Title
A Theory for Continua Undergoing Phase Transitions

Authors
Auricchio, Ferdinando
Lubliner, Jacob
Taylor, Robert

Publication Date
1993-12-01

A THEORY FOR CONTINUA UNDERGOING PHASE TRANSITIONS

F. Auricchio
J. Lubliner
R.L. Taylor
Department of Civil Engineering
University of California at Berkeley, Berkeley, CA 94720 USA

Contents

F.Auricchio, J.Lubliner, R.L.Taylor

1 Introduction

In the present work we present a continuum mechanics theory for materials undergoing solid-solid, rate-independent phase transition phenomena. The motivation for the work is originated from the increased practical interest in such materials, a classical examples being the so called *shape memory alloys* (Wayman 1990, Wayman 1992, Wayman 1993). Due to a reversible martensite phase transformation, the shape memory alloys present and intrinsic capacity of remembering their original shape and such properties have open novel frontiers in terms of applications (Duerig 1990).

To model shape memory alloys as well as many other materials undergoing (reversible with or without hysteresis, and/or multiple) phase transformations, we clearly need a general and flexible framework. Following the approach used in the *generalized plasticity* theory by Lubliner (1984, 1993), we will base our development on the concepts of *elastic range* and *loading direction.*

The paper is organized as follow: we start by discussing the assumptions and the basic framework within the context of an internal variable approach. We then describe the form of the evolutionary equation for the internal variables. Finally, as an example, an application to the case of martensitic transformations in shape memory alloys is outlined.

2 Theory

In this Section we develop the core of a theory for continua undergoing solid-solid rate-independent phase transitions. We first describe the assumptions and the basic framework, then we discuss an appropriate form for the flow rule.

2.1 Assumptions and basic framework

We start by assuming that:

- the material undergoing phase transition has a rate-independent behavior,

- its inelastic behavior can be modeled through the use of internal variables

Following Rice (1971) by internal variables we mean quantities which should be properly specified depending on the material considered and on the evolution processes we wish to model. The choice of internal variables may also depend on the scale at which we wish to describe the material; accordingly, they may be some macroscopic average quantities or microscopic parameters.

We assume that:

- a local thermo-mechanical state is represented by an ordered pair $(\mathbf{G}, \mathbf{q})$, where $\mathbf{G}$ stands for the set of controllable state variables (*control state*) and $\mathbf{q} = (q_1, q_2, ..., q_n)$ is the set of internal variables. An example of control state variables is:

$$\mathbf{G} = (\mathbf{E}, \theta)$$

where: $\mathbf{E}$ is the Green (Lagrangian) strain tensor and θ is the relative temperature. The set of all possible (realizable) states $(\mathbf{G}, \mathbf{q})$ is called *state space* $\mathcal{S}$. Moreover, we define:

$$\mathcal{S}_{\mathbf{q}} = \mathcal{S}|_{\mathbf{q}} \stackrel{\text{def}}{=} \{ \mathbf{G} \mid (\mathbf{G}, \mathbf{q}) \in \mathcal{S} \}$$

such that $\mathcal{S}_{\mathbf{q}}$ is the set of all possible control states $\mathbf{G}$ given a specific value to $\mathbf{q}$. Finally, note that the set of internal variables $\mathbf{q}$ is, like $\mathbf{E}$, unaffected by a Euclidean transformation of the spatial frame (they may be scalar or components of material tensors).

- there exists a sufficiently smooth function $\psi : \mathcal{S} \rightarrow \mathcal{R}$, such that $\psi = \psi(\mathbf{G}, \mathbf{q})$ is the the *Helmholtz free energy* per unit volume.

