

## Chapter 2

# General concepts

### 2.1. Formulation of the constitutive equations

The constitutive equations presented in this section are generally expressed as ordinary differential equations (ODE), so that the mechanical response depends on the present load and on its history (hereditary behavior). Two strategies can be considered to account for this property. The first one introduces a functional dependency between the variables. In the second one, it is assumed that the history can be “summarised” in a series of internal variables, which allow the material to have the memory of the past history, and to have a full description of its present state. Except for the case of linear viscoelasticity, the models developed in the framework of the second approach are easier to manage than those coming from the first approach.

The other important assumptions which are classically used in the development of constitutive equations are the following:

1. the *local state* principle, which assumes that the behavior in a given point  $M$  depends only on the variables defined in this point, and not from the surrounding;
2. the principle of *material simplicity*, which states that the mechanical behavior depends only on the first gradient of the transformation tensor;
3. the *objectivity* principle, expressing that the behavior does not depend on the observer; a consequence is that time cannot be used explicitly in the constitutive equations.

All these assumptions will be revisited in Chap. 6. For the case of homogeneous isotropic materials, they can be summarized by a relation between stresses and strains,

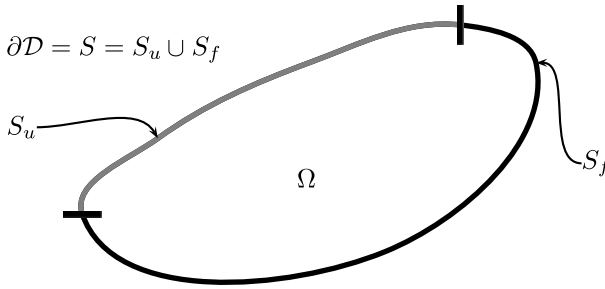
such as:

$$\sigma(t) = \mathcal{F}_{-\infty}^t(\varepsilon(\tau)) \quad (2.1)$$

## 2.2. Principle of virtual power

The present section gives a “minimal overview” of the principle of virtual power. The reader is invited to go to Chap. 6 on finite transformations for a comprehensive description.

The principle tells us that, for any field of virtual velocity,  $\mathbf{v}'(\mathbf{x}, t)$ , the sum of the power of internal and external forces is equal to zero, for any domain  $\mathcal{D}$ , with an external surface  $\partial\mathcal{D} = S$  in a solid  $\Omega$  (Fig. 2.1).



**Figure 2.1.** Domain  $\mathcal{D}$ , surface  $\partial\mathcal{D} = S$  in a solid  $\Omega$

The virtual power of the internal forces can be evaluated by means of the double dot product of the *real stress tensor* by the virtual strain-rate tensor:

$$\mathcal{P}^{(i)} = - \int_{\mathcal{D}} \underline{\underline{\sigma}} : \underline{\underline{\dot{\varepsilon}}}' dV \quad (2.2)$$

A classical series of manipulations can be made on this equation:

- due to the symmetry of the stress tensor  $\underline{\underline{\sigma}}$ , the strain rate can be replaced by the symmetrical part of the velocity gradient, then by the velocity gradient;
- the resulting equation can be integrated by parts, and the divergence theorem can be applied, resulting in a partition of the internal power into a volumetric term and a surface term;
- denoting by  $\underline{n}$  the outgoing normal at a given point of  $\partial\mathcal{D}$ , the following new expression can be obtained:

$$\mathcal{P}^{(i)} = \int_{\mathcal{D}} (\underline{\text{div}} \underline{\sigma}) \cdot \underline{v}' dV - \int_{\partial \mathcal{D}} (\underline{\sigma} \cdot \underline{n}) : \underline{v}' dS \quad (2.3)$$

The power of external forces consists in a volumetric term, for long distance actions, and a second one, representing surface forces:

$$\mathcal{P}^{(e)} = \int_{\mathcal{D}} \underline{f} \cdot \underline{v}' dV + \int_{\partial \mathcal{D}} \underline{F} \cdot \underline{v}' dS \quad (2.4)$$

The application of the principle leads to

$$\int_{\mathcal{D}} (\underline{\text{div}} \underline{\sigma} + \underline{f}) \cdot \underline{v}' dV - \int_{\partial \mathcal{D}} (\underline{\sigma} \cdot \underline{n} - \underline{F}) \cdot \underline{v}' dS = 0 \quad (2.5)$$

Since no assumption has been made on the velocity field, it can now be taken equal to zero successively on the external surface, or in any domain inside the body; it follows that

$$\underline{\text{div}} \underline{\sigma} + \underline{f} = \underline{0} \quad \forall \underline{x} \in \mathcal{D} \quad (2.6)$$

$$\underline{\sigma} \cdot \underline{n} = \underline{F} \quad \forall \underline{x} \in \partial \mathcal{D} \quad (2.7)$$

$$(2.8)$$

The equilibrium of the body and the boundary conditions are obtained by choosing for  $\mathcal{D}$  the solid  $\Omega$  itself.

The same calculation scheme can be applied by using a *statically admissible* (SA) stress tensor field  $\underline{\sigma}^*$ , and a *kinematically admissible* (KA) velocity field  $\underline{v}'$  (*a priori* they are not linked by a constitutive equation), which verify respectively (2.9) and (2.10), and (2.11):

$$\underline{\text{div}} \underline{\sigma}^* + \underline{f} = \underline{0} \quad \text{in } \Omega \quad (2.9)$$

$$\sigma_{ij}^* n_j = F_i \quad \text{on } S_f^i \quad (2.10)$$

$$v'_i = v_i^d \quad \text{on } S_d^i \quad (2.11)$$

where  $S_f^i$  are the surfaces where the component  $i$  of the surface forces  $\underline{F}$  is imposed,  $S_d^i$  are the surfaces where the component  $i$  of the velocity is imposed, with

$$S_d^i \cap S_f^i = \emptyset \quad S_d^i \cup S_f^i = S \quad \forall i = 1, 2, 3$$

It follows that

$$\int_{\Omega} \underline{\sigma}^* : \underline{\dot{\epsilon}}' dV = \int_{\Omega} \underline{f} \cdot \underline{\dot{u}}' dV + \int_S \underline{F} \cdot \underline{\dot{u}}' dS \quad (2.12)$$

### 2.3. Thermodynamics of irreversible processes

The purpose of this section is to provide a few comments on the balance of energy, taking into account the effect of a mechanical strain, and the interactions between temperature, elastic strain and plastic strain. The development uses the small strain formalism; additional terms needed for a finite strain framework are shown in Chap. 6.

#### 2.3.1. First and second principles of thermodynamics

After mass conservation and equilibrium equations, the third fundamental conservation law in Continuum Mechanics expresses conservation of energy: this is the first principle of thermodynamics. After this principle, *for any time during the thermomechanical process, the convective derivative of the total energy of a system is the sum of the power of the external applied forces,  $\mathcal{P}^{(e)}$ , and of the rate of received heat,  $\dot{Q}$* . The energy is the sum of the internal energy and of the kinetic energy, which will not be considered here. The principle can be written by introducing the notation  $E$  for the internal energy on a domain  $\mathcal{D}$ , and the specific internal energy  $e$ :

$$\frac{dE}{dt} = \int_{\mathcal{D}} \rho \frac{de}{dt} dV = \mathcal{P}^{(e)} + \dot{Q} \quad (2.13)$$

If the real fields are used as SA and KA fields, the power of the external forces can be written as

$$\mathcal{P}^{(e)} = \int_{\mathcal{D}} \underline{\sigma} : \underline{\dot{\epsilon}} dV \quad (2.14)$$

The term  $\dot{Q}$  is the sum of two different contributions. The first one corresponds to the surface density of the received heat rate, which is defined as a function of the vector  $\underline{q}$  and of the normal  $\underline{n}$ , ( $\underline{n}$  is the outgoing normal to the surface  $\partial\mathcal{D}$  of the domain considered). The second,  $r$ , represents a volumetric heat coming from sources like chemical reactions, phase transformation or induction heating:

$$\dot{Q} = \int_{\mathcal{D}} r dV - \int_{\partial\mathcal{D}} \underline{q} \cdot \underline{n} dS = \int_{\mathcal{D}} (r - \text{div } \underline{q}) dV \quad (2.15)$$

The variation of the specific internal energy is then the sum of the rate of specific energy due to internal forces and of the rate of specific heat received:

$$\rho \frac{de}{dt} = \underline{\sigma} : \underline{\dot{\epsilon}} + r - \text{div } \underline{q} \quad (2.16)$$

The second principle provides an upper bound of the rate of heat that can be taken by the volume  $\mathcal{D}$  at a temperature  $T$ , and can be written as a function of the entropy  $S$  or of the specific entropy  $s$ :

$$\frac{dS}{dt} \geq \int_{\mathcal{D}} \frac{r}{T} dV - \int_{\partial\mathcal{D}} \frac{\underline{q} \cdot \underline{n}}{T} dS \quad (2.17)$$

$$\text{hence } \int_{\mathcal{D}} \left( \rho \frac{ds}{dt} - \frac{r}{T} + \text{div} \left( \frac{\underline{q}}{T} \right) \right) dV \geq 0 \quad (2.18)$$

Helmoltz free energy is denoted by  $\psi$ . Using the equality  $e = \psi + Ts$ , together with (2.16) and (2.18), the following relation, called Clausius' inequality, holds:

$$\underline{\sigma} : \underline{\dot{\epsilon}} - \rho \frac{d\psi}{dt} - \rho s \dot{T} - \frac{1}{T} \underline{q} \cdot \underline{\text{grad}}(T) \geq 0 \quad (2.19)$$

### 2.3.2. Dissipation

The *method of local state* assumes that the thermodynamical state of the continuum at a given point and a given time is completely given by the definition of a set of *state variables*. The free energy defines a potential, depending on temperature and on the state variables  $\alpha_I$ , which characterizes the mechanical system. It is then possible to express the derivative of  $\psi$  as

$$\frac{d\psi}{dt} = \frac{\partial\psi}{\partial T} \dot{T} + \frac{\partial\psi}{\partial\alpha_I} \dot{\alpha}_I \quad (2.20)$$

Using this expression in (2.19), the following expressions are successively found:

$$s = -\frac{\partial\psi}{\partial T} \quad (2.21)$$

$$\underline{\sigma} : \underline{\dot{\epsilon}} - \rho \frac{\partial\psi}{\partial\alpha_I} \dot{\alpha}_I - \frac{1}{T} \underline{q} \cdot \underline{\text{grad}}(T) \geq 0 \quad (2.22)$$

Two types of volumetric contributions can be seen in (2.22), a mechanical “intrinsic” dissipation,  $\phi_1$ , and a purely thermal dissipation,  $\phi_2$ :

$$\phi_1 = \underline{\sigma} : \underline{\dot{\epsilon}} - \rho \frac{\partial\psi}{\partial\alpha_I} \dot{\alpha}_I \quad (2.23)$$

$$\phi_2 = -\frac{1}{T} \underline{q} \cdot \underline{\text{grad}}(T) \quad (2.24)$$

### 2.3.3. Heat equation

Classically, intrinsic and thermal dissipation are assumed to be uncoupled, so that one has to check the positivity of each of them. A positive thermal dissipation is obtained with Fourier's conduction law. For the case of an isotropic material, it can be written:

$$\underline{q} = -k(T, \alpha_I) \underline{\text{grad}}(T) \quad (2.25)$$

The scalar function  $k$ , whose values are strictly positive, describes the conduction (units: W/m/K). The expression enforces the positivity of  $\phi_2$ . If the conduction is anisotropic, it will be described by a non-negative quadratic form:

$$\phi_2 = \frac{1}{T} \underline{\text{grad}} T \cdot \underline{k} \cdot \underline{\text{grad}} T \quad (2.26)$$

Combining Fourier's law with the relation (2.16), and using the expression of the derivative of  $\psi$  with respect to the temperature and the state variables, the so-called *heat equation* is obtained:

$$\text{div} \left( k \underline{\text{grad}}(T) \right) = \rho C_\varepsilon \dot{T} - r - \underline{\sigma} : \underline{\dot{\varepsilon}} + \rho \left( \frac{\partial \psi}{\partial \alpha_I} - T \frac{\partial^2 \psi}{\partial T \partial \alpha_I} \right) \dot{\alpha}_I \quad (2.27)$$

The equation provides a law for the evaluation of the temperature evolution in presence of mechanical deformation. One has defined a specific heat at constant strain,  $C_\varepsilon = T \partial s / \partial T$  (unit: J/kg/K).

The left-hand side of the equation is proportional to the Laplacian  $\Delta T$  for the case of an isotropic conduction with a constant  $k$  coefficient. Since the unit of  $\rho$  is kg/m<sup>3</sup>, the product  $\rho C_\varepsilon$ , which characterizes the amount of heat trapped in the material, thus producing the temperature elevation, is in N/m<sup>2</sup>/K. The product  $\rho C_\varepsilon \dot{T}$  is then consistent with the volumetric term  $r$ , which is expressed in J/m<sup>3</sup>, and with the mechanical terms  $\underline{\sigma} : \underline{\dot{\varepsilon}}$  (Pa/s). The shape of the present expression does not depend on the mechanical constitutive equations. The differences between the materials will simply be taken into account by the shape of the free energy  $\psi$  and by the nature of state variables. The dissipated mechanical energy is represented by the term:  $\phi_1 = \underline{\sigma} : \underline{\dot{\varepsilon}} - \rho (\partial \psi / \partial \alpha_I) \dot{\alpha}_I$ , which takes into account the energy input, and the stored part of this energy (the storage can be temporary or not). A thermomechanical coupling can be exhibited at the level of the state variables if the cross derivative  $\partial^2 \psi / (\partial T \partial \alpha_I)$  is non-zero.

### 2.3.4. Linear thermoelasticity

The strain partition involves only an elastic and a thermal part: the elastic strain is taken equal to zero for a reference state  $\underline{\sigma}^I$ , and the thermal dilatation is equal to

zero for a reference temperature  $T^I$ . For the isotropic case, a linear thermoelasticity law can simply be obtained by choosing the following expression for the free energy, where  $K$  characterizes the compressibility of the material ( $3K = 3\lambda + 2\mu$ ,  $\lambda$  and  $\mu$  are the Lamé constants) and  $\alpha$  the linear coefficient of thermal expansion:

$$\begin{aligned} \psi = & \underline{\sigma}^I : \underline{\varepsilon} + \frac{1}{2} \lambda (\text{trace}(\underline{\varepsilon}))^2 + \mu \underline{\varepsilon} : \underline{\varepsilon} - 3K\alpha \text{trace}(\underline{\varepsilon})(T - T^I) \\ & - \frac{1}{2} \frac{\rho C_\varepsilon}{T^I} (T - T^I)^2 \end{aligned} \quad (2.28)$$

The state variable is the elastic strain. Since the thermoelastic behavior is a reversible process, the intrinsic dissipation,  $\phi_1 = (\underline{\sigma} - \rho \partial \psi / \partial \underline{\varepsilon}) : \dot{\underline{\varepsilon}}$ , must be equal to zero for any infinitesimal increment of the thermoelastic strain. As a consequence, the partial derivative of  $\psi$  provides the elastic constitutive equations:

$$\underline{\sigma} = \rho \frac{\partial \psi}{\partial \underline{\varepsilon}} = \underline{\sigma}^I + \lambda \text{trace}(\underline{\varepsilon}) \underline{\mathbf{I}} + 2\mu \underline{\varepsilon} - 3K\alpha(T - T^I) \underline{\mathbf{I}} \quad (2.29)$$

$$\underline{\sigma} - \underline{\sigma}^I = \lambda(\text{trace}(\underline{\varepsilon}) - 3\alpha(T - T^I)) \underline{\mathbf{I}} + 2\mu \underline{\varepsilon} \quad (2.30)$$

The variation of the specific entropy can also be obtained:

$$\rho s = -\rho \frac{\partial \psi}{\partial T} = 3K\alpha \text{trace}(\underline{\varepsilon}) + \frac{\rho C_\varepsilon}{T^I} (T - T^I) \quad (2.31)$$

This expression is consistent with the definition of  $C_\varepsilon$  in the vicinity of  $T^I$ ,  $C_\varepsilon = T^I \partial s / \partial T$ . One can easily illustrate the difference between the adiabatic and the isothermal evolutions on this simple case. Assuming that  $s = 0$  in (2.31), the variation of temperature produced by a volume change  $\Delta V / V = \text{trace}(\underline{\varepsilon})$  can then be expressed:

$$(T - T^I) = -\frac{3K\alpha T^I}{\rho C_\varepsilon} \text{trace}(\underline{\varepsilon}) \quad (2.32)$$

Introducing this expression in the constitutive equation (2.30), one can compare the responses:

$$\text{– isothermal: } \underline{\sigma} - \underline{\sigma}^I = \lambda \text{trace}(\underline{\varepsilon}) \underline{\mathbf{I}} + 2\mu \underline{\varepsilon} \quad (2.33)$$

$$\text{– adiabatic: } \underline{\sigma} - \underline{\sigma}^I = \left( \lambda + \frac{9K^2 \alpha^2 T^I}{\rho C_\varepsilon} \right) \text{trace}(\underline{\varepsilon}) \underline{\mathbf{I}} + 2\mu \underline{\varepsilon} \quad (2.34)$$

The shear modulus is left unchanged for the two cases (isothermal and adiabatic) but the first Lamé coefficient  $\lambda$  is larger for an adiabatic evolution. On the other hand, (2.31) shows that, for an adiabatic loading, a positive volume change (for instance in pure tension) will produce a decrease of the temperature, and *vice-versa*. The order of magnitude of these perturbations is small. For instance, typical values for steels are

( $E = 200 \text{ GPa}$ ,  $\nu = 0.3$ ,  $\rho = 8000 \text{ kg/m}^3$ ,  $C_\varepsilon = 500 \text{ J/kg/K}$ ,  $\alpha = 15 \times 10^{-6}$ ), such that:

$$\frac{\Delta T}{T} = -\frac{3K\alpha}{\rho C_\varepsilon} \frac{\Delta V}{V} = -1.875 \frac{\Delta V}{V} \quad (2.35)$$

For an anisotropic body, the thermoelastic law can be written as:

$$\underline{\underline{\sigma}} = \underline{\underline{\sigma}}^I + \underline{\underline{\Lambda}} : (\underline{\underline{\varepsilon}} - \underline{\underline{\varepsilon}}^{th}) \quad (2.36)$$

where  $\underline{\underline{\Lambda}}$  is a fourth-rank tensor characterizing the elastic moduli, and  $\underline{\underline{\varepsilon}}^{th}$  a second-rank tensor characterizing thermal dilatation  $\underline{\underline{\varepsilon}}^{th} = \underline{\underline{\alpha}}(T - T^I)$ ,  $\underline{\underline{\alpha}}$  the second-rank elastic strain tensor.

