\mathbf{p}_i \otimes \mathbf{p}_i), \quad (98)$$

in which $h(\)$ is the Heaviside function. The positive projection operator (a fourth-order tensor) is then defined as :

$$\mathbf{P}_{ijkl}^+ \equiv \mathbf{Q}_{ik}^+ \mathbf{Q}_{jl}^+. \quad (99)$$

By inspection, if all eigenvalues of $\boldsymbol{\varepsilon}$ are negative (compressive) then \mathbf{P}^+ will be the null tensor and, conversely, if all eigenvalues are positive (tensile) then \mathbf{P}^+ corresponds to the fourth-order identity. The various combinations of positive and negative eigenvalues projects or annihilates the appropriate components of a second-order tensor in the corresponding principal directions of $\boldsymbol{\varepsilon}$.

For both functions, $\phi_{D1}(\mathbf{Y}, \mathbf{Y}^S)$ and $\phi_{D2}(\mathbf{Y}^H)$, it is proposed that the quadratic form of a HODO function be employed where the tensile strain operator is used in place of the positive semi-definite tensor \mathbf{A} . The damage function can then be expressed as :

$$\Phi_D(\mathbf{Y}, \mathbf{Y}^S, \mathbf{Y}^H) \equiv \sqrt{(\mathbf{Y} - \mathbf{Y}^S) : \mathbf{P}^+ : (\mathbf{Y} - \mathbf{Y}^S)} - (\sqrt{\mathbf{Y}^H : \mathbf{P}^+ : \mathbf{Y}^H} + \omega_0) \leq 0. \quad (100)$$

The anisotropic nature of the problem is preserved due to the directional dependence of the projection operator and the anisotropy of the damage variables. From eqn (37), the damage (or integrity in this case) evolution is :

$$\dot{\boldsymbol{\Omega}} = \dot{\lambda}_D \frac{\mathbf{P}^+ : (\mathbf{Y} - \mathbf{Y}^S)}{\phi_{D1}(\mathbf{Y}, \mathbf{Y}^S)}. \quad (101)$$

Due to the negative sign in eqn (95), as the damage evolves the eigenvalues of $\boldsymbol{\Omega}$

decrease, which is equivalent to an increase in the eigenvalues of \mathbf{D} according to eqn (94). Hence, the damage does evolve in the direction of the tensile strain projection of the damage energy release rate, as intended. The evolution relations for the damaging variables are:

$$\begin{aligned}\dot{\mathbf{D}}^H &= -\dot{\lambda}_D \frac{\mathbf{P}^+ : \mathbf{Y}^H}{\phi_{D_2}(\mathbf{Y}^H)} \\ \dot{\mathbf{D}}^S &= -\dot{\Omega} = -\dot{\lambda}_D \frac{\mathbf{P}^+ : (\mathbf{Y} - \mathbf{Y}^S)}{\phi_{D_1}(\mathbf{Y}, \mathbf{Y}^S)}.\end{aligned}\quad (102)$$

If the initial value of $\mathbf{Y}^H = \mathbf{0}$ then \mathbf{D}^H will remain zero. Therefore, a nonzero initial value must be assumed in order to get the evolutionary process started. An appropriate choice might be an initial isotropic value such as $\mathbf{Y}^H = \xi \mathbf{i}$ where ξ is a “small” constant.

The constitutive relations for the damaging variables must be constructed to complete this formulation. Unlike plasticity theories where isotropic hardening is often representative of the observed phenomenon, damage is generally observed to be anisotropic. As a consequence the scalar degeneration of the damaging variables \mathbf{Y}^H and \mathbf{D}^H would not be appropriate. A linear, but rather general, constitutive relationship could be postulated of the form:

$$\mathbf{Y}^H(\mathbf{D}^H) \equiv -\mathbf{G} : \mathbf{D}^H \quad (103)$$

where \mathbf{G} is a fourth-order tensor to be specified. In the absence of experimental data, simpler relationships for the damaging variables are postulated of the form:

$$\mathbf{Y}^H(\mathbf{D}^H) \equiv -\kappa_a \mathbf{D}^H, \quad \mathbf{Y}^S(\mathbf{D}^S) \equiv -\kappa_k \mathbf{D}^S \quad (104)$$

where κ_a and κ_k are positive scalar material parameters.

Consider the case where $\hat{\sigma}$, Ω , and ε all have the same principle bases and examine the damage energy release rate of eqn (95). Under a uniaxial tensile stress load path, the tensile strains are in the direction of the tensile stress, while in the off-axis direction the stresses are zero. Hence, the only nonzero component of \mathbf{Y} is in the direction of the loading. This produces damage evolution in this direction as was intended. A further evaluation of \mathbf{Y} reveals some problems. For the load path of uniaxial compression stress, the only nonzero component of \mathbf{Y} is again along the loading axis by the same arguments but the strains are compressive in the loading direction. The positive strain projection operator eliminates components in the direction of negative strains and hence no damage evolution is predicted for uniaxial compression. This is the first deficiency for this formulation.

The second is illustrated by considering a triaxial compression load path. For sufficient axial loading, the strains in the directions of the confining pressure will eventually become positive due to the Poisson effect from the axial load. The difference in signs between the stresses and the strains in the off-axis direction causes \mathbf{Y} to have positive eigenvalues which then causes the principle values of Ω to increase, which implies a decrease in the principle values of the damage. A decrease in damage is the same as “healing” of the material which is not allowed based on physical considerations. This condition can be remedied by redefining the positive projection operator used in the damage function with one that is based on both tensile strains and stresses. For this modification, no damage evolution is predicted for the triaxial compression load path.