- there exists a dissipation function $D : \mathcal{S} \times \mathcal{R}^n \rightarrow \mathcal{R}$ defined by:

$$D = D(\mathbf{G}, \mathbf{q}, \mathbf{r}) = -\frac{\partial}{\partial \mathbf{q}} \psi(\mathbf{G}, \mathbf{q}) \cdot \mathbf{r} \quad , \quad \mathbf{r} \in \mathcal{R}^n$$

The value of D at $(\mathbf{G}, \mathbf{q}, \dot{\mathbf{q}})$ is the *specific dissipation* per unit reference volume and it is assumed to satisfy Kelvin inequality:

$$D = D(\mathbf{G}, \mathbf{q}, \dot{\mathbf{q}}) \geq 0$$

For a rate independent material, a **process** may be defined as a path (a directed curve, not necessarily simple) in state space $\mathcal{S}$; accordingly, indicating with t the curve arc-length (or a fictitious measure of time), we may think a process as a mapping:

$$t \mapsto \mathbf{G}(t)$$

In particular we distinguish between elastic and inelastic processes:

- **elastic process:** a process entirely contained in a manifold described by $\mathbf{q} = const.$ or equivalently a process for which $\dot{\mathbf{q}} = \mathbf{0}$;
- **inelastic process:** any other process.

Given a state $(\mathbf{G}, \mathbf{q})$, we may define the **elastic range** of such state as the set:

$$\mathcal{E}(\mathbf{G}, \mathbf{q}) = \{ \mathbf{G}^\star \mid \text{there exists an elastic process from } (\mathbf{G}, \mathbf{q}) \text{ to } (\mathbf{G}^\star, \mathbf{q}) \}$$

As a result of rate-independence, a process with $\mathbf{G} = const.$ is elastic. Accordingly, every control state $\mathbf{G}$ belongs to its elastic range, that is:

$$\mathbf{G} \in \mathcal{E}(\mathbf{G}, \mathbf{q})$$

therefore every state has a non-empty elastic range. It seems reasonable to assume that $\mathcal{E}(\mathbf{G}, \mathbf{q})$ is relatively closed in $\mathcal{S}_\mathbf{q}$; the relative boundary of $\mathcal{E}(\mathbf{G}, \mathbf{q})$ will be denoted by $\partial\mathcal{E}(\mathbf{G}, \mathbf{q})$ and the interior by $\overset{\circ}{\mathcal{E}}(\mathbf{G}, \mathbf{q})$. If the elastic range of a state $(\mathbf{G}, \mathbf{q})$ has a non-empty interior, then its boundary may be assumed to be a piecewise smooth surface in $\mathcal{S}_\mathbf{q}$, called the *loading surface* in the control variable $\mathbf{G}$-space.

It is interesting to observe that in classical plasticity the elastic range depends only on the set of internal variables $\mathbf{q}$:

$$\mathcal{E}(\mathbf{G}, \mathbf{q}) = \mathcal{E}(\mathbf{q}) = \mathcal{S}_\mathbf{q}$$

while in generalized plasticity (Lubliner 1984) it depends on the whole state:

$$\mathcal{E}(\mathbf{G}, \mathbf{q}) = \mathcal{E}(\mathbf{G}, \mathbf{q}) \subseteq \mathcal{S}_\mathbf{q}$$

We may now introduce the definition of elastic and inelastic state:

$$
\begin{array}{lcl}
(\mathbf{G},\mathbf{q}) \text{ is an } \textbf{elastic state} & \Leftrightarrow & \mathbf{G} \in \overset{\circ}{\mathcal{E}}(\mathbf{G},\mathbf{q}) \\
(\mathbf{G},\mathbf{q}) \text{ is an } \textbf{inelastic state} & \Leftrightarrow & \mathbf{G} \in \partial\mathcal{E}(\mathbf{G},\mathbf{q})
\end{array}
$$

It follows from this definition, that every state $(\mathbf{G}^{\star},\mathbf{q})$ in a sufficiently small neighborhood of an elastic state $(\mathbf{G},\mathbf{q})$ is attainable elastically. Consequently, in any possible process passing through an elastic state, $\dot{\mathbf{q}} = \mathbf{0}$.

The set of all elastic states in $\mathcal{S}$ will be called the *elastic domain* and denoted by $\mathcal{S}^E$; the projection of $\mathcal{S}^E$ into $\mathcal{S}_{\mathbf{q}}$ is the elastic domain at $\mathbf{q}$ and will be denoted by $\mathcal{S}^E_{\mathbf{q}}$, that is:

$$
\mathcal{S}^E_{\mathbf{q}} = \mathcal{S}^E|_{\mathbf{q}} \overset{\text{def}}{=} \left\{ \mathbf{G} \mid (\mathbf{G},\mathbf{q}) \in \mathcal{S}^E \right\}
$$