### 2.3.5. Nonlinear behavior

#### Expression of the dissipation

In a small perturbation framework, the total strain is partitioned into a thermoelastic (elastic and thermal dilatation)  $\underline{\underline{\varepsilon}}^e$ , and an inelastic strain  $\underline{\underline{\varepsilon}}^p$ :

$$\underline{\underline{\varepsilon}} = \underline{\underline{\varepsilon}}^e + \underline{\underline{\varepsilon}}^p \quad (2.37)$$

Note the slight notation change, since the term  $\underline{\underline{\varepsilon}}^e$  stands for the sum of the elastic and the thermal parts. Another modification is introduced in order to select between:

- the variables  $\alpha_I$  that characterize the *hardening* in the material (*id est* the evolution of the properties by those mechanisms which preserve the continuum, namely local stress redistribution, variation of the dislocation density, ...). The notation  $\alpha_I$  will still be in use for them;
- variables characterizing *damage*, corresponding to material degradation related to porosity growth or microcrack propagation. These variables will be named  $d_J$ .

The most general set of state variables includes then  $(\underline{\underline{\varepsilon}}^e, \alpha_I, d_J)$ , and the intrinsic dissipation is:

$$\phi_1 = \left( \underline{\underline{\sigma}} - \rho \frac{\partial \psi}{\partial \underline{\underline{\varepsilon}}^e} \right) : \dot{\underline{\underline{\varepsilon}}}^e + \underline{\underline{\sigma}} : \dot{\underline{\underline{\varepsilon}}}^p - \rho \frac{\partial \psi}{\partial \alpha_I} \dot{\alpha}_I - \rho \frac{\partial \psi}{\partial d_J} \dot{d}_J \quad (2.38)$$

Since thermoelasticity is a thermoelastic process:

$$\phi_1 = \underline{\underline{\sigma}} : \dot{\underline{\underline{\varepsilon}}}^p - \rho \frac{\partial \psi}{\partial \alpha_I} \dot{\alpha}_I - \rho \frac{\partial \psi}{\partial d_J} \dot{d}_J \quad (2.39)$$



The complete definition of the model is not made unless a constitutive equation is chosen. This will be made later. The second and third terms in the right-hand side describe the free energy change in the representative material element at constant temperature. This is the amount of energy which is stored in the material due to its hardening, or which is used to break open cracks or defects. This energy must be subtracted from the plastic power to get the amount of energy transformed into heat. Having a positive value of  $\Phi_1$  is equivalent to having an irreversible process that produces heat. The increase of temperature is given by:

$$\operatorname{div}(k \mathbf{grad}(T)) = \rho C_\varepsilon \dot{T} - r - \phi_1 - \phi_{is} \quad (2.40)$$

where  $\phi_{is}$  represents the sum of the so called “isentropic” terms, deduced from the second term in parenthesis in the right-hand side of (2.27). Its origin is the temperature dependence of the thermodynamical forces.

### *Approximated evaluation of the dissipation*

In the literature, there is generally no specific connection between the expression of the constitutive equation and the intrinsic dissipation term, and this dissipation is simply evaluated as a fraction  $\beta$  of the plastic power (with  $0 \leq \beta \leq 1$ ). This is only an approximation, as explained by (2.39). The value classically chosen for  $\beta$  is larger than 0.9, specially for the case of large strains. This has to be related to the fact that, in this type of regime, the hardening is saturated, so that the material is no longer able to store energy by means of any hardening process.

The case of an adiabatic loading can be introduced in this relation, by mentioning that the term  $\operatorname{div}(k \mathbf{grad}(T))$  becomes zero. An estimation of the temperature evolution can be written (assuming no volumetric heat source for the sake of brevity):

$$\rho C_\varepsilon \dot{T} = \phi_1 \quad (2.41)$$

The temperature increase will be significant for the case of rather rapid loadings, preventing any heat leak from happening. Consequently, unless thermoplastic coupling must be taken into account, for instance for forging problems, it will produce temperature increase for any type of external load (e.g., tension or compression).

### *Generalized standard model*

The thermodynamical forces  $A_I$  (resp.  $y_J$ ) associated to the state variables  $\alpha_I$  (resp.  $d_J$ ) are defined by:

$$A_I = \rho \frac{\partial \psi}{\partial \alpha_I} \quad y_J = -\rho \frac{\partial \psi}{\partial d_J} \quad (2.42)$$

where the components of the stress tensor and the variables  $A_I$  and  $y_J$  are put together in the vector  $Z$ . The vector  $z$  collects plastic strain and the state variables. The intrinsic dissipation is now:

$$\Phi_1 = \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}^p - A_I \dot{\alpha}_I + y_J \dot{d}_J = Z \dot{z} \quad (2.43)$$

with:

$$Z = \{\boldsymbol{\sigma}, A_I, y_J\} \quad z = \{\boldsymbol{\varepsilon}^p, -\alpha_I, d_J\} \quad (2.44)$$

A model will be “generalized standard” if and only if there exists a *potential*  $\Omega(Z)$  such that:

$$\dot{z} = \frac{\partial \Omega}{\partial Z} \quad (2.45)$$

If  $\Omega$  is a convex function of  $Z$ , and contains the origin, the intrinsic dissipation is automatically positive, since, for any point  $Z$ , all the points of the domain defined by the equipotential surface  $\Omega = Cte$  are located on the same side of the tangent plane, defined by the normal  $\partial \Omega / \partial Z$ , and:

$$\phi_1 = Z \frac{\partial \Omega}{\partial Z} \quad (2.46)$$

Using the Legendre–Fenchel’s transform, one can introduce  $\Omega^*$ , depending on the time derivative of the state variables:

$$\Omega^*(\dot{z}) = \max_Z (Z \dot{z} - \Omega(Z)) \quad (2.47)$$

The variables  $Z$  can then be obtained as the partial derivative of  $\Omega^*$  with respect to  $\dot{z}$ . The main assumption in the generalized standard models is that the information needed to predict the nonlinear behavior of the material can be captured in two potentials:

- free energy, defining the relationships between the state variables and the associated thermodynamical forces, and the constitutive equations of reversible phenomena;
- the dissipation potential, providing information on the irreversibility of the process.

### *State coupling, dissipation coupling*

A state coupling [MAR89] will result from the presence in the free energy of terms involving the product of two state variables. In the partial derivative, there is an additional term in the thermodynamical force. This type of coupling will be considered for instance between damage and elasticity (see Chap. 4). It introduces a symmetry

of the interactions, for instance for two variables  $A_1$  and  $A_2$ , thermodynamical forces associated with  $\alpha_1$  and  $\alpha_2$ :

$$\frac{\partial A_1}{\partial \alpha_2} = \frac{\partial A_2}{\partial \alpha_1} = \frac{\partial^2 \phi}{\partial \alpha_1 \partial \alpha_2} \quad (2.48)$$

On the other hand, a dissipation coupling is introduced each time there are several potential functions  $\Omega_K$ , such as:

$$\dot{z} = \sum_K \frac{\partial \Omega_K}{\partial Z} \quad (2.49)$$

This last type of coupling can be seen in “multisurface” models, or in crystal plasticity.

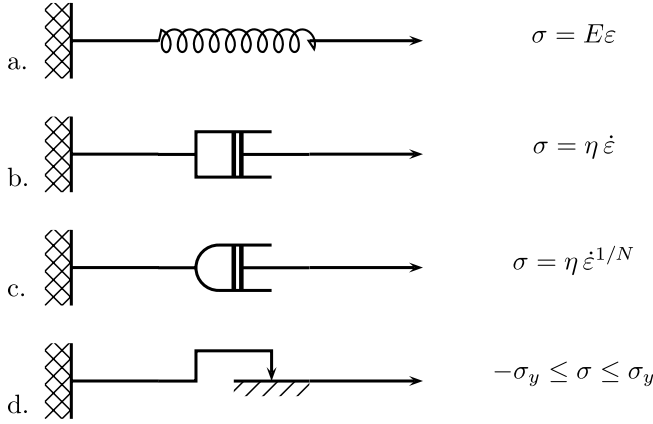
## 2.4. Main class of constitutive equations

### 2.4.1. Basic building blocks

The nature of the qualitative response of materials to a few simple tests allows us to define several classes. These basic material behaviors can be illustrated by elementary mechanical systems. There are only three types, elasticity, time independent plasticity, and viscosity. The corresponding elements are:

1. The spring, characterizing linear elasticity, for which strain is totally reversible after unloading, and providing a one-to-one relation between stress and strain components (Fig. 2.2a).
2. The dash-pot, representing linear (Fig. 2.2b) or nonlinear (Fig. 2.2c) viscosity. The viscosity is said to be “pure” if a one-to-one relation between load and strain rate is introduced. If the relation is linear, the corresponding model is Newton’s law.
3. The frictional device, used to illustrate a basic behavior where permanent strain appears if the load is large enough. If the corresponding threshold does not depend on the subsequent loading, the behavior is perfectly plastic. If, in addition, the elastic strain before permanent strain is neglected, the model is simply rigid-plastic.

All these elements can be combined together in order to form rheological models. They represent mechanical systems that can help to define various models. This type of construction has nothing to do with the deformation mechanisms, but it can be useful to better understand the nature of the stress–strain relations in the various



**Figure 2.2.** The basic “building blocks”

constitutive equations. New models can be developed easily by considering the combination of several elements. This introduction offers the opportunity to present in a simple framework most of the concepts used for tridimensional loading cases.

The mechanical response of the systems will be considered in three different planes, allowing us to illustrate the behavior for various loading types:

- *hardening*, or monotonic increase of the stress or the strain (strain–stress plane,  $\varepsilon$ – $\sigma$ );
- *creep*, a test performed under constant stress (time–strain plane,  $t$ – $\varepsilon$ );
- *relaxation*, a test performed under constant strain (time–stress plane,  $t$ – $\sigma$ ).

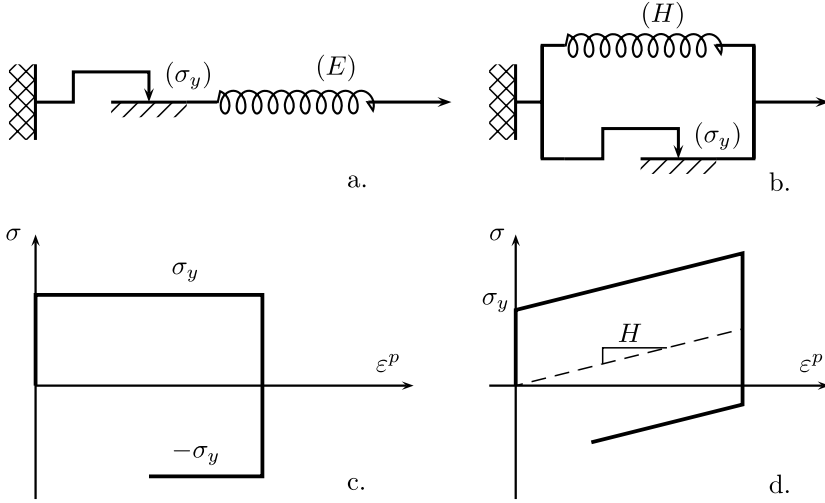
### 2.4.2. One-dimensional plasticity

#### *Elastic perfectly plastic models*

The system shown in Fig. 2.3a produces an elastic perfectly plastic behavior, as modeled in Fig. 2.3c. The maximum absolute value of the stress that can be applied to the system is  $\sigma_y$ .

The model is characterized by a *loading function*  $f$ , depending on the stress  $\sigma$  only, and defined by:

$$f(\sigma) = |\sigma| - \sigma_y \quad (2.50)$$



**Figure 2.3.** Systems built as series or parallel sets of friction device and spring

The elasticity domain corresponds to the negative values of  $f$ , and the behavior of the system can be summarized by the following equations:

- elastic behavior if:  $f < 0$   $(\dot{\epsilon} = \dot{\epsilon}^e = \dot{\sigma}/E)$
- elastic unloading if:  $f = 0$  and  $\dot{f} < 0$   $(\dot{\epsilon} = \dot{\epsilon}^e = \dot{\sigma}/E)$
- plastic flow if:  $f = 0$  and  $\dot{f} = 0$   $(\dot{\epsilon} = \dot{\epsilon}^p)$

For the elastic regime, the plastic strain rate is zero, meanwhile the elastic strain rate becomes zero during plastic flow. As a consequence, the expression of the plastic strain rate cannot take the stress as a critical variable. The total strain rate must be chosen as a driving variable.

The model does not involve any hardening, since the stress level no longer changes when stress reaches the limit of the elastic domain. There is no stored energy during the deformation, and the heat dissipation is equal to the plastic power. The model can reach infinite strain under constant load, leading to breakdown of the system by excessive strain.

### Prager model

When the elements are associated in parallel, as shown in Fig. 2.3b, the resulting behavior is illustrated in Fig. 2.3d. In such a case, a linear hardening is observed. The hardening is said to be “kinematic”, since it depends on the current value of the

plastic strain only. Under the present form, the model is rigid-plastic. It becomes elasto-plastic if a second spring is added, in series with respect to the first assembly. The shape of the curve in the plane  $\sigma$ - $\varepsilon^p$  is related to the stress level in the friction device,  $X$ . Its expression is obtained after writing that, during plastic flow, the stress in the spring is  $X = H\varepsilon^p$ . On the other hand, plastic flow appears only if the absolute value of the stress in the friction device,  $|\sigma - H\varepsilon^p|$ , is equal to  $\sigma_y$ . For a given strain, this stress  $X$  represents an *internal stress* that characterizes the new neutral state in the material.

This second example gives the opportunity to consider a more complete model. The loading function depends now on the applied stress and on the internal stress. It is written:

$$f(\sigma, X) = |\sigma - X| - \sigma_y \quad (2.51)$$

Plastic flow occurs if and only if  $f = 0$  and  $\dot{f} = 0$ . This leads to the following condition:

$$\frac{\partial f}{\partial \sigma} \dot{\sigma} + \frac{\partial f}{\partial X} \dot{X} = 0 \quad (2.52)$$

Thus:

$$\text{sign}(\sigma - X) \dot{\sigma} + \text{sign}(\sigma - X) \dot{X} = 0 \quad (2.53)$$

$$\dot{\sigma} = \dot{X}, \quad \text{and finally: } \dot{\varepsilon}^p = \dot{\sigma}/H \quad (2.54)$$

In this case, there is a stress evolution during plastic flow, such that it can be used as a control variable. One can also express the plastic strain rate as a function of the total strain rate, by using the strain partition assumption, combined with the expression of the plastic strain rate. In this framework, one can return to the perfectly plastic case by choosing  $H = 0$ :

$$\dot{\varepsilon}^p = \frac{E}{E + H} \dot{\varepsilon} \quad (2.55)$$

The amount of dissipated energy for one cycle is exactly the same as for the previous assembly. As a matter of fact, a fraction of the energy is now *temporarily* stored in the material (here in the spring), and *totally* recovered after unloading. This introduces the concept of reversible hardening, for those types of hardenings which do not involve any dissipation for a loading cycle. Most of the hardening rules developed later in this chapter will introduce an energy dissipation (see Sect. 3.7.2).