A remedy to the first deficiency does not follow as easily. Basically, there are two areas that can be modified in the preceding formulations. A new form of the Helmholtz thermodynamic potential could be postulated and/or a new damage function could be constructed. No satisfactory resolution to the above dilemma has been obtained. For tensile loading conditions, the formulation does well, however, for compressive type loading this formulation appears to be deficient. The inability to predict lateral damage due to an unconfined compression results from the form of the damage energy release rate which is actually a consequence of the damage representation theory as postulated by Cordebois

and Sidoroff (1979). This model is a simple approximation to damage based on volume averages of the effective-areas which does not contain sufficient detail to predict the complex mechanisms occurring on the sub-scale. Until a refined damage representation theory is employed, this is probably the best that can be expected from this formulation.

3.2.5. *Fourth-order damage measure.* A class of anisotropic damage theories is evaluated in which the damaged stiffness tensor evolves directly. This approach has been followed by a number of researchers, including Simo and Ju (1987), Ju (1989), and Yazdani and Schreyer (1990). The thermodynamic damage variable is taken to be the damaged stiffness tensor, \mathbf{E} . Since the damage variable is a fourth-order tensor, all damaging variables are assumed to be fourth-order tensors. The damaging variable associated with shift and the surface energy term are neglected for comparison with the formulations of others. The fourth-order damage energy release rate, \mathbf{Y} , is given by :

$$\mathbf{Y} = \frac{-1}{2} (\boldsymbol{\varepsilon}^e \otimes \boldsymbol{\varepsilon}^e). \quad (105)$$

A damage function based on tensile strain projection is again postulated :

$$\Phi_D(\mathbf{Y}, \mathbf{Y}^H) \equiv \sqrt{\mathbf{Y} :: \mathbf{P}^+ :: \mathbf{Y}} - (\sqrt{\mathbf{Y}^H :: \mathbf{P}^+ :: \mathbf{Y}^H} + \omega_0) \leq 0, \quad (106)$$

where \mathbf{P}^+ is now an eighth-order tensor. The damaged stiffness evolution is then :

$$\dot{\mathbf{E}} = \dot{\lambda}_D \frac{\mathbf{P}^+ : \mathbf{Y}}{\phi_{D1}(\mathbf{Y})} = (-\dot{\lambda}_D) \frac{(\boldsymbol{\varepsilon}^{e+} \otimes \boldsymbol{\varepsilon}^{e+})}{\sqrt{(\boldsymbol{\varepsilon}^{e+} \otimes \boldsymbol{\varepsilon}^{e+}) :: (\boldsymbol{\varepsilon}^{e+} \otimes \boldsymbol{\varepsilon}^{e+})}}. \quad (107)$$

The components of the damaged stiffness tensor are degraded corresponding to the directions of principal tensile strains, which correlates to the observations in brittle materials. However, this fourth-order theory also contains inherent deficiencies. Consider a uniaxial tensile load path, where the specimen is taken to the state of significant damage. Now release the load and reload in pure shear. The previous degradation of the stiffness tensor leaves the shear components of the damaged stiffness tensor unaltered, and hence, a shear loading responds as an undamaged material. This may not be physically realistic.

The preceding approach is formulated independent of an effective-stress relation. Recall that the effective-stress relation is the crucial mechanism by which the effective-stress plasticity and damage theories are coupled. Therefore, to calculate the effective stresses, the operator, \mathbf{M} , must be extracted from the evolved damaged stiffness tensor. A method for performing this extraction is outlined below. For any damage representation where the principle of equivalent elastic energy is employed, the damaged stiffness is related to the effective-stress operator by :

$$\mathbf{E} = \mathbf{M}^{-1} : \mathbf{E}^o : \mathbf{M}^{-T}. \quad (108)$$

Next perform a Cholesky decomposition of the damaged and undamaged stiffness tensors :

$$\mathbf{E} = \mathbf{G} : \mathbf{G}^T, \quad \mathbf{E}^o = \mathbf{H} : \mathbf{H}^T. \quad (109)$$

These results can be combined to obtain a relation for the effective-stress operator :

$$\mathbf{G} : \mathbf{G}^T = \mathbf{M}^{-1} : (\mathbf{H} : \mathbf{H}^T) : \mathbf{M}^{-T} \Rightarrow \mathbf{M}^{-1} = \mathbf{G} : \mathbf{H}^{-1}. \quad (110)$$

For plausible evolutions of the stiffness tensor, the effective-stress operator that is extracted using eqn (110) does not always produce a symmetric effective-stress tensor, a result which may not be physically plausible. In addition, the computational effort required to perform this decomposition is extensive and considering that at least one decomposition

would be required for each load step, the practical utility of this approach becomes cost prohibitive. The difficulty in extracting the effective-stress operator may render this approach ineffective for a coupled formulation.