It is easy to show that:

$$
\mathcal{S}^E_{\mathbf{q}} \subset \mathcal{E}(\mathbf{G},\mathbf{q}) \quad \text{for every} \quad \mathbf{G} \in \mathcal{S}_{\mathbf{q}}
$$

As discussed and proved in Reference (Lubliner 1993), under simple topological assumptions:

- $\mathcal{S}^E_{\mathbf{q}}$ is open,
- $\mathcal{S}^E_{\mathbf{q}} = \overset{\circ}{\mathcal{E}}(\mathbf{G},\mathbf{q}) \quad \text{if} \quad \mathbf{G} \in \bar{\mathcal{S}}^E_{\mathbf{q}}$

We may also introduce the following sets:

$$
\begin{array}{rcl}
\mathcal{S}^I & = & \mathcal{S} - \mathcal{S}^E \\
\mathcal{S}^I_{\mathbf{q}} & = & \mathcal{S}_{\mathbf{q}} - \mathcal{S}^E_{\mathbf{q}}
\end{array}
$$

Note that a process whose range is entirely in the elastic region is reversible and not merely quasi-reversible.

Now consider an inelastic state $(\mathbf{G},\mathbf{q})$ and a vector $\mathbf{N}$ pointing away from $\mathcal{E}(\mathbf{G},\mathbf{q})$, called the *loading direction* [1] (refer to Figure 1). A path through $(\mathbf{G},\mathbf{q})$ such that:

$$
\mathbf{N} \cdot \dot{\mathbf{G}} < 0
$$

leads to an elastically attainable state and is therefore, at least locally, an elastic path, so that $\dot{\mathbf{q}} = \mathbf{0}$. On the other hand, a path with:

$$
\mathbf{N} \cdot \dot{\mathbf{G}} \geq 0
$$

[1] Note that $dim(\mathbf{N}) = dim(\mathbf{G})$.

leads toward states that are attainable only inelastically, that is with $\dot{\mathbf{q}} \neq \mathbf{0}$. At this point it is worth noting that we have not stated any assumption about the smoothness of the elastic range.

Moreover, it follows that if $(\mathbf{G}, \mathbf{q})$ is an inelastic state and $(\dot{\mathbf{G}}, \dot{\mathbf{q}})$ is the local tangent velocity of a path through $(\mathbf{G}, \mathbf{q})$, such that:

$$\left(\mathbf{G} + h\dot{\mathbf{G}}, \mathbf{q}\right) \in \mathcal{E}(\mathbf{G}, \mathbf{q})$$

for all sufficiently small positive h, then $\dot{\mathbf{q}} = \mathbf{0}$, so that $D(\mathbf{G}, \mathbf{q}, \dot{\mathbf{q}}) = 0$. We therefore have the following:

Theorem 2.1 *If:*

1. $(\boldsymbol{G}, \boldsymbol{q})$ *is a inelastic state,*
2. $(\dot{\boldsymbol{G}}, \dot{\boldsymbol{q}})$ *is the local tangent velocity of a path through* $(\boldsymbol{G}, \boldsymbol{q})$,
3. $(\boldsymbol{G} + h\dot{\boldsymbol{G}}, \boldsymbol{q}) \in \mathcal{E}(\boldsymbol{G}, \boldsymbol{q})$ *for* $|h|$ *sufficiently small,*
4. $D(\boldsymbol{G}, \boldsymbol{q}, \dot{\boldsymbol{q}}) > 0$

then $h < 0$.

2.2 Flow rule

The preceding considerations give an idea of the nature of the rate equation for $\mathbf{q}$, which governs the inelastic processes. If these equations have the form (Lubliner 1973):

$$\dot{\mathbf{q}} = \mathbf{f}(\mathbf{G}, \mathbf{q}, \dot{\mathbf{G}})$$

then:

- if $(\mathbf{G}, \mathbf{q})$ is an elastic state, that is

  $$\mathbf{G} \in \mathcal{S}^E_{\mathbf{q}}$$

  then: every $(\mathbf{G}^\star, \mathbf{q})$ in a sufficiently small neighborhood of $\mathbf{G}$ is attainable elastically, which implies $\dot{\mathbf{q}} = \mathbf{0}$ and accordingly:

  $$\mathbf{f}(\mathbf{G}, \mathbf{q}, \dot{\mathbf{G}}) = \mathbf{0}$$

- if $(\mathbf{G},\mathbf{q})$ is an inelastic state and if given a vector $\mathbf{N}$ pointing outward from the elastic range we have:

$$\mathbf{N}\cdot\dot{\mathbf{G}} < 0$$

then, again, $\dot{\mathbf{q}} = \mathbf{0}$, which implies:

$$\mathbf{f}(\mathbf{G},\mathbf{q},\dot{\mathbf{G}}) = \mathbf{0}$$

- for all the other cases

$$\mathbf{f}(\mathbf{G},\mathbf{q},\dot{\mathbf{G}}) \neq \mathbf{0}$$

Moreover for a rate independent material, $\mathbf{f}(\mathbf{G},\mathbf{q},\cdot)$ must be homogeneous of the first degree, that is:

$$\mathbf{f}(\mathbf{G},\mathbf{q},a\dot{\mathbf{G}}) = a\mathbf{f}(\mathbf{G},\mathbf{q},\dot{\mathbf{G}})$$

for any non negative real number a. The simplest rate equation having these properties is:

$$\dot{\mathbf{q}} = \mathbf{h} < \mathbf{N}\cdot\dot{\mathbf{G}} >$$

where $< . >$ is the Macauley bracket defined as:

$$< x >= \frac{1}{2}(x + |x|)$$

and:

$$\mathbf{h} = \mathbf{h}(\mathbf{G},\mathbf{q})$$

is a continuous function, null if $\mathbf{G} \in \overset{\circ}{\mathcal{E}}(\mathbf{G},\mathbf{q})$. Note that if $\partial\mathcal{E}(\mathbf{G},\mathbf{q})$ is locally given by:

$$\partial\mathcal{E}(\mathbf{G},\mathbf{q}) = \{\ \mathbf{G}^{\star} \mid f(\mathbf{G}^{\star};\mathbf{G},\mathbf{q}) = const.\ \}$$

then:

$$\mathbf{N} = \left.\frac{\partial f}{\partial \mathbf{G}^{\star}}\right|_{\mathbf{G}^{\star}=\mathbf{G}}$$

such that:

$$\dot{\mathbf{q}} = \mathbf{h} < \overset{\circ}{f} >$$

where:

$$\overset{\circ}{f} = \left. \frac{\partial f}{\partial \mathbf{G}^*} \right|_{\mathbf{G}^* = \mathbf{G}} \cdot \dot{\mathbf{G}}$$

If $\mathbf{h} \propto \mathbf{N}$, then the evolutionary equation for $\dot{\mathbf{q}}$ can be called *associative.*

3 An application to shape memory alloys

Now we wish to specialize the results from Section 2 to a simple case to demonstrate the generality and the flexibility of the theory. To do so, we choose as an example the shape memory alloys. However, to not overload the present discussion, we will only outline how their behavior can be modelled within the theory described in Section 2 and in a forthcoming work we will present all the necessary details.

Shape memory alloys present a peculiar macroscopic behavior (usually known as *shape memory effect* and *pseudo-elasticity* [Wayman 1990, Wayman 1992, Wayman 1993]), due to a microscopic martensitic phase transformation. A martensitic transformation is a solid-solid, diffusion-less crystallographic transformations, between an higher temperature phase, called *austenite*, and a lower temperature phase, called *martensite* (Khachaturyan 1983, Wayman 1964). For the case of shape memory alloys, the martensitic transformation is reversible and rate-independent (Funakubo 1987, Otsuka 1986).

In order to properly describe at least the basic features of shape memory alloys, we need to consider two different phase transformations. The first one is relative to the conversion of austenite into martensite ($A \rightarrow M$), while the second is relative to the conversion of martensite into austenite ($M \rightarrow A$). However, for simplicity we start by modelling only the first transformation and later extend the arguments to both.