### General expression for one-dimensional elastoplasticity

In the general case, the “loading–unloading” conditions are the following;

- elastic behavior if:  $f(\sigma, A_i) < 0$   $(\dot{\varepsilon} = \dot{\sigma}/E)$
- elastic unloading if:  $f(\sigma, A_i) = 0$  and  $\dot{f}(\sigma, A_i) < 0$   $(\dot{\varepsilon} = \dot{\sigma}/E)$
- plastic flow if:  $f(\sigma, A_i) = 0$  and  $\dot{f}(\sigma, A_i) = 0$   $(\dot{\varepsilon} = \dot{\sigma}/E + \dot{\varepsilon}^p)$

The plastic modulus  $H$  generally depends on plastic strain and on the hardening variables. Its value for a given point defined by  $(\sigma, A_i)$  can be found by enforcing the condition that the current point remains at the border of the elastic domain during plastic flow. The resulting equation is the “consistency condition”:

$$\dot{f}(\sigma, A_i) = 0 \quad (2.56)$$

These expressions may appear too complex to define one-dimensional plastic flow (and they are!), nevertheless this framework is useful, since nothing but the same tools will be used for characterizing three-dimensional loadings. In the previously defined examples, the yield domain is either fixed or mobile, but for each case, its size is kept constant. No hardening variable is associated to the perfectly plastic case. Linear kinematic hardening involves a variable  $X$  which depends on the current value of the plastic strain. This variable will become tensorial in the 3D case. The corresponding type of hardening is called *kinematic hardening* (Fig. 2.4b). In the special case illustrated by the rheological model, the evolution of the variable  $X$  is linear with respect to the plastic strain, so that the model is called *linear kinematic* (Prager, 1958).

Instead of being translated and keeping a constant size, the elastic domain can have a constant location (classically, the origin of the deviatoric stress space), and a changing size. This other case describes a material which presents *isotropic hardening* (Taylor and Quinney, 1931). The hardening variable in  $f$  is the size of the elastic domain, denoted  $R$ :

$$f(\sigma, X, R) = |\sigma| - R - \sigma_y \quad (2.57)$$

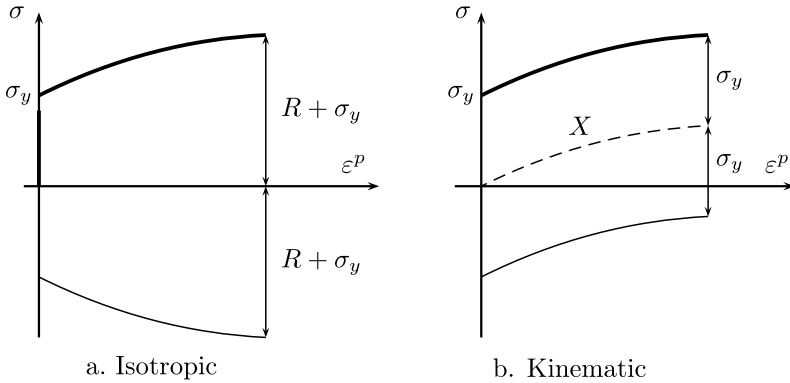
The evolution of this variable does not depend on the sign of the plastic strain rate. It will then be related to the *accumulated plastic strain rate*,  $p$ , whose rate is defined by the absolute value of the plastic strain rate:  $\dot{p} = |\dot{\varepsilon}^p|$ . There is no difference between  $p$  and  $\varepsilon^p$  for a monotonic loading. In such a case, checking the consistency condition is equivalent to enforcing the current value of the applied stress to be on the boundary of the yield domain. For kinematic hardening, one gets  $\sigma = X + \sigma_y$ , and for isotropic hardening,  $\sigma = R + \sigma_y$ . As a consequence, the shape of the tension curve is exactly determined by the shape of the equations governing the evolution of the hardening variables. The two models previously considered provide linear stress-strain curves after the initial yield, with either a null or a constant plastic modulus. Most of the time, one has to consider curves which have decreasing values as a function of the strain (for instance Ramberg–Osgood’s model, with two material parameters,  $K$  and  $m$ ), or an exponential function (this last formulation is often more satisfactory, since it introduces a saturating value, the *ultimate stress*  $\sigma_u$  on the material element; again there are two material parameters,  $\sigma_u$  and  $b$ ):

$$\sigma = \sigma_y + K(\varepsilon^p)^m \quad (2.58)$$

$$\sigma = \sigma_u + (\sigma_y - \sigma_u) \exp(-b \varepsilon^p) \quad (2.59)$$

Let us note that the plastic behavior is often defined by a series of values of the couple (plastic strain–stress), without any reference to an explicit form of the constitutive

equation. The hardening type considered in this case will be isotropic. This type of hardening is predominant for rather large strains (more than 10%). Nevertheless one has to keep in mind that kinematic hardening plays an important role during unloadings, even for large strains, and that it is predominant for small strains and cyclic loadings. This is for instance the only way to successfully represent the Bauschinger effect: the yield limit under a compressive (resp. tensile) loading applied after a tensile (resp. compressive) prehardening is smaller than the reference yield limit for a compression (resp. tension) loading. Isotropic hardening is nevertheless often promoted, since its numerical implementation is easier, compared to kinematic hardening.



**Figure 2.4.** Illustration of the two main hardening types

### 2.4.3. One-dimensional viscoelasticity

#### *A first rheological model*

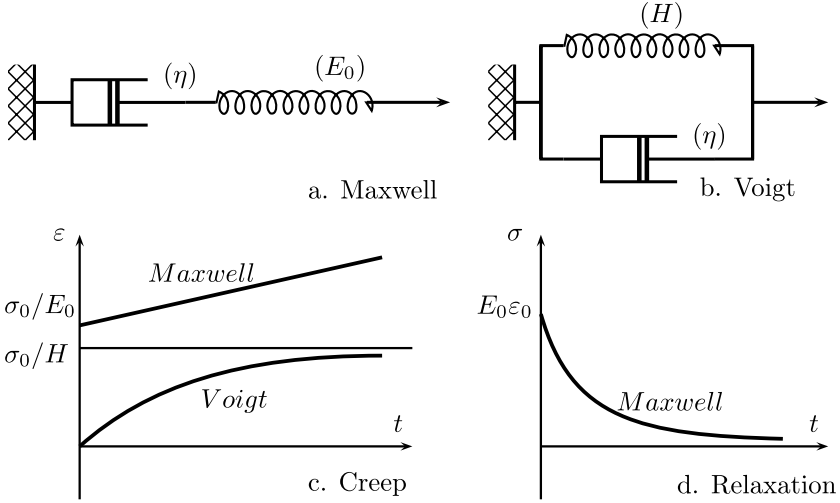
Maxwell's model is formed by gathering a dashpot and a spring as a serial assembly (Fig. 2.5a), meanwhile Voigt's model puts the same elements in parallel (Fig. 2.5b). The constitutive equations for each of these models are respectively:

$$\text{--Maxwell: } \dot{\varepsilon} = \dot{\sigma}/E_0 + \sigma/\eta \quad (2.60)$$

$$\text{--Voigt: } \sigma = H\varepsilon + \eta\dot{\varepsilon} \quad \text{or} \quad \dot{\varepsilon} = (\sigma - H\varepsilon)/\eta \quad (2.61)$$

Voigt's model is not able to produce any instantaneous response (no elasticity). Its relaxation function is then not continuous but piecewise derivable, with a finite jump at the loading time: when applied at  $t = 0$ , a strain jump will produce an infinite stress. The applicability of the model for modeling relaxation is then limited, but it can be used if the strain is applied progressively. In fact, most of the time, an additional spring is added, leading to Kelvin–Voigt's model (see next paragraph). Under a





**Figure 2.5.** Elementary behavior of Maxwell's and Voigt's models

constant stress  $\sigma_0$ , the strain tends to an asymptotic value,  $\sigma_0/H$ , the amount of creep is then limited (Fig. 2.5c). On the other hand, if the strain is slowly increased until  $\varepsilon_0$ , then fixed at this last value, the asymptotic stress will be  $H \varepsilon_0$ . The stress does not disappear completely.

On the contrary, for Maxwell's model, the creep rate is constant (secondary creep, Fig. 2.5c), and the stress tends to zero during a relaxation test (Fig. 2.5d).

When models and loadings are both so simple, the mechanical response can be obtained easily by integration of the differential equations. The expressions are then respectively, for Maxwell's model:

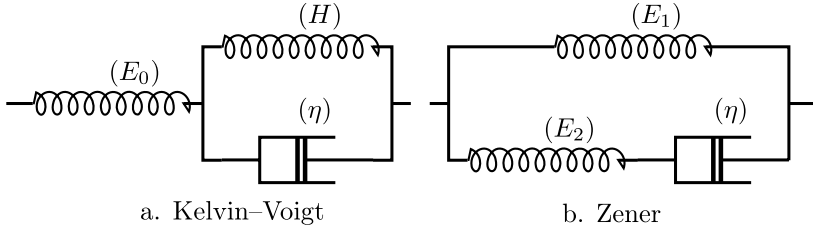
$$\text{—creep for a stress } \sigma_0: \quad \varepsilon = \sigma_0/E_0 + \sigma_0 t / \eta \quad (2.62)$$

$$\text{—relaxation for a strain } \varepsilon_0: \quad \sigma = E_0 \varepsilon_0 \exp[-t/\tau] \quad (2.63)$$

and for Voigt's model:

$$\text{—creep for a stress } \sigma_0: \quad \varepsilon = (\sigma_0 / H)(1 - \exp[-t/\tau']) \quad (2.64)$$

The unit of the constants  $\tau = \eta/E_0$  and  $\tau' = \eta/H$  is second;  $\tau$  stands for the so-called *Maxwell relaxation time*.



**Figure 2.6.** Example of a combined model

#### 2.4.4. Study of a combined model

As shown in Fig. 2.6a, Kelvin-Voigt's model provides the following mechanical responses for  $t > 0$ , under a creep loading at a stress  $\sigma_0$  (denoting by  $\tau_f$  the characteristic time  $\tau_f = \eta/H$ ), and under a relaxation loading at a strain  $\varepsilon_0$  (denoting by  $\tau_r$  the characteristic time  $\tau_r = \eta/(H + E_0)$ ):

$$\varepsilon(t) = C(t) \sigma_0 = \left( \frac{1}{E_0} + \frac{1}{H} (1 - \exp[-t/\tau_f]) \right) \sigma_0 \quad (2.65)$$

$$\sigma(t) = E(t) \varepsilon_0 = \left( \frac{H}{H + E_0} + \frac{E_0}{H + E_0} \exp[-t/\tau_r] \right) E_0 \varepsilon_0 \quad (2.66)$$

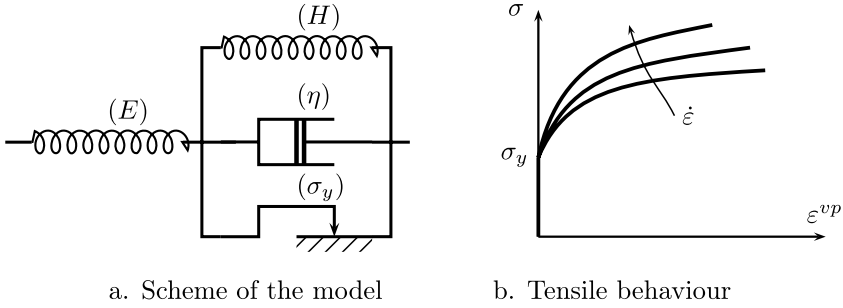
The relaxation time  $\tau_r$  is shorter than the corresponding time for creep loading,  $\tau_f$ . The material evolution toward an asymptotic state is then quicker in relaxation than in creep.

Zener's model (Fig. 2.6b) can be considered as another Kelvin-Voigt's model, after introducing the variable change  $1/E_1 = 1/E_0 + 1/H$ , and  $E_2 = E_0 + H$ . The two models are exactly equivalent under both creep and relaxation loading. This model is sometimes used for representing concrete behavior. The models previously described can be improved:

- a *generalized Kelvin-Voigt's* model is obtained by constructing a series assembly of several sets of dashpot-springs  $(H, \eta)$ ; this type of approach is used for network polymers;
- a *generalized Maxwell's* model is obtained by considering a parallel assembly of several sets of dashpot-springs  $(E_2, \eta)$ . This model is generally in good agreement with the behavior of thermoplastic polymers.

### 2.4.5. One-dimensional viscoplasticity

#### A simple rheological model



**Figure 2.7.** Generalized Bingham's model

In Fig. 2.7a, it is shown how a viscoplastic model can be derived from a plastic model, by adding a dashpot. The resulting model is the so-called generalized Bingham's model. The original Bingham's model involves neither the spring assembled in series ( $E \rightarrow \infty$ , no instantaneous elastic behavior, this is a rigid-viscoplastic model) nor the spring in parallel ( $H = 0$ , no hardening). The elastic strain is characterized by the spring ( $E$ ), the viscoplastic strain, denoted by  $\varepsilon^{vp}$ , is illustrated by the parallel assembly of the friction device and the dashpot. The equations of the model are obtained by combining all the elementary subsets:

$$X = H\varepsilon^{vp} \quad \sigma_v = \eta \dot{\varepsilon}^{vp} \quad \sigma_p \leq \sigma_y \quad (2.67)$$

where  $X$ ,  $\sigma_v$  and  $\sigma_p$  are respectively the stresses in the spring ( $H$ ), in the dashpot and in the friction device, and:

$$\sigma = X + \sigma_v + \sigma_p \quad (2.68)$$

An elastic domain is then present in this viscoplastic model. The border of the domain is reached when  $|\sigma_p| = \sigma_y$ . Three regimes can then be observed, according to the viscoplastic strain rate, that can be either zero, positive, or negative:

$$(a) \quad \dot{\varepsilon}^{vp} = 0 \quad |\sigma_p| = |\sigma - H\varepsilon^{vp}| \leq \sigma_y \quad (2.69)$$

$$(b) \quad \dot{\varepsilon}^{vp} > 0 \quad \sigma_p = \sigma - H\varepsilon^{vp} - \eta \dot{\varepsilon}^{vp} = \sigma_y \quad (2.70)$$

$$(c) \quad \dot{\varepsilon}^{vp} < 0 \quad \sigma_p = \sigma - H\varepsilon^{vp} - \eta \dot{\varepsilon}^{vp} = -\sigma_y \quad (2.71)$$

Case (a) corresponds to the interior of the elastic domain ( $|\sigma_p| < \sigma_y$ ) or to a state of elastic unloading ( $|\sigma_p| = \sigma_y$  and  $|\dot{\sigma}_p| \leq 0$ ). For the two other cases, a viscoplastic flow is present, with ( $|\sigma_p| = \sigma_y$  and  $|\dot{\sigma}_p| = 0$ ). Using the definition  $\langle x \rangle = \max(x, 0)$ , the three cases can be summarized by the unique expression:

$$\eta \dot{\varepsilon}^{vp} = \langle \sigma - X - \sigma_y \rangle \text{sign}(\sigma - X) \quad (2.72)$$

or:

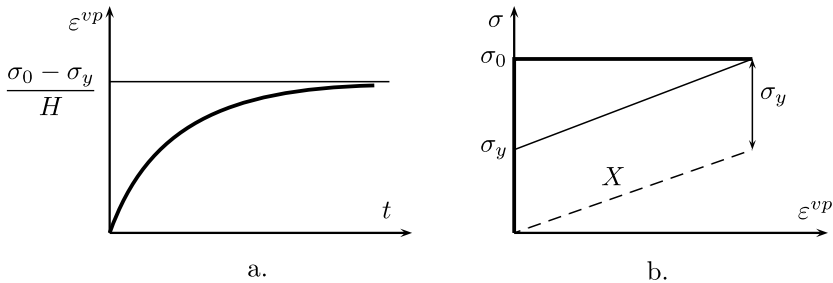
$$\dot{\varepsilon}^{vp} = \frac{\langle f \rangle}{\eta} \text{sign}(\sigma - X) \quad \text{with } f(\sigma, X) = |\sigma - X| - \sigma_y \quad (2.73)$$

If compared to the initial plastic model, the nature of the present one is totally different, since the current point representing the state of stress can be in the area defined by  $f > 0$ . On the other hand, the viscoplastic strain rate is time dependent. It can be non-zero without any stress increment, for instance. This new type of behavior explains why, in Fig. 2.7b, the tensile curve is no longer unique (the larger the strain rate, the higher the viscous stress  $\sigma_v$  and the higher the tension curve). This explains also why, during unloading at prescribed stress, the current stress point does not meet instantaneously the elastic domain (one can have a positive viscoplastic flow for decreasing stress values). Last but not least, this type of model allows the user to model creep tests and relaxation tests.

For a creep test (Fig. 2.8), assuming that stress is applied instantaneously at time  $t = 0$  (with a step between 0 and  $\sigma_o > \sigma_y$ ), starting from a reference state where all the strains are equal to zero, the model predicts that the viscoplastic strain is given by an exponential function of the time  $t$ , with a characteristic time  $\tau_f = \eta/H$  (Fig. 2.8a):

$$\varepsilon^{vp} = \frac{\sigma_o - \sigma_y}{H} \left( 1 - \exp\left(-\frac{t}{\tau_f}\right) \right) \quad (2.74)$$

Figure 2.8b shows, in the stress-viscoplastic strain plane, the evolutions of the internal stress  $X$  and of the viscoplastic threshold  $X + \sigma_y$ . When the threshold reaches the level of the applied stress  $\sigma_o$ , the creep strain rate becomes zero.