An alternate fourth-order approach which does not require a Cholesky decomposition of the stiffness tensor is proposed. Introduce a fourth-order damage tensor, \mathbf{D} , and unlike the previous fourth-order theory which used \mathbf{E} , this tensor is also the thermodynamic damage variable. The damage tensor is required to possess the symmetries $D_{ijkl} = D_{klij} = D_{jkl i} = D_{ijlk}$. An effective-stress relation is postulated to be of the form :

$$\hat{\boldsymbol{\sigma}} \equiv (\mathbf{I} - \mathbf{D})^{-1} : \boldsymbol{\sigma}, \quad (111)$$

where \mathbf{I} is the fourth-order identity tensor. This also provides the definition for the effective-stress operator which is $\mathbf{M} = (\mathbf{I} - \mathbf{D})^{-1}$. Then with the principle of equivalent elastic energy the damaged stiffness is :

$$\mathbf{E} = (\mathbf{I} - \mathbf{D}) : \mathbf{E}^o : (\mathbf{I} - \mathbf{D}). \quad (112)$$

This implies a damage energy release rate of the form :

$$\mathbf{Y} = \frac{1}{2} [\hat{\boldsymbol{\sigma}} \otimes \boldsymbol{\varepsilon}^e + \boldsymbol{\varepsilon}^e \otimes \hat{\boldsymbol{\sigma}}]. \quad (113)$$

The damage function is that given in eqn (106) and the resulting damage evolution is :

$$\dot{\mathbf{D}} = \dot{\lambda}_D \frac{\mathbf{P}^+ : \mathbf{Y}}{\phi_{D1}(\mathbf{Y})}. \quad (114)$$

The problem associated with the extraction of the effective-stress operator is thus avoided using this approach. The effective-stress operator is constructed using the damage tensor as it evolves. Unfortunately, the inability to predict the observed evolution of lateral damage from unconfined compression is exhibited, similar to the results associated with the use of a second-order damage tensor. The problem related to the undamaged shear components of the stiffness tensor, as discussed for the previous fourth-order theory, is again exhibited. Although this fourth-order theory can now be utilized in a coupled formulation because a relation for the effective stress is available, some deficiencies are still apparent. In summary, each of the three phenomenological damage formulations possess an inherent deficiency that results from approximation of the complex microscale mechanisms.

3.3. Specific coupled models

The utility of the thermodynamically consistent framework developed in Section 2 is highlighted when considering specific coupled models. A coupled formulation is thermodynamically consistent if both the elastoplasticity and the damage portions are developed using the framework such as those presented in Sections 3.1 and 3.2, respectively. Therefore, any of the specific plasticity models presented in Section 3.1., or others developed in a similar manner, can be coupled without further derivations to any specific damage model such as those given in Section 3.2.

To illustrate the ease by which different elastoplasticity and damage formulations can be coupled, a specific example is presented. For the elastoplasticity model, the J_2 plasticity model with linear isotropic hardening and linear kinematic (Prager) hardening is selected. The damage model selected is the damage representation theory of Cordebois and Sidoroff (1979) with the anisotropic damage function based on tensile strain projections, as presented in Section 3.2.2. Using these specific models, the pertinent relations for the coupled formulation are summarized as follows :

Damage energy release rate :

$$\mathbf{Y} = -\frac{1}{2} (\hat{\boldsymbol{\sigma}} \cdot \boldsymbol{\Omega} \cdot \boldsymbol{\varepsilon}^e + \boldsymbol{\varepsilon}^e \cdot \boldsymbol{\Omega} \cdot \hat{\boldsymbol{\sigma}}) + 2\gamma_D (\mathbf{D} \cdot \boldsymbol{\Omega} + \boldsymbol{\Omega} \cdot \mathbf{D})$$

Damage function :

$$\Phi_D(\mathbf{Y}, \mathbf{Y}^S, \mathbf{Y}^H) = \sqrt{(\mathbf{Y} - \mathbf{Y}^S) : \mathbf{P}^+ : (\mathbf{Y} - \mathbf{Y}^S)} - (\sqrt{\mathbf{Y}^H : \mathbf{P}^+ : \mathbf{Y}^H} + \omega_0) \leq 0$$

Damage evolution relations :

$$\dot{\Omega} = -\dot{\mathbf{D}}^S = \dot{\lambda}_D \frac{\mathbf{P}^+ : (\mathbf{Y} - \mathbf{Y}^S)}{\phi_{D1}(\mathbf{Y}, \mathbf{Y}^S)}, \quad \dot{\mathbf{D}}^H = -\dot{\lambda}_D \frac{\mathbf{P}^+ : \mathbf{Y}^H}{\phi_{D2}(\mathbf{Y}^H)}$$

Damaging constitutive relations :

$$\mathbf{Y}^H(\mathbf{D}^H) \equiv -\kappa_a \mathbf{D}^H, \quad \mathbf{Y}^S(\mathbf{D}^S) \equiv -\kappa_k \mathbf{D}^S$$

Damage :

$$\mathbf{D} = \mathbf{i} - \Omega^2$$

Damaged stiffness :

$$E_{ijkl} = (\delta_{im} - D_{im})^{1/2} (\delta_{jn} - D_{jn})^{1/2} (\delta_{ko} - D_{ko})^{1/2} (\delta_{lp} - D_{lp})^{1/2} E_{mnpq}^0$$

Effective stress :

$$\hat{\sigma} = (\mathbf{i} - \mathbf{D})^{-1/2} \cdot \boldsymbol{\sigma} \cdot (\mathbf{i} - \mathbf{D})^{-1/2} \equiv \mathbf{M}(\mathbf{D}) : \boldsymbol{\sigma}$$

Yield function :

$$\Phi_p(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\sigma}}^S, \hat{\boldsymbol{\sigma}}^H) = \sqrt{(\hat{\boldsymbol{\sigma}} - \hat{\boldsymbol{\sigma}}^S)^d : (\hat{\boldsymbol{\sigma}} - \hat{\boldsymbol{\sigma}}^S)^d} - (\hat{\sigma}^H + \sigma_y) \leq 0$$