Assume that the material is described by a set of two control variables:

$$\mathbf{G} = \{g_1, g_2\}$$

which for examples could be uniaxial stress (or strain) and temperature. With $\mathbf{q}$ we indicate the set of internal variables (left unspecified), which however are supposed to describe the actual material composition (appropriate phase fraction description). Looking at experimental results (Funakubo 1987, Otsuka 1986), we may immediately note that $\mathcal{S}_{\mathbf{q}}$ is independent of $\mathbf{q}$. Accordingly, in what follows, $\mathcal{S}_{\mathbf{q}}$ will be indicated by $\mathcal{S}$ to simplify the notation, that is, the points of $\mathcal{S}$ are $\mathbf{G}$, not $(\mathbf{G}, \mathbf{q})$. Moreover, there is a well defined subset of $\mathcal{S}$ where the $A \rightarrow M$ transformation may occur; such subset, indicated with $\mathcal{S}^I_{AM}$, may be described by:

$$\mathcal{S}^I_{AM} = \{ \text{ all } \mathbf{G} \text{ such that } F_1 F_2 < 0 \ \}$$

where F_1 and F_2 real-valued functions defined on $\mathcal{S}$ (refer to Figure 2). Accordingly, we may also define $\mathcal{S}^E_{AM}$ as the subset of $\mathcal{S}$ where the $A \rightarrow M$ transformation may never occur, that is:

$$\mathcal{S}^E_{AM} = \mathcal{S} - \mathcal{S}^I_{AM}$$

Note that the set $\mathcal{S}^E_{AM}$ may also be described as:

$$\mathcal{S}^E_{AM} = \{ \text{ all } \mathbf{G} \text{ such that } F_1 F_2 > 0 \ \}$$

and it is <u>not</u> a connected set. Again, looking at the experimental results, it is immediately seen that for any point in $\mathcal{S}^I_{AM}$, we may define a loading direction $\mathbf{N}$, that is, a direction $\mathbf{N}$ such that

$$\mathbf{N} \cdot \dot{\mathbf{G}} \geq 0 \quad \Leftrightarrow \quad \dot{\mathbf{q}} \neq \mathbf{0}$$

Accordingly, for any inelastic state:

$$\dot{\mathbf{q}} = \mathbf{h} < \mathbf{N} \cdot \dot{\mathbf{G}} >$$

Recalling the definition of inelastic region for the present problem, we may rewrite the evolutionary equation for the internal variables as:

$$\dot{\mathbf{q}} = \mathbf{a} < -F_1 F_2 >< \mathbf{N} \cdot \dot{\mathbf{G}} >$$

for any state in the state space $\mathcal{S}$ since the condition $-F_1 F_2 > 0$ imposed by the first Macauley bracket automatically verifies that a state is in $\mathcal{S}^I_{AM}$.

A similar treatment can be adopted for modelling the reverse transformation, i.e. the conversion of martensite into austenite ($M \rightarrow A$). Two functions F_3 and F_4 delimiting the inelastic region and a loading direction $\mathbf{M}$ must be introduced, such that we may write:

$$\begin{aligned} \mathcal{S}^I_{MA} &= \{ \text{ all } \mathbf{G} \text{ such that } F_3F_4 < 0 \ \} \\ \mathcal{S}^E_{MA} &= \mathcal{S} - \mathcal{S}^I_{MA} \\ &= \{ \text{ all } \mathbf{G} \text{ such that } F_3F_4 > 0 \ \} \end{aligned}$$

Accordingly, for this phase transformation we have:

$$\dot{\mathrm{q}} = \mathbf{b} < -F_3F_4 >< \mathbf{M} \cdot \dot{\mathbf{G}} >$$

Finally, the whole set of transformations can be described by the single evolutionary equation:

$$\dot{\mathrm{q}} = \mathbf{a} < -F_1F_2 >< \mathbf{N} \cdot \dot{\mathbf{G}} > + \mathbf{b} < -F_3F_4 >< \mathbf{M} \cdot \dot{\mathbf{G}} >$$

where the different Macauley brackets are able to identify which phase transformation is active. Moreover, we have:

$$\begin{aligned} \mathcal{S}^I &= \mathcal{S}^I_{AM} \cup \mathcal{S}^I_{MA} \\ \mathcal{S}^E &= \mathcal{S} - \mathcal{S}^I = \mathcal{S}^E_{AM} \cap \mathcal{S}^E_{MA} \end{aligned}$$

We observe that due to the flexibility described in Section 2, no complication of any sort is introduced from the relative position of the two inelastic regions $\mathcal{S}^I_{AM}$ and $\mathcal{S}^I_{MA}$; accordingly, they may or may not intersect and may even perfectly overlap.