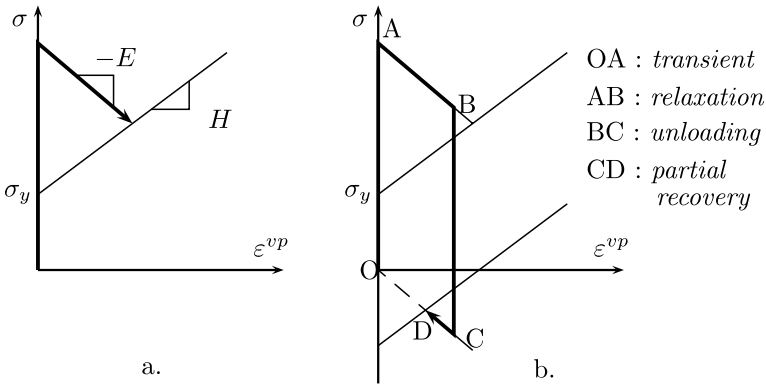
**Figure 2.8.** *Creep with a Bingham's model*

For a relaxation test, the mechanical response to a strain step (from 0 to  $\varepsilon_o$  such as  $E\varepsilon_o > \sigma_y$ ) involves now a characteristic relaxation time  $\tau_r = \eta/(E + H)$ :

$$\sigma = \sigma_y \frac{E}{E + H} \left( 1 - \exp\left(-\frac{t}{\tau_r}\right) \right) + \frac{E\varepsilon_o}{E + H} \left( H + E \exp\left(-\frac{t}{\tau_r}\right) \right) \quad (2.75)$$

Figure 2.9a shows the path followed by the current stress point during the relaxation (slope  $-E$ , since  $\dot{\varepsilon}^{vp} + \dot{\sigma}/E = 0$ ). On the other hand, Fig. 2.9b shows the path

followed during a *fading memory* test, when *recovery* is present. The test consists of the following steps: OA, load from 0 to  $\varepsilon_{max}$ ; AB, constant strain  $\varepsilon_{max}$ ; BC, unloading from  $\varepsilon_{max}$  to 0; CD, constant strain equal to zero. Depending on the initial loading level, the viscoplastic strain increment after unloading (branch BC of the loading path) can be either negative or equal to zero, nevertheless, the viscoplastic strain will never reach zero, unless the initial yield  $\sigma_y$  is zero itself. In fact, the asymptotic steady state is obtained as the intersection of the segment of slope  $-E$  joining C and O, and of the lower yield,  $-\sigma_y + H\varepsilon_p$ . If there is no initial threshold, the strain partition is no longer relevant: besides, the resulting model is Kelvin–Voigt’s model, a viscoelastic one, and there is no residual stress in the fading memory test.



**Figure 2.9.** Typical behavior of Bingham’s model under prescribed strain (relaxation, then fading memory test)

### A few classical viscoplastic models

In the previous example, the viscoplastic strain rate is proportional to a given “active” stress, obtained as the difference between the applied stress and the threshold. This active stress represents the distance between the current point and the border of the elastic domain. This is then nothing but the value taken by the function  $f$ . The linear relationship can be also replaced by a more general one, by introducing the viscosity function  $\phi$ , which provides then for a tensile test:

$$\dot{\varepsilon}^{vp} = \phi(f) \quad (2.76)$$

For a model presenting both kinematic and isotropic hardening, this relation can be inverted to get, for a tensile test:

$$\sigma = \sigma_y + X + R + \phi^{-1}(\dot{\varepsilon}^{vp}) = \sigma_y + X + R + \sigma_v \quad (2.77)$$

The tension curve is determined by the evolution of the threshold, as in the plastic case (through the  $X$  and  $R$  variables), but also by a viscosity function, which determines the value of the viscous stress  $\sigma_v$ . For obvious experimental reasons, it is assumed that  $\phi(0) = 0$ , and that  $\phi$  is a monotonic, always increasing function. For the limiting case where  $\sigma_v$  becomes zero, a time independent plastic model is recovered.

In a viscoplastic framework, one can have two ways to introduce hardening. As previously shown, the variables  $X$  and  $R$ , defined for plastic models, are still a good choice for a viscoplastic one. A second opportunity is to change the shape of the viscous stress. The models using the first approach are said to have an *additive hardening*, meanwhile those which use the second choice have a *product hardening*. A combined approach is also possible.

Finally, opposite to the plastic case, a model in which the elastic domain is reduced to the origin ( $\sigma = 0$ ), is admissible, even without hardening. In fact, this is nothing but the most simple model, Norton's model (with two material parameters,  $K$  and  $n$ ):

$$\dot{\varepsilon}^{vp} = \left( \frac{|\sigma|}{K} \right)^n \text{sign}(\sigma) \quad (2.78)$$

It can be generalized to produce a model with a threshold, but still without hardening, and also to produce more complex models, using again  $X$  and  $R$  (additive hardening, [LEM85b])

$$\dot{\varepsilon}^{vp} = \left\langle \frac{|\sigma| - \sigma_y}{K} \right\rangle^n \text{sign}(\sigma) \quad (2.79)$$

$$\dot{\varepsilon}^{vp} = \left\langle \frac{|\sigma - X| - R - \sigma_y}{K} \right\rangle^n \text{sign}(\sigma - X) \quad (2.80)$$

Many other solutions can be found in the literature, for instance, a hyperbolic sinus is used by Sellars and Teggart instead of a power function (no hardening, material parameters  $A$  and  $K$ ):

$$\dot{\varepsilon}^{vp} = A \sinh \left( \frac{|\sigma|}{K} \right) \text{sign}(\sigma) \quad (2.81)$$

Product type hardening is obtained when the function  $\phi$  does not depend only on  $f$ . For instance, Lemaitre's rule can be written with three material parameters,  $K$ ,  $m$  and  $n$  (all of them positive):

$$\dot{\varepsilon}^{vp} = \left( \frac{|\sigma|}{K} \right)^n p^{-n/m} \text{sign}(\sigma) \quad \text{with } \dot{p} = |\dot{\varepsilon}^{vp}| \quad (2.82)$$

### 2.4.6. Temperature influence

All the material parameters previously defined are likely to depend on temperature. The relationships can be defined by tables, for instance each time they are identified for a series of constant temperature tests. For those cases where the physical mechanisms are well defined, one can write precise explicit expressions for the temperature evolution. The most popular rule is Arrhenius'. This rule is valid for creep tests. It introduces a thermal activation energy,  $Q$ , and  $R$ , the constant of perfect gas (note that the ratio  $Q/R$  is homogeneous to a temperature), and assumes that, for a given load, the higher the temperature, the higher the viscoplastic strain rate:

$$\dot{\varepsilon}^{vp} = \dot{\varepsilon}_0 \exp(-Q/RT) f(\sigma) \quad (2.83)$$

Time–temperature equivalences can then be founded on this rule, so that researchers often use laboratory tests at temperatures higher than the target temperature in the applications, in order to collect in a short time a series of data valid for longer time periods. The approach must be taken with care. For materials presenting aging, new physical phenomena can appear, so that the extrapolation is no longer valid. Anyway, this approach cannot generally be extended to large temperature ranges.

### 2.4.7. Summary

The very general expressions presented in this section summarize the nature of the models for elasticity, plasticity, viscoplasticity. For the last two, an elastic domain has been defined (possibly reduced to the origin for the viscoplastic model) and hardening variables have been introduced. They have thus a lot of points in common. But they also have a big difference: plastic flow is time independent, viscoplastic flow is delayed:

$$d\varepsilon^p = g(\sigma, \dots) d\sigma \quad d\varepsilon^{vp} = g(\sigma, \dots) dt \quad (2.84)$$

This last remark will have important consequences for the expression of the elasto-(visco-)plastic tangent moduli.

For the moment, only very naive forms of the constitutive equations have been considered. More realistic expressions will be given later on, directly in the three-dimensional form.

## 2.5. Yield criteria

The one-dimensional expressions delivered in the preceding section for the inelastic models exhibit an elastic domain, which is an area in the space formed by the stress

and the hardening variables in which there is no plastic flow. For the one-dimensional case, this domain is restricted to a segment that can be transformed by translation or expansion (sometimes, this is only one point). On the other hand, several models are able to describe the presence of a maximum stress level, limiting the admissible state for a material element. The aim of the present section is then to extend these concepts for 3D loadings. It starts with a short review of the tools available in the framework of continuum mechanics, then the main classes of criteria are shown. Each material will have its relevant model built from the combination of a given flow rule, taken from the previous section, and of a criterion chosen in the following list.

### 2.5.1. Available tools

For the one-dimensional case, the elastic domain was characterized by two critical stress values, one in tension and one in compression, for which plastic flow may happen. For Prager's model, the initial yield is defined by the segment  $[-\sigma_y, \sigma_y]$ , and the current location for a strain  $\varepsilon^p$  is  $[-\sigma_y + X, \sigma_y + X]$ , with  $X = H\varepsilon^p$ . It is described by the loading function  $f$  (defined from  $\mathbb{R}^2$  into  $\mathbb{R}$ ,  $f : (\sigma, X) \rightarrow f(\sigma, X)$ ). Under multiaxial loadings,  $f$  will be a function of the stress tensor  $\underline{\sigma}$  and of the tensor  $\underline{X} = H\underline{\varepsilon}^p$ , (from  $\mathbb{R}^{12}$  in  $\mathbb{R}$ ), such as: if  $f(\underline{\sigma}, \underline{X}) < 0$ , the stress state is elastic, if  $f(\underline{\sigma}, \underline{X}) = 0$ , the current point is at the boundary of the domain, meanwhile the condition  $f(\underline{\sigma}, \underline{X}) > 0$  defines the outside of the domain. In the general case,  $f$  will operate on stresses and on tensorial or scalar hardening variables, thus a set  $(\underline{\sigma}, A_i)$ . The present section will not consider hardening: only the restrictions of the function  $f$  to the stress space is described.

Experiments show that, for most materials, the initial elastic domain is convex (this is specially true for metals and alloys whose deformation mechanisms are crystallographic slip). The loading function is then convex with respect to  $\underline{\sigma}$ , so that, for any real  $\lambda$  between 0 and 1, and for any couple of stresses  $(\underline{\sigma}_1, \underline{\sigma}_2)$  belonging to the boundary:

$$f(\lambda \underline{\sigma}_1 + (1 - \lambda) \underline{\sigma}_2) \leq \lambda f(\underline{\sigma}_1) + (1 - \lambda) f(\underline{\sigma}_2) \quad (2.85)$$

Material symmetries have to be respected by the plasticity criteria. In particular, for isotropic materials,  $f$  must be a symmetric function defined by means of eigenstresses, or a function of the stress invariants given by the characteristic polynomial:

$$I_1 = \text{Tr}(\underline{\sigma}) = \sigma_{ii} \quad (2.86)$$

$$I_2 = (1/2) \text{Tr}(\underline{\sigma}^2) = (1/2) \sigma_{ij} \sigma_{ji} \quad (2.87)$$

$$I_3 = (1/3) \text{Tr}(\underline{\sigma}^3) = (1/3) \sigma_{ij} \sigma_{jk} \sigma_{ki} \quad (2.88)$$

It is experimentally observed that plastic flow does not depend on hydrostatic pressure. As a consequence, the deviatoric stress and its invariants are chosen instead of



the stress tensor itself in the expressions of the plasticity criteria. The deviatoric stress tensor  $\underline{s}$  is obtained by removing the hydrostatic pressure from  $\underline{\sigma}$ .

$$\underline{s} = \underline{\sigma} - (I_1/3) \underline{I} \quad (2.89)$$

$$J_1 = \text{Tr}(\underline{s}) = 0 \quad (2.90)$$

$$J_2 = (1/2) \text{Tr}(\underline{s}^2) = (1/2) s_{ij} s_{ji} \quad (2.91)$$

$$J_3 = (1/3) \text{Tr}(\underline{s}^3) = (1/3) s_{ij} s_{jk} s_{ki} \quad (2.92)$$

It is often convenient, for identification purpose, to define criteria by means of expressions homogeneous to stresses. This leads to writing  $f$  as a function of  $J$  instead of  $J_2$ : as shown by (2.93),  $J$  is nothing but the absolute value of the applied stress for a tensile loading, and can be expressed with respect to the eigenstresses  $\sigma_1, \sigma_2, \sigma_3$  for 3D cases:

$$J = ((3/2) s_{ij} s_{ji})^{1/2} = ((1/2)((\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2))^{1/2} = |\sigma| \quad (2.93)$$

The preceding value is related to the so-called octahedral shear stress. The octahedral planes are those whose normal vector is  $\{1, 1, 1\}$  in the principal stress space. The normal and the tangential components of the stress vector for this plane can be expressed as a function of the three eigenstresses  $\sigma_1, \sigma_2, \sigma_3$ :

$$\sigma_{oct} = (1/3) I_1 \quad \tau_{oct} = (\sqrt{2}/3) J \quad (2.94)$$

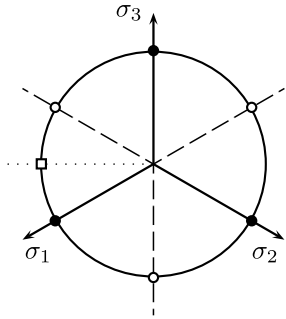
The value of  $J$  is then nothing but octahedral shear. The plane with a normal  $(1, 1, 1)$  will then play a specific role for the representation of the criteria. The reason is that all the stress states which differ only by a spherical tensor (they are equivalent with respect to a criterion which does not introduce hydrostatic pressure as a critical variable) have the same projection on this plane. This plane is shown in Fig. 2.10. Its equation is  $\sigma_1 + \sigma_2 + \sigma_3 = -I_1/3$ . The projections of the three principal axes determine sectors of  $2\pi/3$ .

## 2.5.2. Criteria without hydrostatic pressure

### *Von Mises criterion*

Since the trace of the stress tensor does not play an active role, a simple idea consists in using the second invariant of the deviatoric stress tensor, or  $J$ . This produces an ellipsoid in the space of the symmetric stress tensors (since  $J$  has a quadratic expression with respect to the components  $s_{ij}$ ). Denoting by  $\sigma_y$  the initial yield stress in tension, the expression is then:

$$f(\underline{\sigma}) = J - \sigma_y \quad (2.95)$$



- show the points equivalent to a simple tension, ○ show the points equivalent to a simple compression (for instance an equibiaxial loading, since the state  $\sigma_1 = \sigma_2 = \sigma$  is equivalent to  $\sigma_3 = -\sigma$ ), ◻ is a shear state

**Figure 2.10.** Definition of the deviatoric plane

### Tresca criterion

Von Mises criterion involves the maximum shear in each principal plane,  $(\sigma_i - \sigma_j)$ . On the contrary, Tresca criterion takes only the largest. As in the preceding case, hydrostatic pressure does not modify the value of the criterion. Instead of having a regular surface like von Mises, the Tresca criterion is only piecewise linear.

$$f(\underline{\sigma}) = \max_{i,j} |\sigma_i - \sigma_j| - \sigma_y \quad (2.96)$$

### Comparison of Tresca and von Mises criteria

Several subspaces are convenient to compare von Mises and Tresca criteria. The limit surfaces defined by each of them are drawn in given planes. The most popular choices are:

- the tension–shear plane (Fig. 2.11a), when the only non-zero components are  $\sigma = \sigma_{11}$  and  $\tau = \sigma_{12}$ ; the expressions of the criteria are then:

$$\text{– von Mises: } f(\sigma, \tau) = (\sigma^2 + 3\tau^2)^{1/2} - \sigma_y \quad (2.97)$$

$$\text{– Tresca: } f(\sigma, \tau) = (\sigma^2 + 4\tau^2)^{1/2} - \sigma_y \quad (2.98)$$

- the plane of two eigenstresses  $(\sigma_1, \sigma_2)$  (Fig. 2.11b), when the third one  $\sigma_3$  is zero:

$$\text{– von Mises: } f(\sigma_1, \sigma_2) = (\sigma_1^2 + \sigma_2^2 - \sigma_1\sigma_2)^{1/2} - \sigma_y \quad (2.99)$$

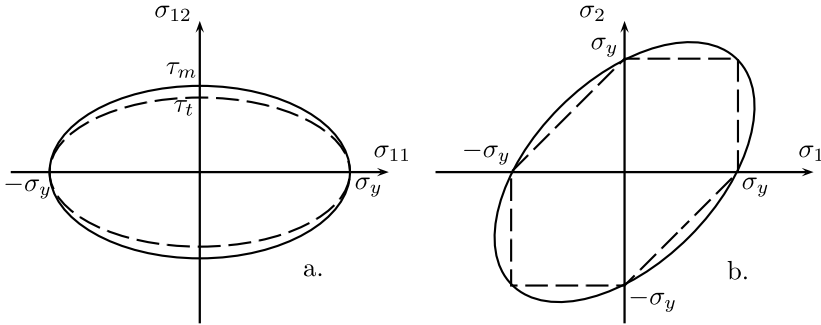
$$\text{– Tresca: } f(\sigma_1, \sigma_2) = \sigma_2 - \sigma_y \quad \text{if } 0 \leq \sigma_1 \leq \sigma_2 \quad (2.100)$$

$$f(\sigma_1, \sigma_2) = \sigma_1 - \sigma_y \quad \text{if } 0 \leq \sigma_2 \leq \sigma_1 \quad (2.101)$$

$$f(\sigma_1, \sigma_2) = \sigma_1 - \sigma_2 - \sigma_y \quad \text{if } \sigma_2 \leq 0 \leq \sigma_1 \quad (2.102)$$

$$\text{(symmetry wrt } \sigma_1 = \sigma_2 \text{ axis)} \quad (2.103)$$

- in the deviatoric plane (Fig. 2.10), the von Mises criterion is represented by a circle (this is consistent with the interpretation in terms of octahedral shear); Tresca's criterion is a hexagon;
- in the principal space, each criterion is represented by a cylinder whose axis is  $(1, 1, 1)$ , and the section is the curve defined in the deviatoric plane.