Plasticity evolutions relations :

$$\dot{\hat{\boldsymbol{\epsilon}}}^P = \dot{\lambda}_p \frac{(\hat{\boldsymbol{\sigma}} - \hat{\boldsymbol{\sigma}}^S)^d}{\sqrt{(\hat{\boldsymbol{\sigma}} - \hat{\boldsymbol{\sigma}}^S)^d : (\hat{\boldsymbol{\sigma}} - \hat{\boldsymbol{\sigma}}^S)^d}}, \quad \dot{\hat{\boldsymbol{\epsilon}}}^S = -\dot{\hat{\boldsymbol{\epsilon}}}^P, \quad \dot{\hat{\boldsymbol{\epsilon}}}^H = -\dot{\lambda}_p$$

Hardening constitutive relations :

$$\hat{\boldsymbol{\sigma}}^S(\hat{\boldsymbol{\epsilon}}^S) \equiv -K_K \hat{\boldsymbol{\epsilon}}^S, \quad \hat{\boldsymbol{\sigma}}^H(\hat{\boldsymbol{\epsilon}}^H) \equiv -K_I \hat{\boldsymbol{\epsilon}}^H$$

Effective strain rate mapping :

$$\dot{\hat{\boldsymbol{\epsilon}}}^P = \mathbf{M}(\mathbf{D}) : \dot{\hat{\boldsymbol{\epsilon}}}^P.$$

In addition to the undamaged elasticity parameters, the number of material parameters required for this coupled formulation is at most seven: $(\gamma_D, \omega_0, \kappa_a, \kappa_k, \sigma_y, K_I, K_K)$, while the minimum required is four material parameters $(\omega_0, \kappa_a, \sigma_y, K_I)$.

4. SPECIFIC APPLICATIONS

Most metals behave in a ductile manner, which is typically represented by plasticity models, although given sufficient load metals will eventually damage and break. Therefore, both plastic and damage processes occur, and hence, a coupled formulation is needed. To demonstrate the applicability of coupled formulations to ductile failure, a method of evaluating the individual mechanisms must be established. As part of the experimental procedure, a technique for the direct measurement of the damage is required. Other than fatigue damage, little experimental effort has been dedicated to the investigation of the damage in ductile materials such as metals.

The term "direct" measurement does not imply that the size, distribution, and orientations of the microcracks are measured throughout a test. Instead, the state of the material integrity is determined in an averaged sense by easily measured parameters. For example, the change in elastic modulus, which is an explicit function of the damage state, is the direct damage measurement employed by Cordebois and Sidoroff (1979), Chow and Wang (1987b), and Lemaitre (1983). Elastic unloading at periodic intervals in the test sequence is used to determine the changes in the elastic modulus. This approach is in contrast to other investigators that correlate a material state parameter just prior to rupture as a

measure of damage. For example, a measure of the plastic-strain is sometimes used to infer the failure. This is an example of an indirect measure of damage.

At least for some materials, the damage process appears to only occur in the post-peak or softening regime of the stress-strain response. In this response regime, a specimen may separate or localize into two distinct material domains. In a relatively small region as the load (strain) increases the material cannot sustain the previous value of stress because of damage, a feature often called material softening. It is within this localized zone that extensive amounts of damage occur. The remaining material that surrounds the localized zone unloads without further plastic flow or damage. The determination of how the localized region is established and the evolution of the region is still a very active research area. A completely satisfactory method for experimentally evaluating the damaging process as it occurs within the localized zone has yet to be established.

Once a material localizes or softens the boundary value problem becomes ill-posed as exhibited by lack of convergence with mesh refinement for numerical solutions. Some researchers, e.g. Chen and Schreyer (1990) and Bazant (1990), have proposed nonlocal approaches as a means of circumventing these problems. This particular subject was not a part of this research. For now, assume that the materials data considered in this section are correct for a representative volume element and that it is understood that an incorporation into the constitutive formulation of softening requires an additional feature such as nonlocal terms.

4.1. *Aluminum alloy 2024*

Fortunately, all three sets of data provided by Cordebois and Sidoroff (1979), Chow and Wang (1987b), and Lemaitre (1983) contain results for the aluminum alloy 2024. These are the data to be considered for comparison with the models. Chow and Wang (1987b) indicated the alloy temper as 2024-T3 but the others did not specify the temper. Results are similar using independent testing techniques; thus, either the tempers were the same, or the temper had little effect on the results. The first explanation is the most probable.

Unfortunately, all the authors presented incomplete descriptions of the test setup and results. All samples were tested using a load path of uniaxial tensile stress. Chow and Wang (1987b) used sheet stock aluminum in a “dog-bone” shape. The samples were marked with reference lines for inferring strains. Lemaitre (1983) conducted tests using cylindrical tensile bars, and the material strains were measured using small (0.5 mm) strain gages. Cordebois and Sidoroff (1979) did not describe their test setup. It is assumed that all tests were conducted at room temperature. With respect to the test results, none of the researchers presented a corresponding stress-strain response, but all provided results relative to the damage evolution (either a damage measure or the change in the elastic modulus). Chow and Wang (1987b), along with Cordebois and Sidoroff (1979), also presented results on the evolution of the lateral strains which is used to determine the apparent Poisson’s ratio. None indicated whether or not a zone of localization appeared.

Since no stress-strain data were provided, typical full-range stress-strain response data for 2024-T3 aluminum alloy were taken from a standard metals handbook (MIL-HDBK-5E, 1987). A slight inconsistency between the handbook and the experimental data was observed. The maximum strain from the handbook was approximately 0.19, while the experimental results indicated maximum strains of at least 0.25. It is assumed that the experimental results up to a strain value of 0.19 are consistent with the handbook data.