4 Closure

In the present work we presented a theory for continua undergoing solid-solid, rate independent phase transitions. The inelastic behavior is modeled through the use of internal variables and the developments are based on the concepts of *elastic range* and *loading direction.* The theory may describe the behavior of materials undergoing reversible and/or multiple phase transformations. As example, an application to the case of martensitic transformations in shape memory alloys is outlined. In forthcoming works, we will detail such application, presenting numerical results and discussing the numerical implementation of the model within a valid computational framework.

References

[] T.W. Duerig, K.N. Melton, D.Stockel, and C.M.Wayman, (1990) *Engineering aspects of shape memory alloys*, Butterworth-Heinemann.

[] H. Funakubo, (1987) *Shape memory alloys*, Gordon and Breach Science Publishers, translated from the Japanese by J.B. Kennedy.

[] A.G. Khachaturyan, (1983) *Theory of structural transformations in solids*, John Wiley & Sons .

[] J. Lubliner, (1993) *Generalized plasticity theory*, Unpublished work.

[] J. Lubliner, (1973) *On the structure of the rate equations of materials with internal variables*, Acta Mechanica **17**, 109–119.

[] J. Lubliner, (1984) *A maximum-dissipation principle in generalized plasticity*, Acta Mechanica **52**, 225–237.

[] K. Otsuka and K.Shimizu, (1986) *Pseudoelasticity and shape memory effects in alloys*, International Metals Reviews **31** , 93–114.

[] J.R. Rice, (1971) *Inelastic constitutive relations for solids: an internal variable theory and its application to metal plasticity*, Journal of the Mechanics and Physics of Solids **19**.

[] C.M. Wayman, (1964) *Introduction to the crystallography of martensitic transformations*, Macmillan.

[] C.M. Wayman, (1992) *Shape memory and related phenomena*, Progress in Material science **36**, 203–224.

[] C.M. Wayman, (1993) *Shape memory alloys*, MRS bulletin **April**, 49–56.

[] C.M. Wayman and T.W. Duerig, (1990) *An introduction to martensite and shape memory*, Engineering aspects of shape memory alloys (T.W. Duerig, K.N. Melton, D.Stockel, and C.M.Wayman, eds.), pp. 3–20.

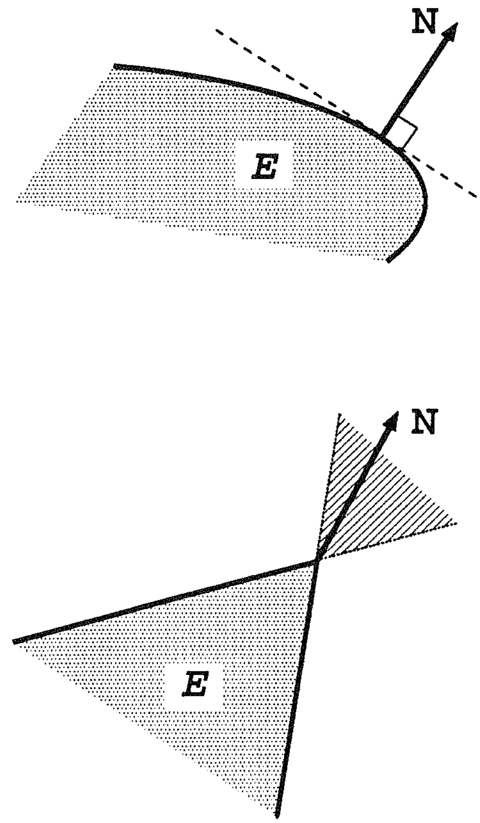

Figure 1: Graphical description of a vector $\mathbf{N}$ pointing away from the elastic range $\mathcal{E}(\mathbf{G}, \mathbf{q})$ (outward normal). For the case of a locally smooth elastic range, the outward normal is uniquely defined. For the case of a locally non-smooth elastic range, the outward normal must be contained in the cone of outward normals.