**Figure 2.11.** Comparison of Tresca (dashed lines) and von Mises (continuous lines) criteria (a) in tension–shear (von Mises:  $\tau_m = \sigma_y/\sqrt{3}$ , Tresca:  $\tau_t = \sigma_y/2$ ), (b) in biaxial tension

### 2.5.3. Criteria involving hydrostatic pressure

This type of criterion is needed to represent the behavior of powders, soils, and also for models that take damage into account. Classically, a compressive hydrostatic pressure makes plastic flow more difficult. One of the consequences of this formulation is the dissymmetry between tension and compression. This section shows a few examples, and Sect. 3.11 will complete the description for porous materials.

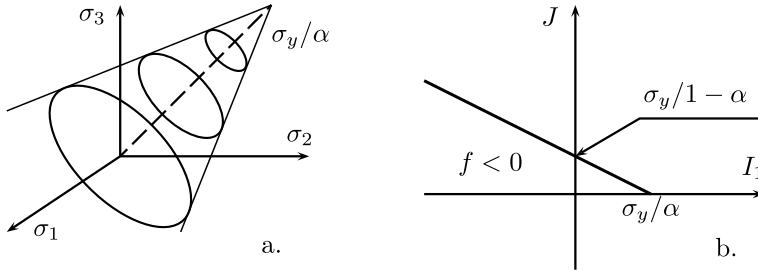
#### *Drucker–Prager criterion*

This is a generalization of the von Mises criterion, formed by means of a linear combination of the second invariant of the stress deviator and of the trace of the stress

tensor. It is still represented by a circle in the deviatoric plane, nevertheless, the radius of this circle depends on the location on the axis ( $\sigma_1 = \sigma_2 = \sigma_3$ ) (see Fig. 2.12a):

$$f(\boldsymbol{\sigma}) = (1 - \alpha)J + \alpha I_1 - \sigma_y \quad (2.104)$$

The yield stress under simple tension is still  $\sigma_y$ , and the yield stress in simple compression is  $-\sigma_y/(1 - 2\alpha)$ . The parameter  $\alpha$  depends on the material. It has to be between 0 and 1/2. For  $\alpha = 0$ , this model is equivalent to von Mises (Fig. 2.12b).



**Figure 2.12.** Drucker–Prager criterion, (a) in eigenstress space, (b) in the plane  $I_1 - J$

### Mohr–Coulomb criterion

The Mohr–Coulomb model is an extension of the Tresca criterion, involving, like the latter, the maximum shear stress, but also a pressure dependent stress, chosen as the centre of the Mohr circle corresponding to the maximum shear:

$$f(\boldsymbol{\sigma}) = \sigma_1 - \sigma_3 + (\sigma_1 + \sigma_3) \sin \phi - 2C \cos \phi \quad (\text{with } \sigma_3 \leq \sigma_2 \leq \sigma_1) \quad (2.105)$$

The idea behind the criterion is the presence of internal friction in the material, so that the ultimate shear that can be applied to the material ( $T_t$  in Fig. 2.13a) is larger and larger if the normal compression stress is high. The admissible limit forms an *intrinsic curve* in Mohr's plane. The expression (2.105) is obtained with the assumption of a linear friction coefficient:

$$|T_t| < -\tan(\phi) T_n + C \quad (2.106)$$

The material parameter  $C$  is called *cohesion*, and illustrates the shear stress which can be applied on the material under zero mean stress. The angle  $\phi$  denotes the *internal friction* in the material. If  $C$  is zero and not  $\phi$ , the material is pulverulent. If  $\phi$  is zero and not  $C$ , as for the Tresca criterion, the material is purely *cohesive*.

Another expression of the criterion can be written as a function of the *coefficient of passive earth pressure*,  $K_p$ , and of the compressive yield stress,  $R_p$ :

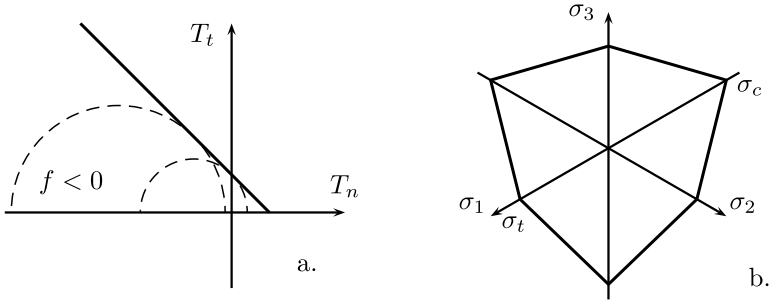
$$f(\underline{\sigma}) = K_p \sigma_1 - \sigma_3 - R_p \quad \text{with} \quad (2.107)$$

$$K_p = \frac{1 + \sin \phi}{1 - \sin \phi} \quad R_p = \frac{2 C \cos \phi}{1 - \sin \phi} \quad (2.108)$$

In the deviatoric plane (Fig. 2.13b), the criterion is represented by an irregular hexagon, characterized by the following values (with  $p = (-1/3)I_1$ ):

$$\sigma_t = 2 \sqrt{6}(C \cos \phi - p \sin \phi)/(3 + \sin \phi) \quad (2.109)$$

$$\sigma_c = 2 \sqrt{6}(-C \cos \phi + p \sin \phi)/(3 - \sin \phi) \quad (2.110)$$



**Figure 2.13.** Illustration of the Mohr–Coulomb criterion, (a) in the Mohr plane, (b) in the deviatoric plane

### “Closed” criteria

The two preceding criteria predict that the material cannot meet the yield limit under triaxial compression. This behavior is generally not true for real materials sensitive to hydrostatic pressure. More complex models have then been developed to correctly simulate stress states coming from operations like high isostatic pressure experiments for instance. The limit curve is decomposed into two parts, that are pasted for a critical negative value of the hydrostatic pressure. The most popular of these models is the *cap model*, which starts from Drucker–Prager, and introduces an elliptic “cap” for the negative values of the hydrostatic pressure, or the *Cam–clay* model (developed at Cambridge, used mainly for clays), whose limit curve is defined by two ellipses in the plane  $(I_1 - J)$ .

### 2.5.4. Anisotropic criteria

Literature shows that experimental yield surfaces of metallic alloys determined after plastic deformation exhibit an expansion, a translation and a distortion. The two first events are taken into account by isotropic and kinematic hardening. This is not the case of the last one, which is not described by classical models. Kinematic hardening and distortion induce anisotropy. On the other hand, materials like composites are fundamentally anisotropic, due to the local material geometry (e.g., long fibers composites). These two types of anisotropy are generally well represented by heterogeneous material models, (see Chap. 5), nevertheless they are still not too manageable in an industrial environment. This is why other solutions are investigated here to represent anisotropic materials. The most general approach consists in defining the criterion as a function of the components of the stress tensor in a given manifold. The chosen form must be intrinsic, in order to have a result independent of manifold change. A guide to build this type of model is provided by the invariant theory (see for instance [BOE78]).

The most popular criterion is a quadratic form, used to generalize the von Mises criterion, through the following expression:

$$J_B(\underline{\sigma}) = (\underline{\sigma} : \underline{\underline{B}} : \underline{\sigma})^{1/2} \quad (2.111)$$

where  $\underline{\underline{B}}$  is a fourth-order tensor. Taking the tensor  $\underline{\underline{J}}$  instead of a general  $\underline{\underline{B}}$  ( $\underline{\underline{J}}$  such as  $\underline{\underline{s}} = \underline{\underline{J}} : \underline{\underline{\sigma}}$  ( $\underline{\underline{s}}$  deviatoric stress tensor)) will lead to the classical von Mises criterion. The number of independent components of the tensor can be reduced according to the symmetries of the material. In addition to the classical symmetries  $B_{ijkl} = B_{ijlk} = B_{jikl} = B_{klij}$ , one has to enforce the condition  $B_{jjkl} = 0$  in order to ensure plastic incompressibility (the plastic strain rate is proportional to  $\underline{\underline{B}} : \underline{\underline{\sigma}}$ ). There are still 15 free material parameters (as for a symmetric  $5 \times 5$  matrix). If the material has three perpendicular planes of symmetry, the coupling terms between axial and shear components (such as  $B_{1112}$ ), are equal to zero, so that there are only six independent terms. In the symmetry frame, one finds the classical expression:

$$f(\underline{\sigma}) = (F(\sigma_{11} - \sigma_{22})^2 + G(\sigma_{22} - \sigma_{33})^2 + H(\sigma_{33} - \sigma_{11})^2 + 2L\sigma_{12}^2 + 2M\sigma_{23}^2 + 2N\sigma_{13}^2)^{1/2} - \sigma_y = f_H(\underline{\sigma}) \quad (2.112)$$

Writing the fourth-order tensor like a  $6 \times 6$  matrix, the terms in  $\underline{\underline{B}}$  are:

$$\begin{pmatrix} F+H & -F & -H & 0 & 0 & 0 \\ -F & G+F & -G & 0 & 0 & 0 \\ -H & -G & H+G & 0 & 0 & 0 \\ 0 & 0 & 0 & 2L & 0 & 0 \\ 0 & 0 & 0 & 0 & 2M & 0 \\ 0 & 0 & 0 & 0 & 0 & 2N \end{pmatrix} \quad (2.113)$$

A simple manipulation allows us to check that the same criterion can be expressed as a function of the components of the stress deviator associated to  $\underline{\sigma}$ ,  $J_B(\underline{\sigma}) = (\underline{s} : \underline{\underline{B}}' : \underline{s})^{1/2}$ , where the components of  $\underline{\underline{B}}'$  are:

$$\begin{pmatrix} 2F - G + 2H & 0 & 0 & 0 & 0 & 0 \\ 0 & 2F + 2G - H & 0 & 0 & 0 & 0 \\ 0 & 0 & -F + 2G + 2H & 0 & 0 & 0 \\ 0 & 0 & 0 & 2L & 0 & 0 \\ 0 & 0 & 0 & 0 & 2M & 0 \\ 0 & 0 & 0 & 0 & 0 & 2N \end{pmatrix} \quad (2.114)$$

For a transverse isotropy around the axis  $x_3$ , only three independent coefficients are left, since  $F = G$ ,  $L = M$ ,  $N = F + 2H$ . A full isotropy will induce also  $F = H$ ,  $L = N$ ,  $N = 3F$ , and leads to the tensor  $\underline{\underline{J}}$  and von Mises criterion, as already mentioned.

For the case of asymmetric yield in tension and compression, one may consider an expression involving a linear expression of the stress components, e.g., for Tsai criterion:

$$f(\underline{\sigma}) = f_H(\underline{\sigma}) + Q(\sigma_{22} - \sigma_{33}) + P(\sigma_{11} - \sigma_{33}) \quad (2.115)$$

The type of generalization shown for the case of criteria expressed in terms of invariants can also be applied to criteria expressed in terms of eigenstresses. A classical case in geotechnical applications deals with transversely isotropic materials: they have a criterion expressed in terms of eigenstresses, plus  $N$  and  $T$ , which are respectively the normal stress and the shear stress for a facet perpendicular to the axis defining slices schistosity, defined by the normed vector  $\underline{n}$ :

$$N = \underline{n} \cdot \underline{\underline{\sigma}} \cdot \underline{n} \quad T = (\|\underline{\underline{\sigma}} \cdot \underline{n}\|^2 - N^2)^{1/2} \quad (2.116)$$

For instance, Coulomb criterion is written:

$$f(\underline{\sigma}) = \max(K_p \max \sigma_i - \min \sigma_i - R_c, T + N \tan \phi' - C') \quad (2.117)$$

$$K_p = \frac{1 + \sin \phi}{1 - \sin \phi} \quad R_c = \frac{2C \cos \phi}{1 - \sin \phi} \quad (2.118)$$

where  $\phi$  denotes the friction angle in the slice planes,  $C$  the cohesion,  $\phi'$  the friction angle for the slip of two slices,  $C'$  the cohesion.

## 2.6. Numerical methods for nonlinear equations

We present here general methods to solve systems of nonlinear equations that can be written as:

$$\{R\}(\{U\}) = \{0\} \quad (2.119)$$

where  $\{R\}$  depends on  $\{U\}$ . The linear case is encountered when  $\{R\}(\{U\})$  can be written  $[K].\{U\} + \{A\}$ , where  $[K]$  does not depend on  $\{U\}$ . One should note that an equation written in the form  $\{R\}(\{U\}) = \{A\}$  can be rewritten in the general form of (2.119). In the framework of the finite element method, these methods will be used simultaneously to solve both the “global” problem and to integrate the constitutive equations (“local” problem).

### 2.6.1. Newton-type methods/modified Newton

Newton method is iterative and consists in linearizing the previous equation ( $k$ : iteration number):

$$\{R\}(\{U\}) = \{R\}(\{U\}_k) + \left. \frac{\partial \{R\}}{\partial \{U\}} \right|_{\{U\}=\{U\}_k} . (\{U\} - \{U\}_k) \quad (2.120)$$

We will note:

$$[K](\{U\}) = \frac{\partial \{R\}}{\partial \{U\}} \quad (2.121)$$

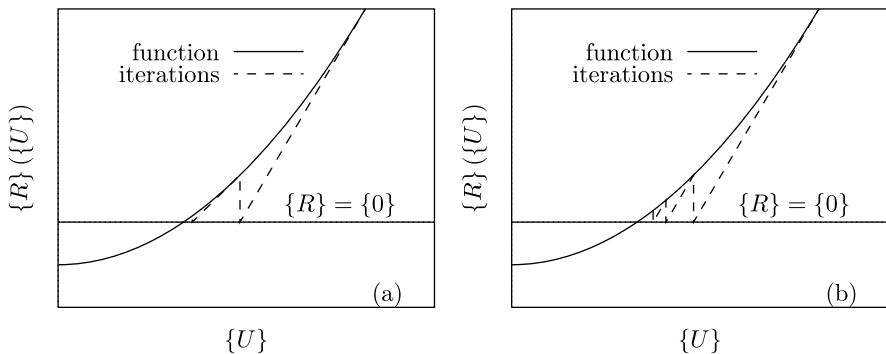
or

$$K_{ij}(\{U\}) = \frac{\partial R_i}{\partial U_j} \quad (2.122)$$

One gets after resolution of the linearized problem (Fig. 2.14a):

$$\{U\}_{k+1} = \{U\}_k - [K]_k^{-1} . \{R\}_k \quad (2.123)$$

In the case of systems with a large number of variables, the calculation cost of  $[K]^{-1}$  may become very high compared to the calculation cost of  $\{R\}(\{U\})$  and the



**Figure 2.14.** (a) Newton method, (b) Modified Newton method



product  $[K]^{-1} \cdot \{R\}_k$ . It sometimes is interesting to keep the inverse matrix obtained for the first iteration and to use it for the following iterations without recalculating the tangent matrix and its inverse (see Fig. 2.14b). In this case, the convergence is slower in terms of iterations number but can be faster in terms of calculation time. One get, then:

$$\{U\}_{k+1} = \{U\}_k - [K]_0^{-1} \cdot \{R\}_k \quad (2.124)$$

It is possible to use other variants of the modified-Newton method. Thus, by inverting the matrix  $[K]$  during the first two increments, and by keeping it constant afterwards, one gets:

$$\begin{aligned} \{U\}_1 &= \{U\}_0 - [K]_0^{-1} \cdot \{R\}_0 \\ \{U\}_2 &= \{U\}_1 - [K]_1^{-1} \cdot \{R\}_1 \\ &\vdots \\ \{U\}_{k+1} &= \{U\}_k - [K]_1^{-1} \cdot \{R\}_k \end{aligned} \quad (2.125)$$

If one has to calculate several increments for which a nonlinear system must be solved (as for the finite element method), one can keep the  $[K]^{-1}$  matrix that was calculated for the first increment. However, it is clear that this solution will not be valid in the case of contact, finite strain, etc.

## 2.6.2. One unknown case, order of convergence

### *Fixed point method*

The equation to solve here is:

$$f(x) = 0 \quad (2.126)$$

where  $x$  is scalar. A first method consists in writing this equation as:

$$x = g(x) \quad (2.127)$$

whose solution is called a *fixed point*. One gets the solution in an iterative way by choosing an initial value  $x_0$ :

$$x_{n+1} = g(x_n) \quad (2.128)$$

Let  $s$  be the solution of  $x = g(x)$ . If there exists an interval around  $s$  such that  $|g'| \leq K < 1$  then the series  $x_n$  converges towards  $s$ . To prove this, we notice first of all that there always exists  $t$  verifying  $t \in [x, s]$  such that (mean value theorem):

$$g(x) - g(s) = g'(t)(x - s) \quad (2.129)$$

As  $g(s) = s$  and  $x_n = g(x_{n-1})$ , one gets:

$$\begin{aligned} |x_n - s| &= |g(x_{n-1}) - g(s)| \leq |g'(t)| |x_{n-1} - s| \\ &\leq K |x_{n-1} - s| \\ &\leq \dots \leq K^n |x_0 - s| \end{aligned} \quad (2.130)$$

As  $K < 1$ ,  $\lim_{n \rightarrow \infty} |x_n - s| = 0$ .