Two different coupled formulations are used to simulate the experimental data. The J_2 plasticity model of Section 3.1. with isotropic hardening is used for both coupled formulations. The nonlinear hardening relation of eqn (73) is employed because it represents the shape of the hardening response for this alloy better than the linear relation. No kinematic hardening is incorporated since the load path is strictly monotonically increasing. The difference between the two coupled formulations is the damage portion of the model. The scalar isotropic damage model of Section 3.2.1. and the second-order damage model of Section 3.2.2. are considered. The latter is identical to the coupled formulation summarized in Section 3.2.3. with the inclusion of the nonlinear isotropic hardening. For both of the damage models the surface energy contribution to the Helmholtz function and the

Table 1. Parameter list for room temperature 2024-T3 aluminum alloy

Parameter name	Nomenclature	Units	Value
Elasticity parameters			
Undamaged elastic modulus	E_0	(MPa)	74,500
Undamaged Poisson's ratio	ν_0	None	0.33
Plasticity parameters			
Initial yield strength	σ_y	(MPa)	250
Linear isotropic hardening modulus	K_I	(MPa)	200
Isotropic hardening exponent	n_p	None	0.4
Damage parameters			
Damage threshold value	ω_0	(MPa)	1.9
Linear damaging parameter	κ_d	(MPa)	15

kinematic damaging are neglected by setting $\gamma_D = 0$ and $\kappa_k = 0$, respectively. The fourth-order damage model was not considered because for this one simple load path the response would be indistinguishable from the model using the second-order damage.

On the figures, the experimental data of Cordebois and Sidoroff (1979), Chow and Wang (1987b), and Lemaitre (1983) are referred to by "C&S79", "C&W87", and "LEM83", respectively. To distinguish the damage model, the nomenclature of "Zero" and "2nd" is used to indicate theories based on zeroth-order (scalar) and second-order damage tensors, respectively.

The experimental and the handbook data are simulated using these two coupled formulations. The set of simulation parameters, presented in Table 1, was deduced using manual trial and error. Both formulations represent the behavior with the same set of material parameters. The simulation of the stress-strain response corresponding to the handbook and experimental data is given in Fig. 2. The correlations between the simulated and the experimental evolution of the elastic modulus is given in Fig. 3. The model simulations are representative of the material behavior, especially considering the scatter in the experimental data.

The correlation for the apparent Poisson ratio is given in Fig. 4. The apparent Poisson ratio is defined by :

$$\nu_{21} \equiv - \frac{\epsilon_{22}^c}{\epsilon_{11}^c} . \tag{115}$$

For the scalar model the apparent Poisson ratio remains constant with increasing damage. This is contrary to the observed behavior that is represented well by the anisotropic model. This deficiency is inherent in a model based on scalar damage as explained by Ju (1990). Other deficiencies would become evident for other load paths.

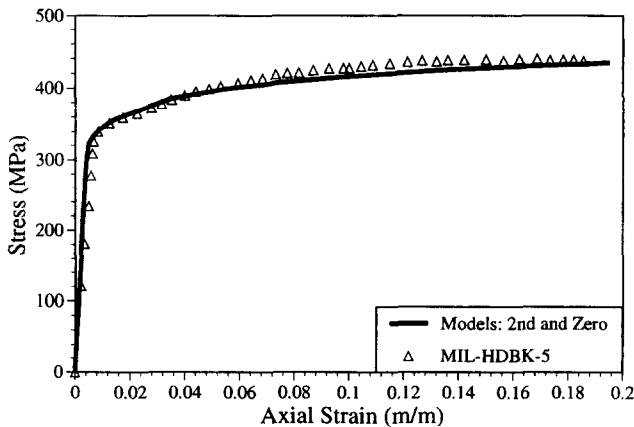


Fig. 2. Stress-strain response of 2024-T3 aluminum alloy.

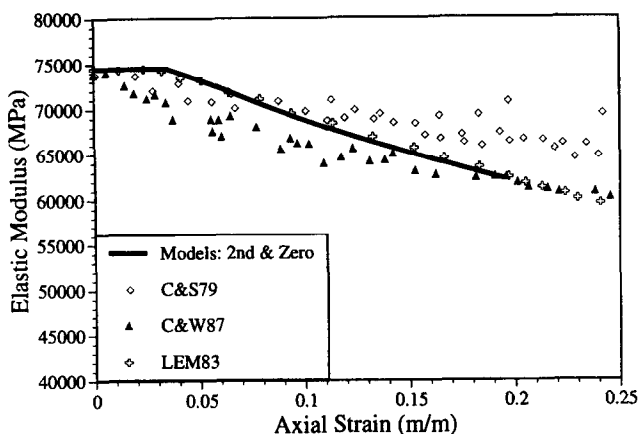


Fig. 3. Evolution of the elastic modulus for 2024-T3 aluminum alloy.

An alternative, but not independent, representation of the model performance is illustrated by the damage measures. One damage measure using the measured elastic modulus is defined by :

$$\omega \equiv 1 - \frac{E^{\text{Measured}}}{E^{\text{Original}}} \tag{116}$$

For the scalar isotropic model the damage variable, d , is related to the measured damage, ω , by :

$$(1 - d)^2 = \frac{E}{E^o} \Rightarrow (1 - d)^2 = (1 - \omega). \tag{117}$$

Similarly, by suitable manipulation of eqn (91) for uniaxial stress, the only nonzero component of the damage tensor, \mathbf{D} , is given by :

$$(1 - D_{11})^2 = \frac{E}{E^o} \Rightarrow (1 - D_{11})^2 = (1 - \omega). \tag{118}$$

Hence, the change in the elastic modulus is directly related to the respective measures of damage. The correlation between the simulated and the measured damage is shown in Fig.