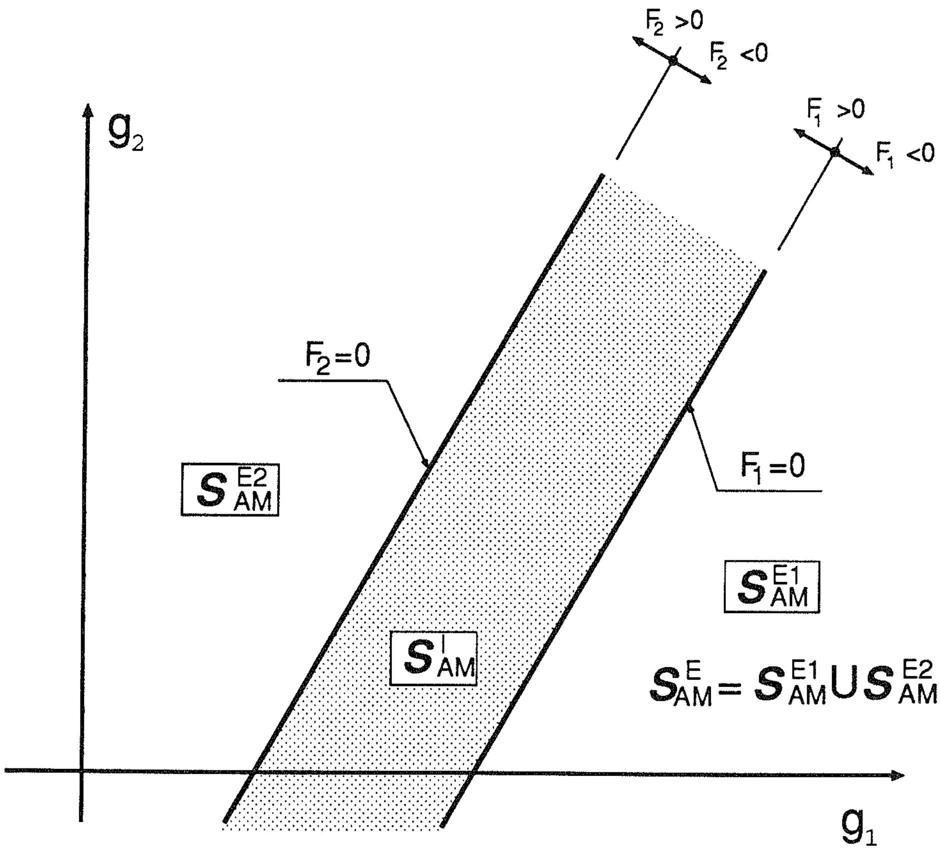

Figure 2: Definition of the F_1 and F_2 functions. $F_1 F_2 < 0$ represent the part of the state space where the $A \rightarrow M$ transformation may occur $(\mathcal{S}^I_{AM})$. $F_1 F_2 > 0$ represent the part of the state space where the $A \rightarrow M$ transformation may not occur, that is the elastic region relative to the transformation $(\mathcal{S}^E_{AM})$.