*Iterative methods: order of convergence*

Let  $\epsilon_n$  be the error on  $x_n$

$$x_n = s + \epsilon_n \quad (2.131)$$

By taking the Taylor expansion of  $x_{n+1}$  one gets

$$\begin{aligned} x_{n+1} &= g(x_n) = g(s) + g'(s)(x_n - s) + \frac{1}{2}g''(s)(x_n - s)^2 \\ &= g(s) + g'(s)\epsilon_n + \frac{1}{2}g''(s)\epsilon_n^2 \end{aligned} \quad (2.132)$$

And then

$$x_{n+1} - g(s) = x_{n+1} - s = \epsilon_{n+1} = g'(s)\epsilon_n + \frac{1}{2}g''(s)\epsilon_n^2 \quad (2.133)$$

The order of an iterative method gives a measurement of its rate of convergence. At first-order it comes:

$$\epsilon_{n+1} \approx g'(s)\epsilon_n \quad (2.134)$$

and at second-order:

$$\epsilon_{n+1} \approx \frac{1}{2}g''(s)\epsilon_n^2 \quad (2.135)$$

*Application to Newton method*

For the Newton method, the Taylor expansion of  $x_n$  about 0 is used to find  $x_{n+1}$ :

$$f(x_{n+1}) = f(x_n) + (x_{n+1} - x_n) f'(x_n) = 0 \quad (2.136)$$

Then:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad (2.137)$$

This is equivalent to the fixed point method with:

$$g(x) = x - \frac{f(x)}{f'(x)} \quad (2.138)$$

One gets

$$g'(x) = \frac{f(x)f''(x)}{f'(x)^2} \quad (2.139)$$

and

$$g''(x) = \frac{f''(x)}{f'(x)} - 2 \frac{f(x)f''(x)^2}{f'(x)^3} + \frac{f(x)f'''(x)}{f'(x)^2} \quad (2.140)$$

One notes that:

$$g'(s) = 0 \quad (2.141)$$

The Newton method is then of the second-order. Moreover, there exists an interval around  $s$  such that  $|g'(s)| < 1$ . Newton method always converges for an initial value  $x_0$  close enough to the solution.

### *Application to the modified Newton method*

In the case of the modified Newton method, one rewrites (2.136) as:

$$f(x_{n+1}) = f(x_n) + (x_{n+1} - x_n) K = 0 \quad (2.142)$$

where  $K$  is a constant. We then have:

$$x_{n+1} = x_n - \frac{f(x_n)}{K} \quad (2.143)$$

Let

$$g(x) = x - \frac{f(x)}{K} \quad (2.144)$$

and

$$g'(x) = 1 - \frac{f'(x)}{K} \quad (2.145)$$

As  $g'(s) \neq 0$ , this method is of first-order. It converges for  $K$  such that:

$$-1 < 1 - \frac{f'(s)}{K} < 1 \quad (2.146)$$

$K$  being a derivative evaluated at iteration  $i$ , let  $K = f'(x_i)$ ,  $K$  and  $f'(s)$  must then be of the same sign and  $f'(s) < 2K$ .

### **2.6.3. BFGS method (Broyden–Fletcher–Goldfarb–Shanno)**

An alternative to Newton method [BAT82, MAT79, PRE88] consists in updating the inverse of  $[K]$  in order to obtain, at iteration  $k$ , a secant approximation of the matrix at iteration  $k - 1$ . If one defines the increment of the unknown  $\{U\}$  as:

$$\{\delta\}_k = \{U\}_k - \{U\}_{k-1} \quad (2.147)$$

and the increment of the deviation from the solution:

$$\{\gamma\}_k = \{R\}_k - \{R\}_{k-1} \quad (2.148)$$

the updated matrix  $[K]_k$  must verify:

$$\{\gamma\}_k = [K]_k \cdot \{\delta\}_k \quad (2.149)$$

BFGS method consists of three steps:

1. Evaluation of an increment of vector  $\{U\}$ :

$$\Delta\{\bar{U}\} = -[K]_{k-1}^{-1} \cdot \{R\}_{k-1} \quad (2.150)$$

$\Delta\{\bar{U}\}$  gives the direction of the actual increment.

2. Seeking a solution in direction  $\Delta\{\bar{U}\}$  in the form:

$$\{U\}_k = \{U\}_{k-1} + \beta \Delta\{\bar{U}\} \quad (2.151)$$

where  $\beta$  is a scalar.  $\beta$  is calculated to minimize the scalar product:

$$-\Delta\{\bar{U}\} \cdot \{R\}_k \quad (2.152)$$

However, one can skip this step and consider  $\beta = 1$ .  $\{\delta\}_k$  and  $\{\gamma\}_k$  can then be calculated.

3. Evaluation of the corrected matrix  $[K]_k^{-1}$ . In BFGS method, the updated matrix has to be written as:

$$[K]_k^{-1} = [A]_k^T \cdot [K]_{k-1}^{-1} \cdot [A]_k \quad (2.153)$$

where  $[A]$  is a matrix expressed as:

$$[A]_k = [I] + \{v\}_k \otimes \{w\}_k \quad (2.154)$$

where  $\{v\}_k$  and  $\{w\}_k$  are vectors calculated from known quantities:

$$\{v\}_k = - \left[ \frac{\{\delta\}_k \cdot \{\gamma\}_k}{\{\delta\}_k \cdot [K]_{k-1}^{-1} \cdot \{\delta\}_k} \right]^{1/2} [K]_{k-1}^{-1} \cdot \{\delta\}_k - \{\gamma\}_k \quad (2.155)$$

$$\{w\}_k = \frac{1}{\{\delta\}_k \cdot \{\gamma\}_k} \{\delta\}_k \quad (2.156)$$

Direction  $\Delta\{\bar{U}\}$  is obtained by the product:

$$\begin{aligned} \Delta\{\bar{U}\} = & -([I] + \{w\}_{k-1} \otimes \{v\}_{k-1}) \cdots ([I] + \{w\}_1 \otimes \{v\}_1) \cdot [K]_0^{-1} \\ & \cdot ([I] + \{v\}_1 \otimes \{w\}_1) \cdots ([I] + \{v\}_{k-1} \otimes \{w\}_{k-1}) \cdot \{R\}_{k-1} \end{aligned} \quad (2.157)$$

It is not necessary to calculate  $[K]_k^{-1}$ . It is sufficient to multiply  $\{R\}_{k-1}$  by  $[I] + \{v\}_{k-1} \otimes \{w\}_{k-1} \cdots [I] + \{v\}_1 \otimes \{w\}_1$ , then the result by  $[K]_0^{-1}$  and to multiply in inverse order  $[I] + \{w\}_1 \otimes \{v\}_1 \cdots [I] + \{w\}_{k-1} \otimes \{v\}_{k-1}$  to get  $\Delta\{\bar{U}\}$ .

#### 2.6.4. Iterative method—conjugate gradient

This method [LUE84] was developed to solve systems of the form:

$$V(\{U\}) = \frac{1}{2} \{U\} \cdot [K] \cdot \{U\} - \{U\} \cdot \{F\} \quad (2.158)$$

The derivative of the previous equation is:

$$\frac{\partial V}{\partial \{U\}} = [K] \cdot \{U\} - \{F\} = \{r\} \quad (2.159)$$

where  $\{r\}$  is the residual or “gradient” of (2.158). The method uses the following approximation of the solution:

$$\begin{aligned} \{d\}_i &= \{r\}_i + \beta_i \{d\}_{i-1} \\ \{U\}_{i+1} &= \{U\}_i + \alpha_i \{d\}_i = \{U\}_i + \alpha_i \{r\}_i + \alpha_i \beta_i \{d\}_{i-1} \end{aligned} \quad (2.160)$$

where  $\{d\}_i$  is the direction of research. One searches for  $\alpha_i$  and  $\beta_i$  in order to minimize  $V(\{u\})$ , so that:

$$\frac{\partial V}{\partial \alpha_i} = 0 \quad \frac{\partial V}{\partial \beta_i} = 0 \quad (2.161)$$

with:

$$\frac{\partial V}{\partial \alpha_i} = \frac{\partial V}{\partial \{U\}_{i+1}} \cdot \frac{\partial \{U\}_{i+1}}{\partial \alpha_i} = ([K] \cdot \{U\}_{i+1} - \{F\}) \cdot \{d\}_i = \{r\}_{i+1} \cdot \{d\}_i = 0 \quad (2.162)$$

and

$$\frac{\partial V}{\partial \beta_i} = \frac{\partial V}{\partial \{U\}_{i+1}} \cdot \frac{\partial \{U\}_{i+1}}{\partial \beta_i} = \{r\}_{i+1} \cdot (\alpha_i \{d\}_{i-1}) = \alpha_i \{r\}_{i+1} \cdot \{d\}_{i-1} = 0 \quad (2.163)$$

The problem is equivalent to finding  $\alpha_i$  and  $\beta_i$  such that

$$\{r\}_{i+1} \cdot \{d\}_i = 0 \quad (2.164)$$

$$\{r\}_{i+1} \cdot \{d\}_{i-1} = 0 \quad (2.165)$$

with

$$\begin{aligned} \{r\}_{i+1} &= -\{F\} + [K] \cdot \{U\}_{i+1} \\ &= -\{F\} + [K] \cdot (\{U\}_i + \alpha_i \{r\}_i + \alpha_i \beta_i \{d\}_{i-1}) \\ &= -\{F\} + [K] \cdot \{U\}_i + \alpha_i [K] \cdot (\{r\}_i + \beta_i \{d\}_{i-1}) \\ &= \{r\}_i + \alpha_i [K] \cdot (\{r\}_i + \beta_i \{d\}_{i-1}) \\ &= \{r\}_i + \alpha_i [K] \cdot \{d\}_i \end{aligned} \quad (2.166)$$

Noting that, by construction,  $\{r\}_i \cdot \{d\}_{i-1} = 0$ , one gets

$$\{r\}_{i+1} \cdot \{d\}_{i-1} = \alpha_i \{d\}_{i-1} \cdot [K] \cdot (\{r\}_i + \beta_i \{d\}_{i-1}) = 0 \quad (2.167)$$

Then

$$\beta_i = - \frac{\{d\}_{i-1} \cdot [K] \cdot \{r\}_i}{\{d\}_{i-1} \cdot [K] \cdot \{d\}_{i-1}} \quad (2.168)$$

$\beta_i$  corresponds to the coefficient of optimal descent. From  $\beta_i$ , it is possible to calculate  $\{d\}_i$ . Using relations  $\{r\}_{i+1} = \{r\}_i + \alpha_i [K] \cdot \{d\}_i$  and  $\{r\}_{i+1} \cdot \{d\}_i = 0$ , one gets:

$$\alpha_i = - \frac{\{r\}_i \cdot \{d\}_i}{\{d\}_i \cdot [K] \cdot \{d\}_i} \quad (2.169)$$

This corresponds to an orthogonalization, in the sense of the scalar product associated to  $[K]$ , of the new gradient with respect to the previous direction. The main advantage of the conjugate gradient method is that it is not necessary to calculate the matrix  $[K]^{-1}$ . The main computational effort lies in the matrix–vector product. The other phases of the calculation are mere scalar products and linear combinations of vectors. In a finite element context, it is not necessary to store the global matrix  $[K]$  but it is then necessary to recalculate the elementary matrices for each iteration.

### 2.6.5. *Riks method*

This method [RIK79], because of its formulation, applies particularly well to mechanics.

#### *Control of the problem*

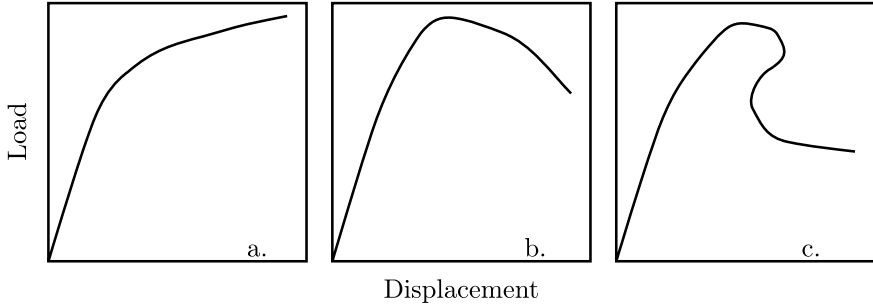
The most natural control of the finite element problem consists in applying external loads ( $\{F_e\}$ ). The system to solve is then:

$$\{F_i\} (\{U\}) = \{F_e\} \quad (2.170)$$

This kind of control is suitable for the case where load is a monotonic function of displacement.

Should a maximum load (plasticity) or a structure instability (plate buckling) exist, load ceases to be monotonic and control is achieved imposing displacement where the load is applied. This boils down to decreasing the number of degrees of freedom of the system. Reactions corresponding to these degrees of freedom are then free.

In the case where there exists a snap-back point on the load–displacement (e.g., buckling), one must exert a mixed (load and displacement) control. Figure 2.15 sums up the three situations.



**Figure 2.15.** Type of response load–displacement: (a) Increasing load. Single  $U$ – $F$  couple in all cases. (b) Load–displacement curve with maximum loading. Several solutions in displacement for a given load. (c) Load–displacement curve with snap-back in displacement. Several solutions in load for a given displacement

### Control of the load–displacement response

The equilibrium equation is written:

$$\lambda \{\mathcal{F}_e\} = \{F_i\} (\{U\}) \quad (2.171)$$

$\{\mathcal{F}_e\}$  is the reference loads vector and is known.  $\lambda$  is the scalar load parameter. From now on, there is an additional unknown. The path  $(\{U\}, \lambda)$  can be parameterized by the arc length  $s$  such that

$$\lambda = \lambda(s) \quad \{U\} = \{U\}(s) \quad (2.172)$$

The unitary vector tangent to the curve writes:

$$\{t\} = \left\{ \begin{array}{c} \frac{1}{m} \frac{d\{U\}}{ds} \\ \frac{1}{m} \frac{d\lambda}{ds} \end{array} \right\} \quad \text{with } m = \sqrt{\frac{d\{U\}}{ds} \cdot \frac{d\{U\}}{ds} + \frac{d\lambda}{ds} \frac{d\lambda}{ds}} \quad (2.173)$$

The derivative of the equilibrium equation (2.171) with respect to  $s$  leads to:

$$\frac{d\lambda}{ds} \{\mathcal{F}_e\} - [K] \cdot \frac{d\{U\}}{ds} = \{0\} \quad (2.174)$$

with  $[K] = d\{F_i\}/d\{U\}$ . We will write

$$\{U_F\} = [K^{-1}] \cdot \{\mathcal{F}_e\} \quad (2.175)$$

Then, one gets:

$$\frac{d\{U\}}{ds} = \frac{d\lambda}{ds} \{U_F\} \quad (2.176)$$

*Imposed arc length*

The mixed-control technique with imposed arc length consists in realizing increments in  $(\{U\}, \lambda)$  space such that the increment has a given length  $\Delta s$ :

$$\Delta\{U\} \cdot \Delta\{U\} + (\Delta\lambda)^2 = (\Delta s)^2 \quad (2.177)$$

by writing

$$\Delta\{U\} \approx \Delta s \frac{d\lambda}{ds} \{U_F\} \quad (2.178)$$

$$\Delta\lambda \approx \Delta s \frac{d\lambda}{ds} \quad (2.179)$$

the previous equation becomes:

$$\left(\frac{d\lambda}{ds}\right)^2 \{U_F\} \cdot \{U_F\} + \left(\frac{d\lambda}{ds}\right)^2 = 1 \quad (2.180)$$

Then, one gets for the first iteration:

$$\Delta\lambda^1 = \frac{\pm \Delta s}{\sqrt{1 + \{U_F\} \cdot \{U_F\}}} \quad (2.181)$$

$$\Delta\{U\}^1 = \Delta\lambda^1 \{U_F\} \quad (2.182)$$

The sign of  $\Delta\lambda^1$  is chosen to maximize the scalar product

$$\Delta\lambda_p \Delta\lambda^1 + \Delta\{U\}_p \cdot \Delta\{U\}^1 \quad (2.183)$$

where  $(\Delta\lambda_p, \Delta\{U\}_p)$  are the solution increments of the previous step. We assume that there is a certain continuity of the displacement direction from one increment to the other.