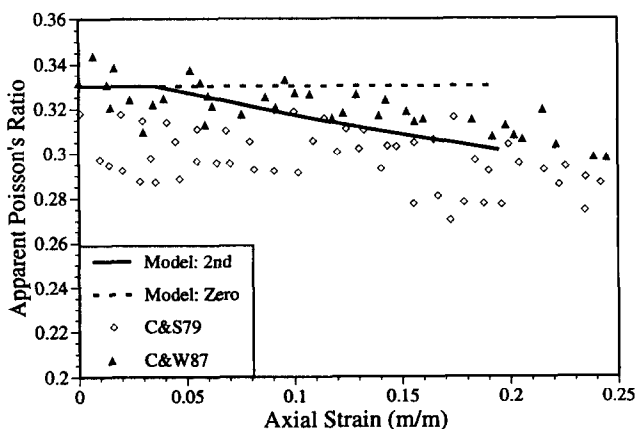


Fig. 4. Evolution of the apparent Poisson's ratio for 2024-T3 aluminum alloy.

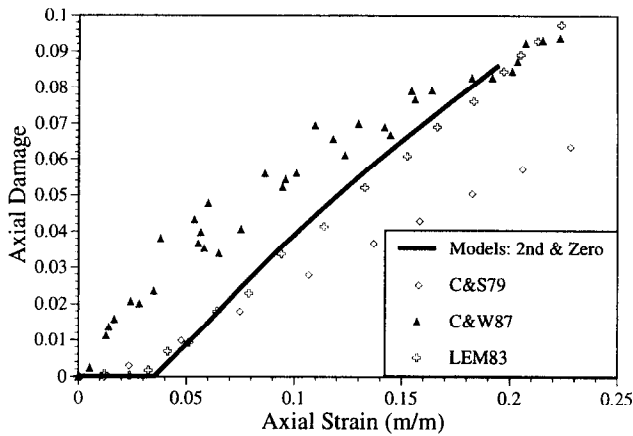


Fig. 5. Evolution of the damage for 2024-T3 aluminum alloy.

5. The agreement is similar to the correlation with the elastic modulus because of the dependencies given in eqns (117) and (118). Overall, the anisotropic coupled formulation correlates well with all measured quantities especially considering the scatter in the experimental data. The scalar isotropic model correlates with all quantities except the change in the apparent Poisson's ratio. These data sets represent a significant fraction of the available data on the direct measurement of damage in ductile materials, yet additional experimental investigations are warranted.

4.2. Temperature-induced ductile-brittle transition

It is generally observed that the ductility or amount of strain to failure (total damage) decreases with decreasing temperature of the material specimen or conversely increases with increasing temperature. In addition, the ultimate strength has an inverse relation with temperature. Therefore, a single material can at one temperature (cold) behave as a brittle material with relatively little strain to failure, while at another temperature (hot) the same material can have significantly greater strains to failure. This temperature-induced transition from ductile to brittle can have significant engineering implications but the transition is not always considered in design and analysis.

The intent of this section is to show that with some easily measurable material parameters, such as the initial yield strength and the undamaged elastic modulus as a function of temperature, the trends observed in the temperature induced ductile to brittle transition are replicated using the coupled elastoplasticity and damage formulation. Although the specimen temperature for each simulation is different, the process is still considered isothermal as was assumed in the theory formulation. In other words, the simulations are of a specimen already at a constant temperature.

Building on the results of the previous section, the aluminum alloy 2024-T3 is considered again. The coupled formulation with a second-order damage tensor is used exclusively in the remaining simulations. The previous parameters (Table 1) and results (Fig. 2) were at room temperature and are used as a baseline. The yield strength and the elastic modulus are temperature dependent. The ratio of these parameters to their values at room temperature (27°C) are taken from a handbook (MIL-HDBK-5E, 1987), and shown in Fig. 6 and Fig. 7, respectively. Both parameters are monotonically decreasing functions of the temperature.

The stress-strain response at two excursion temperatures, one lower and one higher than room temperature, are considered. The upper temperature selected is 160°C and the lower temperature is -160°C. At the lower temperature, the yield strength and the Young's modulus are 116% and 110% of the room temperature values, respectively. At the upper temperature, the yield strength and the Young's modulus are 88% and 94% of the room temperature values, respectively. Other material parameters including the damage parameters may possibly exhibit a temperature dependency but are considered to be temperature independent here.

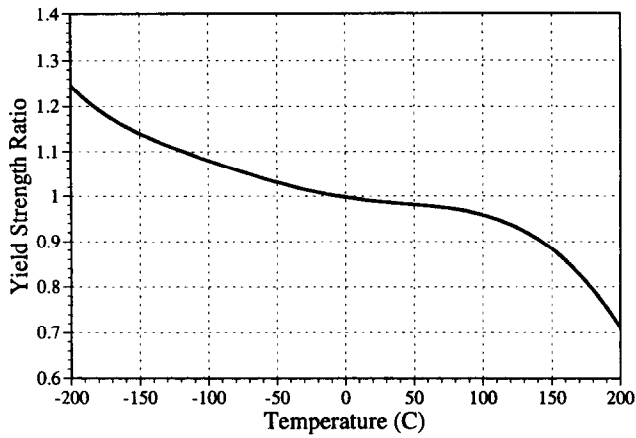


Fig. 6. Yield strength temperature dependency for 2024-T3 aluminum alloy sheet.