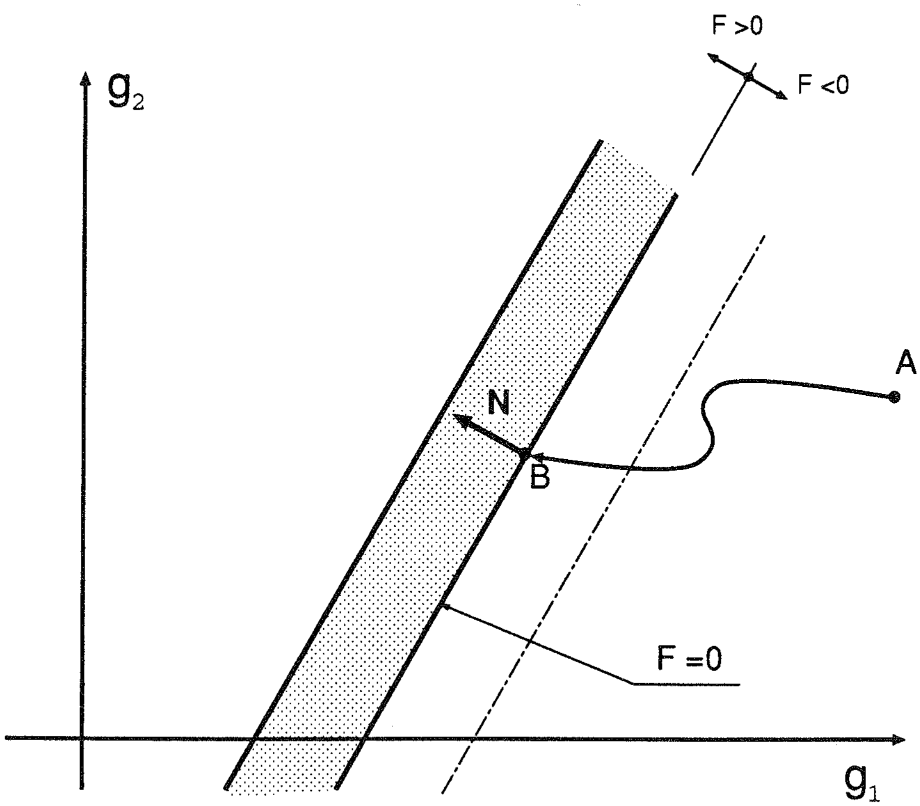

Figure 3: Definition of the loading function F.

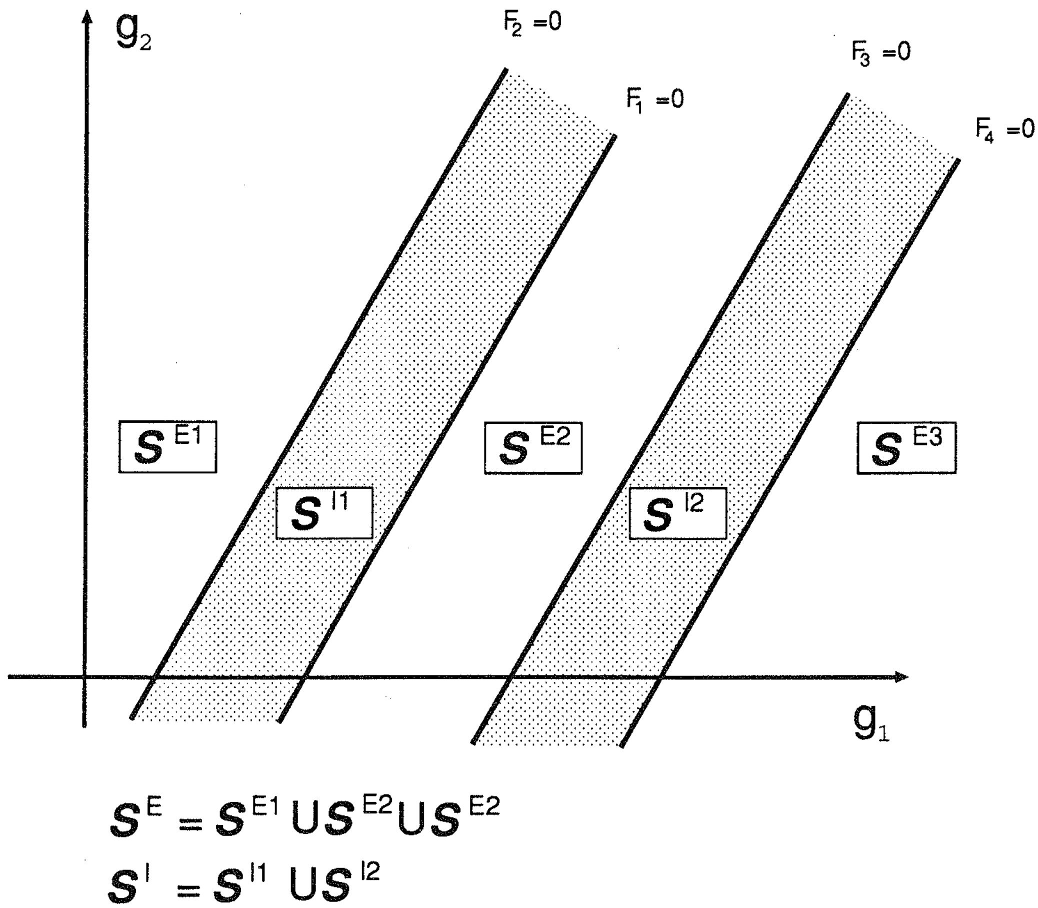

Figure 4: Distinction between elastic ($\mathcal{S}^E$) and inelastic region ($\mathcal{S}^I$)