For the following iterations, we note that

$$\delta\{U_R\}^i = [K]^{-1} \cdot (\lambda^i \{F_e\} - \{F_i\}^i) \quad (2.184)$$

$\delta\{U_R\}^i$  corresponds to the displacement increment calculated for a simple load control. The iteration solution is written

$$\Delta\{U\}^{i+1} = \Delta\{U\}^i + \delta\{U_R\}^i + \delta\lambda\{U_F\} \quad (2.185)$$

with the condition:

$$\Delta\{U\}^{i+1} \cdot \Delta\{U\}^{i+1} + (\Delta\lambda^{i+1})^2 = (\Delta s)^2 \quad (2.186)$$



$\delta\lambda\{U_F\}$  corresponds to a change in the applied load. Furthermore, we have

$$\Delta\lambda^{i+1} = \Delta\lambda^i + \delta\lambda \quad (2.187)$$

We get then a quadratic equation in  $\delta\lambda$  (taking into account that  $\Delta\{U\}^i \cdot \Delta\{U\}^i + (\Delta\lambda^i)^2 = (\Delta s)^2$ ):

$$\begin{aligned} a(\delta\lambda)^2 + b\delta\lambda + c &= 0 \\ \text{with } a &= 1 + \{U_F\} \cdot \{U_F\} \\ b &= 2(\{v\}^i \cdot \{U_F\} + \Delta\lambda^i) \\ c &= 2\{U\}^i \cdot \delta\{U_R\}^i + \delta\{U_R\}^i \cdot \delta\{U_R\}^i \\ \text{with } \{v\}^i &= \{U\}^i + \delta\{U_R\}^i \end{aligned} \quad (2.188)$$

There are then two roots. One chooses the root that allows minimizing the angle between two successive evaluations of  $\Delta\{U\}$ . Then, we maximize:

$$\cos\theta^i = \frac{\{\Delta\{U\}^i, \Delta\lambda^i\} \cdot \{\Delta\{U\}^{i+1}, \Delta\lambda^{i+1}\}}{(\Delta s)^2} \quad (2.189)$$

### 2.6.6. Convergence

Various ways exist to test the convergence of the iterative resolution.

- We try to solve  $\{R\} = \{0\}$ . The resolution process is stopped when  $\{R\}$  is small enough.

$$\|\{R\}\|_n < R_\epsilon \quad (2.190)$$

where  $R_\epsilon$  is the required precision, with:

$$\|\{R\}\|_n = \left( \sum_i R_i^n \right)^{1/n} \quad (2.191)$$

The norm  $\|\cdot\|_\infty$  is often used,

$$\|\{R\}\|_\infty = \max_i |R_i| \quad (2.192)$$

- In numerous cases, equation  $\{R\} = \{0\}$  can be written  $\{R\}_i - \{R\}_e = \{0\}$  where  $\{R\}_e$  is imposed (external load). A relative error can then be defined:

$$\frac{\|\{R\}_i - \{R\}_e\|}{\|\{R\}_e\|} < r_\epsilon \quad (2.193)$$

where  $r_\epsilon$  is the required relative precision. In mechanics  $\{R\}_i$  corresponds to internal forces and  $\{R\}_e$  to external forces. In some cases (for example thermal cooling)  $\{R\}_e = \{0\}$ ; it is then impossible to define a relative error and one must use the absolute error  $\|\{R\}_i - \{R\}_e\| < R_\epsilon$ .

- The search for a solution can be stopped when  $\{U\}$  becomes stable; i.e.:

$$\|\{U\}_{k+1} - \{U\}_k\|_n < U_\epsilon \quad (2.194)$$

*Note:* In some cases, vectors  $\{R\}$  and  $\{U\}$  are not homogeneous. For example in the coupled thermo-mechanical case,  $\{U\}$  contains displacements and temperatures, and  $\{R\}$  forces and thermal flows. One must then normalize with predetermined quantities the various components of  $\{R\}$  and  $\{U\}$  before testing the convergence.

## 2.7. Numerical solution of differential equations

### 2.7.1. General overview

In this section first-order differential equations (DE) are dealt with. Higher order DE can generally be rewritten as first-order ones. For example:

$$\frac{d^2y}{dt^2} + q(t)\frac{dy}{dt} = r(t) \quad (2.195)$$

can be rewritten:

$$\frac{dy}{dt} = z(t) \quad (2.196)$$

$$\frac{dz}{dt} = r(t) - q(t)z(t) \quad (2.197)$$

where  $z$  is a new variable.

In a more general way, a first-order DE can be written:

$$\dot{v}_i = f_i(t, v_1, \dots, v_n) \quad i = 1 \dots n \quad (2.198)$$

or

$$\{\dot{v}\} = \{f\}(t, \{v\}) \quad (2.199)$$

where  $\{v\}$  is a vector of size  $n$ . It is necessary to add to these equations initial conditions written as:

$$\{v\}(t = t_0) = \{v\}_0 \quad (2.200)$$

In this section, we introduce two integration methods: (a) explicit Runge–Kutta type methods [PRE88, TOU93], (b) implicit middle point methods (or  $\theta$ -method). We want to integrate over a time interval  $\Delta t$  from initial time  $t$ . The easiest method is Euler's method which is entirely explicit and very easy to formulate as it relies only on the vector  $\{v\}$  given at initial time  $t$ :

$$\{v\}(t + \Delta t) = \{v\}(t) + \Delta t \{\dot{v}\}(t, \{v\}) \quad (2.201)$$

However, this method is neither precise nor stable.

### 2.7.2. Runge–Kutta method

#### Basic method

By using a Taylor series, one gets:

$$\{v\}(t + \Delta t) = \{v\}(t) + \{\dot{v}\}(t)\Delta t + O(\Delta t^2) \quad (2.202)$$

Euler's integration has a second-order precision  $O(\Delta t^2)$ . One can however test the middle point  $(t + \Delta t/2)$ . Let

$$\{\delta v_1\} = \Delta t \{\dot{v}\}(t) \quad (2.203)$$

and

$$\begin{aligned} \{\delta v_2\} &= \Delta t \{\dot{v}\} \left( t + \frac{\Delta t}{2}, \{v\}(t) + \frac{1}{2}\{\delta v_1\} \right) \\ &= \Delta t \left( \{\dot{v}\}(t) + \frac{\Delta t}{2}\{\ddot{v}\}(t) \right) \\ &= \{\delta v_1\} + \frac{\Delta t^2}{2}\{\ddot{v}\}(t) \end{aligned} \quad (2.204)$$

The second-order Taylor series gives:

$$\{v\}(t + \Delta t) = \{v\}(t) + \{\dot{v}\}(t)\Delta t + \{\ddot{v}\}(t)\frac{\Delta t^2}{2} + O(\Delta t^3) \quad (2.205)$$

One can eliminate the term  $\{\ddot{v}\}(t)\Delta t^2/2$  which is equal to  $\{\delta v_2\} - \{\delta v_1\}$ . It follows that:

$$\{v\}(t + \Delta t) = \{v\}(t) + \{\delta v_2\} + O(\Delta t^3) \quad (2.206)$$

This formula shows that the first-order term in  $\Delta t$  has been eliminated: one gets, then, a second-order method. This procedure can be generalized. The fourth-order Runge–Kutta method writes:

$$\begin{aligned} \{\delta v_1\} &= \Delta t \{\dot{v}\}(t, \{v\}) \\ \{\delta v_2\} &= \Delta t \{\dot{v}\} \left( t + \frac{\Delta t}{2}, \{v\} + \frac{1}{2}\{\delta v_1\} \right) \\ \{\delta v_3\} &= \Delta t \{\dot{v}\} \left( t + \frac{\Delta t}{2}, \{v\} + \frac{1}{2}\{\delta v_2\} \right) \\ \{\delta v_4\} &= \Delta t \{\dot{v}\}(t + \Delta t, \{v\} + \{\delta v_3\}) \\ \{v\}(t + \Delta t) &= \{v\}(t) + \frac{1}{6}\{\delta v_1\} + \frac{1}{3}\{\delta v_2\} + \frac{1}{3}\{\delta v_3\} + \frac{1}{6}\{\delta v_4\} + O(\Delta t^5) \end{aligned} \quad (2.207)$$

This method requires then four evaluations of the function  $\{\dot{v}\}$  whereas the second-order method requires only two. It will then be more efficient if it allows a time step twice as large for the same precision. However high order does not always mean high precision. This is why it is appropriate to use an adaptive time step method.

*Adaptive time step method*

The aim of such a method is to get a predetermined precision while minimizing the calculation time. The method is designed to use a large time step where the function  $\{v\}$  varies little and to reduce it when  $\{v\}$  varies a lot.

Let  $\Delta t$  be the time increment over which the integration has to be made. It is divided into  $n$   $\delta t_k$  increments so that:  $\Delta t = \sum_{k=1}^{k=n} \delta t_k$ . In the case of the fourth-order method, the technique consists in using one time step of magnitude  $\delta t$  and two time steps of magnitude  $\delta t/2$ . Thus, one needs 11 evaluations of  $\{v\}$  (both integrations share the same starting point) that must be compared to the eight evaluations needed to integrate using only two time steps  $\delta t/2$  (as one gets the same precision). Extra cost is then  $11/8 = 1.375$ . Let  $\{v\}(t + \delta t)$  be the exact solution, then:

$$\{v\}(t + \delta t) = \{v_1\} + (\delta t)^5 \{\phi\} + O(\delta t^6) \quad (2.208)$$

$$\{v\}(t + \delta t) = \{v_2\} + 2(\delta t/2)^5 \{\phi\} + O(\delta t^6) \quad (2.209)$$

where  $\{v_1\}$  is the result obtained with a time step  $\delta t$  and  $\{v_2\}$  the result obtained with two time steps  $\delta t/2$ . At fifth-order,  $\phi$  remains constant over the increment. The fifth-order Taylor series indicates that:  $\{\phi\} \simeq \frac{1}{5!} \{v^{(5)}\}$ . The difference between both estimations is an indicator of the error:

$$\{E\} = \{v_2\} - \{v_1\} \quad (2.210)$$

This difference is what must be kept smaller than a given precision by adjusting  $\delta t$ . Ignoring the  $(\delta t)^6$  terms, (2.208) and (2.209) can be solved. It follows that:

$$\{v\}(t + \delta t) = \{v_2\} + \frac{1}{15} \{E\} + O(\delta t^6) \quad (2.211)$$

We get then a fifth-order estimate. However, it is not always possible to control integration error. One can use  $\{E\}$  to modify the time step. Let  $\{E^0\}$  be the desired precision. Here, the precision is a vector. If  $E_i < E_i^0$ ,  $\forall i$ , then we can increase the next time step, if  $\exists i$ ,  $E_i > E_i^0$ , the current time step must be decreased. In both cases, correction factor ( $\alpha$ ) is given by:

$$\alpha = \frac{\delta t_{i+1}}{\delta t_i} = \min_i \left| \frac{E_i^0}{E_i} \right|^{1/5} \quad (2.212)$$

because the error is in  $\delta t^5$  for the fourth-order Runge–Kutta method. We take, of course, the most penalizing factor.  $\{E^0\}$  is still to be chosen. We can choose a fraction of the value of  $\{v\}$ :

$$E_i^0 = \epsilon |v_i| \quad (2.213)$$

In cases where  $v_i$  is close to 0 one can use the following trick:

$$E_i^0 = \epsilon \left( |v_i| + \delta t \left| \frac{dv_i}{dt} \right| \right) \quad (2.214)$$

Precision  $\{E^0\}$  must be obtained over the full increment  $\Delta t$  and not over the local ones  $\delta t_k$ . The smaller  $\delta t$  is compared to  $\Delta t$ , the smaller the imposed precision must be. In such a case, one can use the following approximation:

$$E_i^0 = \epsilon \delta t \left| \frac{dv_i}{dt} \right| \quad (2.215)$$

Then, the local precision depends on  $\delta t$ , and the exponent of the formula (2.212) is not exact anymore: it must be  $(1/4)$  instead of  $(1/5)$ .

Considering that these two exponents differ little and that error estimates are not exact, we use, in a practical approach, the following formula to update the local time step:

$$\alpha = \begin{cases} S \min_i \left| \frac{E_i^0}{E_i} \right|^{1/5} & \text{if } \delta t \text{ increases} \\ S \min_i \left| \frac{E_i^0}{E_i} \right|^{1/4} & \text{if } \delta t \text{ decreases} \end{cases} \quad (2.216)$$

where  $S$  is a security factor slightly smaller than 1.

### 2.7.3. $\theta$ -methods

$\theta$ -methods are implicit methods. There exist two possible formulations. Let  $\{\Delta v\}$  be the increment of the vector  $\{v\}$  over the time step  $\Delta t$ ; there are two ways to evaluate  $\{\Delta v\}$  (with  $0 \leq \theta \leq 1$ ):

$$\{\Delta v\} = \begin{cases} \Delta t \{\dot{v}\}(t + \theta \Delta t) & \text{type 1} \\ \Delta t ((1 - \theta)\{\dot{v}\}(t) + \theta\{\dot{v}\}(t + \Delta t)) & \text{type 2} \end{cases} \quad (2.217)$$

$\theta = 0$  corresponds to the explicit Euler method. If  $\theta \neq 0$  the previous equations must be solved in order to find  $\Delta v$  that satisfies them.

*Type 1 method (generalized middle point)*

$\{\Delta v\}$  is obtained by solving the following nonlinear system of equations:

$$\{R\} = \{\Delta v\} - \Delta t \{\dot{v}\}(t + \theta \Delta t) = \{0\} \quad (2.218)$$

One uses a Newton method for which one has to evaluate the Jacobian  $[J]$  of the previous equation:

$$[J] = \frac{\partial \{R\}}{\partial \{\Delta v\}} = [1] - \Delta t \left. \frac{\partial \{\dot{v}\}}{\partial \{\Delta v\}} \right|_{t+\theta \Delta t} \quad (2.219)$$

*Type 2 method (generalized trapezoidal rule)*

$\{\Delta v\}$  is obtained by solving the following nonlinear system of equations:

$$\{R\} = \{\Delta v\} - \Delta t ((1 - \theta)\{\dot{v}\}(t) + \theta\{\dot{v}\}(t + \Delta t)) = \{0\} \quad (2.220)$$

As previously, one uses the Newton method for which one has to evaluate the Jacobian:

$$[J] = \frac{\partial \{R\}}{\partial \{\Delta v\}} = [1] - \Delta t \theta \frac{\partial \{\dot{v}\}}{\partial \Delta v} \Big|_{t+\Delta t} \quad (2.221)$$

**2.7.4. Comment**

*Which integrating method?*

$\theta$ -methods require calculation of the Jacobian, which cannot always be evaluated. In that case, one must use Runge–Kutta type methods. Calculation of  $[J]$  can be quite heavy. However one can use an approximation: the convergence towards a solution is then slower.

*Calculation of the consistent tangent matrix with a  $\theta$ -method*

For the small strain case where one postulates the additive separation of strains, i.e.,

$$\underline{\underline{\epsilon}} = \underline{\underline{\epsilon}}_e + \underline{\underline{\epsilon}}_p + \underline{\underline{\epsilon}}_{th} \quad (2.222)$$

The  $\theta$ -method allows also, after convergence, calculation of the consistent tangent matrix of the behavior. One can rewrite the system of equations to solve as:

$$\{R\} = \{R_0\} \quad (2.223)$$

with

$$\{R\} = \left\{ \begin{array}{c} \Delta \underline{\underline{\epsilon}}_e + \Delta \underline{\underline{\epsilon}}_p \\ \Delta v_i - \Delta t \dot{v}_i \end{array} \right\} \quad (2.224)$$

and

$$\{R_0\} = \left\{ \begin{array}{c} \Delta \underline{\underline{\epsilon}} - \Delta \underline{\underline{\epsilon}}_{th} \\ 0 \end{array} \right\} \quad (2.225)$$

$\{R_0\}$  is imposed. One can then apply an infinitesimal variation to  $\{R_0\}$ .

$$\{\delta R_0\} = \left\{ \begin{array}{c} \delta \Delta \underline{\underline{\epsilon}} \\ 0 \end{array} \right\} \quad (2.226)$$

and we get

$$\{\delta R\} = [J].\{\delta v\} = \{\delta R_0\} \quad (2.227)$$

then

$$\{\delta v\} = [J]^{-1}.\{\delta R_0\} \quad (2.228)$$

It follows that

$$\delta \Delta \underline{\underline{\varepsilon}}_e = \underline{\underline{J}}_e^* : \delta \Delta \underline{\underline{\varepsilon}} \quad (2.229)$$

where  $[\underline{\underline{J}}_e^*]$  is the part of  $[J]^{-1}$  corresponding to the upper left part:

$$[J]^{-1} = [J^*] = \begin{bmatrix} [\underline{\underline{J}}_e^*] & [J_{12}^*] \\ [J_{21}^*] & [J_{22}^*] \end{bmatrix} \quad (2.230)$$

As  $\underline{\underline{\sigma}} = \underline{\underline{\Lambda}} : \underline{\underline{\varepsilon}}_e$  we get:

$$\delta \Delta \underline{\underline{\sigma}} = \underline{\underline{\Lambda}} : \delta \Delta \underline{\underline{\varepsilon}}_e \quad (2.231)$$

The tangent operator of the behavior is then:

$$\underline{\underline{L}}_c = \underline{\underline{\Lambda}} : \underline{\underline{J}}_e^* \quad (2.232)$$

It is important to note that we will get from  $\underline{\underline{L}}_c$  a “consistent tangent matrix” [SIM85a] (*i.e.*, consistent with the integration scheme), estimated from the incremental form of the differential system. The resulting form is different from the one resulting from rate formulation, as we will see below.