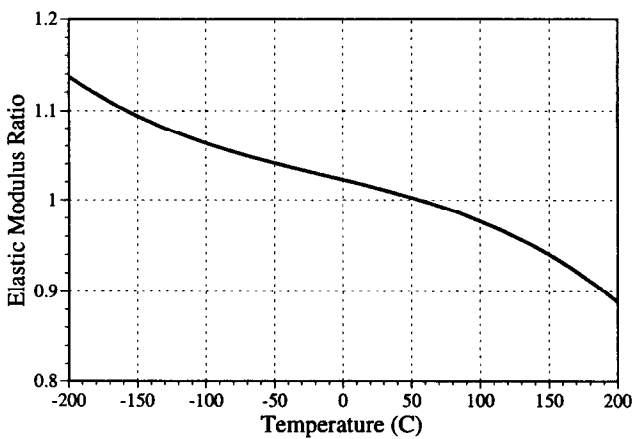


Fig. 7. Elastic modulus temperature dependency for 2024-T3 aluminum alloy sheet.

The stress-strain simulations for the reduced and elevated temperature excursions are given in Fig. 8. Notice the trend of increased ultimate strength with decreasing temperature is preserved. This is as expected due to the variation in the yield strength. Of greater interest is the terminal values of the damage. All simulations were performed to the same maximum strain values. The terminal value of the damage tensor component D_{11} is noted on the figure for each simulation. For most materials the specimen breaks in the post-peak or softening regime at a nonzero stress state. Although the experimental data in this regime

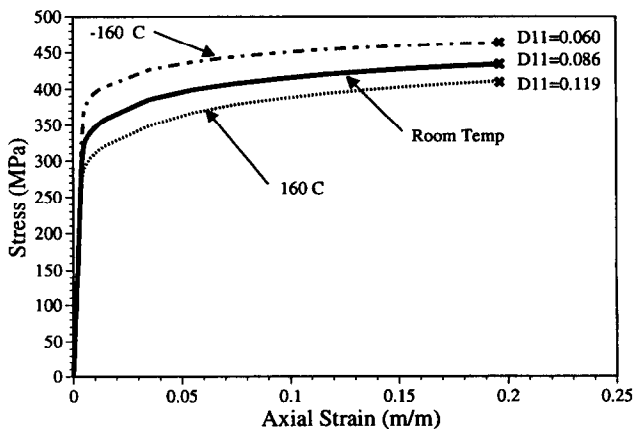


Fig. 8. Stress-strain response simulations of 2024-T3 aluminum alloy for various temperatures (constant total strain).

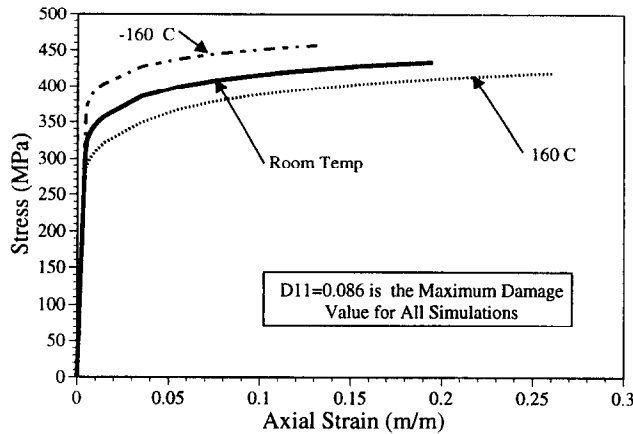


Fig. 9. Stress-strain response simulations of 2024-T3 aluminum alloy for various temperatures (constant damage value).

are scarce, at some point the damage becomes so intense that any further loading results in a catastrophic failure of the material. Therefore, a critical state is defined that, if exceeded, will result in drastic change in the material behavior. As opposed to the complete and immediate failure of the specimen at this critical state other phenomena may occur, for example the microcracks coalesce into dominate macrocracks and/or additional damage mechanisms are activated resulting in accelerated growth of damage. For the current simulations, if the critical damage value at room temperature is 0.086, then at the lower temperature the critical damage value is achieved before the terminal strain value, and for the higher temperature, additional loading is required before this critical value is reached. Hence, the strain corresponding to the critical damage value is less for the lower temperature and greater for the higher temperature. This trend is consistent with the observed phenomenon in many materials. An alternate presentation of the data is given in Fig. 9, where the simulations are terminated when the damage value of 0.086 is achieved. The variations in the strains to achieve the critical damage value are clearly exhibited.

Based on this simple analysis using handbook material parameters it is clear that the use of a coupled elastoplastic and damage formulation has potential utility for predicting the ductile to brittle transitional behavior as induced by the change in the temperature of the material specimen. Either theory used separately would not indicate trends such as these. Additional research into areas such as the temperature dependence on the damage evolution parameters is warranted.

5. SUMMARY

A rigorous thermodynamic framework for coupled elastoplastic and damage theories has been developed. This unified framework is sufficiently general to encompass elastoplastic and damage separately, in addition to coupled theories. General forms of effective-stress-space yield functions and damage functions were constructed using homogeneous functions of degree one. With Lagrange minimization and the principle of maximum entropy, the evolutionary relations were derived. In addition, the convexity of the undamaging elastic domain was shown. For plasticity the resulting evolution of the plastic strains corresponds to an associative flow. The satisfaction of the first and second laws were demonstrated.

This general framework was shown to be sufficiently general to describe many of the popular theories for both plasticity and damage. The broad class of J_2 plasticity with linear isotropic and kinematic hardening was presented. The procedure for extending this formulation to other hardening relations or adding a pressure dependence was discussed. Three common damage theories were considered. Theories using a scalar, a second-order tensor, and a fourth-order tensor as damage measures were formulated using the framework. Limitations of these damage theories were discussed. The utility of the framework was

demonstrated by formulating a coupled model using the specific elastoplasticity and damage models.

Two coupled formulations involving scalar isotropic damage and the second-order anisotropic damage were then used to replicate the experimental data for 2024-T3 aluminum alloy. The anisotropic formulation was shown to replicate all of the observed behavior. As expected, the scalar isotropic model did not correlate well with data exhibiting a change in the apparent Poisson ratio. In addition, the observed trends for temperature induced ductile-to-brittle transition was replicated using easily measured material parameters.

Acknowledgements—This work was supported by the U.S. Department of Energy under Contract DE-AC04-76DP00789. This paper is based in part on the first author's Ph.D. dissertation.

REFERENCES

- Bazant, Z. P. (1990). Recent advances in failure localization and nonlocal models. In *Micromechanics of Failure of Quasi-Brittle Materials* (Edited by S. P. Shah, S. E. Swartz and M. L. Wang), pp. 12–37. Elsevier, Amsterdam.
- Callen, H. B. (1985). *Thermodynamics and an Introduction to Thermostatistics*. Wiley, New York.
- Chen, Z. and Schreyer, H. L. (1990). Failure-controlled solution strategies for damage softening with localization. In *Micromechanics of Failure of Quasi-Brittle Materials* (Edited by S. P. Shah, S. E. Swartz and M. L. Wang), pp. 135–145. Elsevier, Amsterdam.
- Chow, C. L. and Wang, J. (1987a). An anisotropic theory of continuum damage mechanics for ductile fracture. *Engng Fract. Mech.* **27**, 547–558.
- Chow, C. L. and Wang, J. (1987b). An anisotropic theory of elasticity for continuum damage mechanics. *Int. J. Fract.* **33**, 3–16.
- Coleman, B. D. and Gurtin, M. E. (1967). Thermodynamics with internal state variables. *J. Chem. Phys.* **47**, 597–613.
- Cordebois, J. L. and Sidoroff, F. (1979). Damage Induced Elastic Anisotropy. Colloque Euromech 115, Villard de Lans. Also appearing in Boehler J. P. (eds) (1983), *Mechanical Behavior of Anisotropic Solids*. pp. 761–774, Martinus Nijhoff.
- Davis, H. T. (1960). *Introduction to Nonlinear Differential and Integral Equations*. U.S. Atomic Energy Commission.
- Hill, R. (1950). *The Mathematical Theory of Plasticity*. Oxford University Press, London.
- Ju, J. W. (1989). On energy-based coupled elastoplastic damage theories: constitutive modeling and computational aspects. *Int. J. Solids Structures* **25**, 803–833.
- Ju, J. W. (1990). Isotropic and anisotropic damage variables in continuum damage mechanics. *J. Engng Mech. ASCE*, **116**, 2764–2770.
- Kachanar, L. M. (1958). Time of rupture process under creep conditions. *IVZ Akad. Nauk, S.S.R., Otd. Tech. Nauk* **8**, 26–31.
- Krajcinovic, D. (1989). Damage mechanics. *Mech. Mater.* **8**, 117–197.
- Lemaitre, J. (1971). Evaluation of dissipation and damage in metals. *Proc. I.C.M.*, Vol. 1, Kyoto, Japan.
- Lemaitre, J. (1983). A continuous damage mechanics model for ductile fracture. *J. Engng Mater. Tech., ASME* **107**, 83–89.
- Lubliner, J. (1972). On the thermodynamic foundations of nonlinear solids mechanics. *Int. J. Non-Linear Mech.* **7**, 237–254.
- Lubliner, J. (1984). A maximum-dissipation principle in generalized plasticity. *Acta Mech.* **52**, 225–237.
- MIL-HDBK-5E (1987). *Metallic Materials and Elements for Aerospace Vehicle Structures*. U.S. Department of Defense.
- Simo, J. C. and Hughes, T. J. R. (1988). *Plasticity and Viscoplasticity: Numerical Analysis and Computational Aspects*. Springer, Berlin.
- Simo, J. C. and Ju, J. W. (1987). Stress and strain based continuum damage models, parts I and II. *Int. J. Solids Structures* **23**, 821–869.
- Stevens, D. J. and Liu, D. (1992). Strain based constitutive model with mixed evolution rules for concrete. *J. Engng Mech. ASCE* **118**, 1184–1200.
- Strang, G. (1986). *Introduction to Applied Mathematics*. Wellesley-Cambridge, Wellesley, MA.
- Yazdani, S. and Schreyer, H. L. (1990). Combined plasticity and damage mechanics model for plain concrete. *J. Engng Mech., ASCE* **116**, 1435–1450.
- Yazdani, S. and Karnawat, S. (1992). A constitutive theory for brittle solids with elementary coupled theory relations. In *Recent Advances in Damage Mechanics and Plasticity* (Edited by J. W. Ju) ASME AMD-Vol. 132, pp. 249–263.
- Ziegler, H. (1963). Some extremum principles in irreversible thermodynamics with application to continuum mechanics. In *Progress in Solid Mechanics*, Vol. 4 (Edited by I. N. Sneddon and R. Hill), North-Holland, Amsterdam.