### *Consistent tangent matrix calculation for Runge–Kutta type integration*

In the framework of the Runge–Kutta method, it is not possible to get directly the tangent matrix of the behavior. One can however calculate it by perturbing the solution. Let  $\Delta \underline{\underline{\varepsilon}}$  be the strain increment leading to the stress increment  $\Delta \underline{\underline{\sigma}}$ . The tangent matrix is calculated by perturbation of  $\Delta \underline{\underline{\varepsilon}}$ . Using Voigt notation, one gets:

$$L_{cij} = \frac{\sigma_i(\{\Delta \underline{\underline{\varepsilon}}\} + \delta \varepsilon_j \{b^j\}) - \sigma_i(\{\Delta \underline{\underline{\varepsilon}}\})}{\delta \varepsilon_j} \quad (2.233)$$

where  $\{b^j\}$  is the vector such that:

$$b_i^j = \delta_{ij} \quad (2.234)$$

## 2.8. Finite element

The aim of this section is to recall some notions of the finite element method. The role of material behavior in the framework of this method will be particularly emphasized. For a more comprehensive presentation, one can refer to [BAT82, HUG87, BAT91, ZIE00].

### 2.8.1. Spatial discretization

Space is discretized by sets of nodes, edges, faces (3D) and elements. A point in space is located using its spatial coordinates in physical space  $\underline{x} = (x, y, z)$  and by its coordinates in the reference space of the element to which it belongs  $\underline{\eta} = (\eta, \zeta, \xi)$  (Fig. 2.16). Nodal coordinates of an element are given by  $\underline{x}^i$ ,  $i = 1 \dots n$ . Coordinates  $\underline{x}$  are linked to coordinates  $\underline{\eta}$  by the relation (sum from 1 to  $n$  over  $i$ ):

$$\underline{x} = N^i(\underline{\eta}) \underline{x}^i \quad (2.235)$$

where  $N^i$  are shape functions such that:

$$N^i(\underline{\eta}_j) = \delta_{ij} \quad \text{and} \quad \sum_i N^i(\underline{\eta}) = 1, \quad \forall \underline{\eta} \quad (2.236)$$

One will find in the books given as references the expression of the shape functions of each element.

Furthermore, one defines the Jacobian matrix of the transformation from the reference space to the physical space  $\underline{\eta} \rightarrow \underline{x}$ :

$$\underline{J} = \frac{\partial \underline{x}}{\partial \underline{\eta}} \quad (2.237)$$

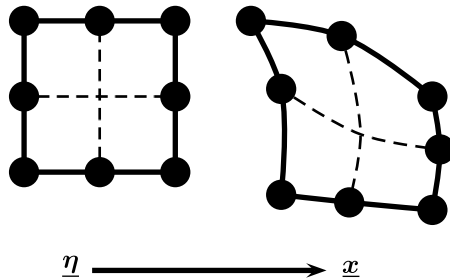


Figure 2.16. Spatial discretization (• = node)



Then:

$$J_{ij} = \frac{\partial x_i}{\partial \eta_j} = \frac{\partial (N^k x_i^k)}{\partial \eta_j} = x_i^k \frac{\partial N^k}{\partial \eta_j} \quad (2.238)$$

The Jacobian is:

$$J = \det(\underline{J}) \quad (2.239)$$

### 2.8.2. Discrete integration method

To perform spatial integrations (volume, surfaces, lines), a Gauss method is generally used which consists in replacing a continuous integration by a discrete sum such as (written here in the one-dimensional case):

$$\int_{-1}^1 f(x) dx = w^i f(x^i) \quad (2.240)$$

where the  $x^i$  are the predetermined positions where the function  $f$  is evaluated (Gauss points) and the  $w^i$  are weights associated to each of them. A discrete integration with  $n$  Gauss points allows us to exactly evaluate a polynomial of order  $2n - 1$ . In the case of higher-order polynomials or non-polynomial functions, Gauss integration gives an approximation of the actual integral. The method is then easily extended to 2D and 3D cases for the case of square reference elements (2D) or brick elements (3D) for which positions and weight of Gauss points can be calculated from the 1D case. For the triangular elements (2D) and tetrahedron (3D), positions and weight must be calculated specifically.

An integral over a finite element volume  $V_e$  can be written as follows, in order to integrate over the reference element  $V_r$ :

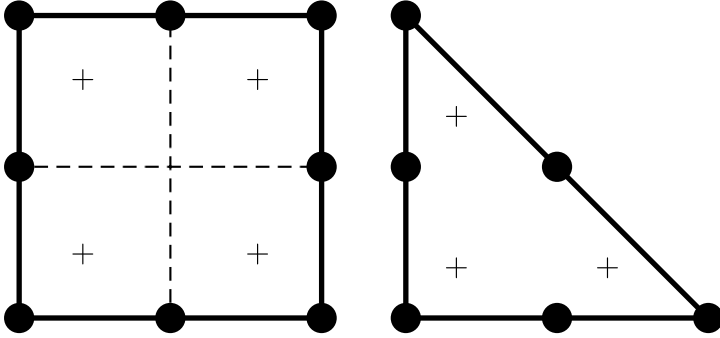
$$\int_{V_e} f(\underline{x}) dx = \int_{V_r} f(\underline{\eta}) J d\eta = f(\underline{\eta}^i) (J w^i) \quad (2.241)$$

where  $\underline{\eta}^i$  is the position of the various Gauss points in the reference space (Fig. 2.17).  $v^i$  will be the volume of the Gauss point  $i$ , with:

$$v^i = J w^i \quad (2.242)$$

### 2.8.3. Discretization of fields of unknowns

Fields of unknowns (displacement, temperature, pressure, etc.) must also be discretized in order to allow problem solving. These fields can be defined:



**Figure 2.17.** Gauss points in parent element: 8 nodes square element, 6 nodes triangle element: (● = node, + = Gauss point)

at the nodes: displacement, pressure, temperature

for each element: pressure when treating incompressibility case, plane strain.

by group of elements: periodical elements used in the framework of homogenization  
plane elements including torsion/flexion effects in a third direction

In general, one uses fields defined at the nodes. These are interpolated using the interpolation function of the elements. They are then continuous. In the case of a displacement field, we have:

$$\underline{u}(\underline{\eta}) = N^k(\underline{\eta})\underline{u}^k \quad \text{or} \quad u_i(\underline{\eta}) = N^k(\underline{\eta})u_i^k \quad (2.243)$$

where  $\underline{u}^k$  is the displacement at the node  $k$ . In the same way, for thermal problems we will have ( $T$ : temperature):

$$T(\underline{\eta}) = N^i(\underline{\eta})T^i \quad (2.244)$$

Gradients can then be simply calculated:

$$(\underline{\text{grad}} \underline{u})_{ij} = \frac{\partial u_i}{\partial x_j} = \frac{\partial u_i}{\partial \eta_k} \frac{\partial \eta_k}{\partial x_j} = u_i^n \frac{\partial N^n}{\partial \eta_k} \frac{\partial \eta_k}{\partial x_j} \quad (2.245)$$

$$(\underline{\text{grad}} T)_i = \frac{\partial T}{\partial x_i} = \frac{\partial T}{\partial \eta_k} \frac{\partial \eta_k}{\partial x_i} = T^n \frac{\partial N^n}{\partial \eta_k} \frac{\partial \eta_k}{\partial x_i} \quad (2.246)$$

By writing second-order tensors as vectors (Voigt notation), one can rewrite the previous equations as matrix-vector products:

$$\{\underline{\text{grad}} \underline{u}\} = [B] \cdot \{u^e\} \quad (2.247)$$

$$\underline{\text{grad}} T = [A] \cdot \{T^e\} \quad (2.248)$$

where  $\{u^e\}$  and  $\{T^e\}$  represent vectors of variables relative to the considered element:

$$\{u^e\} = \begin{Bmatrix} u_1^1 \\ u_2^1 \\ u_3^1 \\ \vdots \\ u_1^n \\ u_2^n \\ u_3^n \end{Bmatrix} \quad \{T^e\} = \begin{Bmatrix} T^1 \\ \vdots \\ T^n \end{Bmatrix} \quad (2.249)$$

Processing incompressibility (7.3.2), plain stress (7.3.3) and periodic conditions (7.3.4) will provide examples of fields of unknowns defined per element or per group of elements.

*Isoparametric elements:* we call isoparametric elements, the elements for which unknowns and coordinates are interpolated in the same way: interpolation function and shape functions are then identical.

#### 2.8.4. Application to mechanics

##### *Equilibrium formulation*

The discretized displacement field sought by the finite element method verifies the displacement boundary conditions: it is then kinematically admissible. However the stress field associated to the displacement field is not necessarily statically admissible. Then, the problem reduces to finding the displacement field that satisfies the PVP: the associated stress field will be statistically admissible.

##### *Greenberg's minimum principle*

We assume here that material behaviors can be written as (elasticity and plasticity cases):

$$\underset{\sim}{\dot{\sigma}} = \underset{\sim}{L} : \underset{\sim}{\dot{\epsilon}} \quad (2.250)$$

If  $\underset{\sim}{\dot{u}}'$  is a kinematically admissible field, it is associated with a strain field  $\underset{\sim}{\dot{\epsilon}}'$  and a stress field  $\underset{\sim}{\dot{\sigma}}' = \underset{\sim}{L} : \underset{\sim}{\dot{\epsilon}}'$ . Generally the field  $\underset{\sim}{\dot{\sigma}}'$  is not statically admissible. One shows that the solution field  $\underset{\sim}{\dot{u}}$  minimizes the functional  $\mathcal{B}(\underset{\sim}{\dot{u}})$  defined by:

$$\mathcal{B} = \frac{1}{2} \int_{\Omega} \underset{\sim}{\dot{\sigma}}' : \underset{\sim}{\dot{\epsilon}}' d\Omega - \int_{\Omega} \underset{\sim}{f} \cdot \underset{\sim}{\dot{u}} d\Omega - \sum_{i=1}^3 \int_{S_f^i} \dot{T}_i^d \dot{u}_i^d dS \quad (2.251)$$

Actually, it suffices to calculate the difference  $\mathcal{B}(\underline{\dot{\mathbf{u}}}') - \mathcal{B}(\underline{\dot{\mathbf{u}}})$ :

$$\begin{aligned} \mathcal{B}(\underline{\dot{\mathbf{u}}}') - \mathcal{B}(\underline{\dot{\mathbf{u}}}) &= \frac{1}{2} \int_{\Omega} (\underline{\dot{\boldsymbol{\sigma}}}' : \underline{\dot{\boldsymbol{\varepsilon}}}' - \underline{\dot{\boldsymbol{\sigma}}} : \underline{\dot{\boldsymbol{\varepsilon}}}) d\Omega \\ &\quad - \int_{\Omega} \underline{\dot{\mathbf{f}}} \cdot (\underline{\dot{\mathbf{u}}}' - \underline{\dot{\mathbf{u}}}) d\Omega - \sum_{i=1}^3 \int_{S_f^i} \dot{T}_i^d (\dot{u}_i' - \dot{u}_i) dS \end{aligned} \quad (2.252)$$

The field  $\underline{\dot{\boldsymbol{\sigma}}}$  being a solution, it is statically admissible. In addition,  $\underline{\boldsymbol{\varepsilon}}$  and  $\underline{\boldsymbol{\varepsilon}}'$  are statically admissible. By applying the PVP, we get that:

$$\int_{\Omega} \underline{\dot{\boldsymbol{\sigma}}} : (\underline{\dot{\boldsymbol{\varepsilon}}}' - \underline{\dot{\boldsymbol{\varepsilon}}}) d\Omega = \int_{\Omega} \underline{\dot{\mathbf{f}}} \cdot (\underline{\dot{\mathbf{u}}}' - \underline{\dot{\mathbf{u}}}) d\Omega + \sum_{i=1}^3 \int_{S_f^i} \dot{T}_i^d (\dot{u}_i' - \dot{u}_i) dS \quad (2.253)$$

where

$$\mathcal{B}(\underline{\dot{\mathbf{u}}}') - \mathcal{B}(\underline{\dot{\mathbf{u}}}) = \frac{1}{2} \int_{\Omega} (\underline{\dot{\boldsymbol{\sigma}}}' : \underline{\dot{\boldsymbol{\varepsilon}}}' + \underline{\dot{\boldsymbol{\sigma}}} : \underline{\dot{\boldsymbol{\varepsilon}}} - 2\underline{\dot{\boldsymbol{\sigma}}} : \underline{\dot{\boldsymbol{\varepsilon}}}') dV \quad (2.254)$$

$\underline{\dot{\boldsymbol{\sigma}}}$  and  $\underline{\dot{\boldsymbol{\sigma}}}'$  are then replaced by using the constitutive equation (2.250):  $\underline{\dot{\boldsymbol{\sigma}}} = \underline{\tilde{\mathbf{L}}} : \underline{\dot{\boldsymbol{\varepsilon}}}$  and  $\underline{\dot{\boldsymbol{\sigma}}}' = \underline{\tilde{\mathbf{L}}}' : \underline{\dot{\boldsymbol{\varepsilon}}}'$ . If both tensors  $\underline{\tilde{\mathbf{L}}}$  and  $\underline{\tilde{\mathbf{L}}}'$  are elastic, or both of them elastoplastic, the expression can be written in a simple form, which is positive, provided that the tensors are positive definite:

$$\mathcal{B}(\underline{\dot{\mathbf{u}}}') - \mathcal{B}(\underline{\dot{\mathbf{u}}}) = \frac{1}{2} \int_{\Omega} (\underline{\dot{\boldsymbol{\varepsilon}}}' - \underline{\dot{\boldsymbol{\varepsilon}}}) : \underline{\tilde{\mathbf{L}}} : (\underline{\dot{\boldsymbol{\varepsilon}}}' - \underline{\dot{\boldsymbol{\varepsilon}}}) dV \quad (2.255)$$

The positivity is still verified if one of the two tensors is elastic and the second one elastoplastic. This demonstrates that the solution field  $\underline{\dot{\mathbf{u}}}$  minimizes the functional  $\mathcal{B}$  and is unique. Should the elastoplastic tensor not be positive, the proposition is not exact anymore.

### 2.8.5. Finite element discretization of Greenberg's principle

We introduce here the matrix  $[B]$  that enables strain-rate calculation from nodal displacements (cf. Sect. 2.8.3):

$$\{\underline{\dot{\boldsymbol{\varepsilon}}}\} = [B] \cdot \{\dot{\mathbf{u}}^e\} \quad (2.256)$$

We get that  $\underline{\dot{\boldsymbol{\sigma}}} = \underline{\tilde{\mathbf{L}}} : \underline{\dot{\boldsymbol{\varepsilon}}}$  or  $\{\underline{\dot{\boldsymbol{\sigma}}}\} = [\underline{\tilde{\mathbf{L}}}] \cdot [B] \cdot \{\dot{\mathbf{u}}^e\}$ . The functional  $\mathcal{B}$  is then calculated element by element:

$$\mathcal{B} = \sum_e \mathcal{B}^e \quad (2.257)$$

with

$$\begin{aligned}
 \mathcal{B}^e(\underline{\dot{u}}) &= \frac{1}{2} \int_{V^e} \{\dot{u}^e\} \cdot [B]^T \cdot [\underline{\underline{L}}] \cdot [B] \cdot \{\dot{u}^e\} dV \\
 &\quad - \int_{V^e} \underline{\underline{f}} \cdot ([N] \cdot \{\dot{u}^e\}) dV - \sum_{i=1}^3 \int_{S_f^{e,i}} \dot{T}_i^d(\{N\} \cdot \{\dot{u}_i^e\}) dS \\
 &= \frac{1}{2} \{\dot{u}^e\} \cdot \left( \int_{V^e} [B]^T \cdot [\underline{\underline{L}}] \cdot [B] dV \right) \cdot \{\dot{u}^e\} \\
 &\quad - \left( \int_{V^e} [N]^T \cdot \underline{\underline{f}} dV \right) \cdot \{\dot{u}^e\} - \sum_{i=1}^3 \left( \int_{S_f^{e,i}} \dot{T}_i^d \{N\} dS \right) \cdot \{\dot{u}_i^e\} \quad (2.258)
 \end{aligned}$$

where  $[N]$  is the shape functions matrix given by:

$$[N] = \begin{bmatrix} N^1 & 0 & 0 & \dots & N^n & 0 & 0 \\ 0 & N^1 & 0 & \dots & 0 & N^n & 0 \\ 0 & 0 & N^1 & \dots & 0 & 0 & N^n \end{bmatrix} \quad (2.259)$$

$\{N\}$  is the shape functions vector:

$$\{N\} = \begin{Bmatrix} N^1 \\ \vdots \\ N^n \end{Bmatrix} \quad (2.260)$$

and  $\{u_i^e\}$  is the vector formed by the components  $i$  of the nodal displacement of the element:

$$\{u_i^e\} = \begin{Bmatrix} u_i^1 \\ \vdots \\ u_i^n \end{Bmatrix} \quad (2.261)$$

Each element contributes to the external  $\{F_e^e\}$  and internal  $\{F_i^e\}$  forces and to the stiffness  $[K^e]$ :

$$\{\dot{F}_e^e\} = \left( \int_{V^e} [N]^T \cdot \underline{\underline{f}} dV \right) + \sum_{i=1}^3 \left( \int_{S_f^{e,i}} \dot{T}_i^d \{N\} dS \right) \quad (2.262)$$

$$[K^e] = \int_{V^e} [B]^T \cdot [\underline{\underline{L}}] \cdot [B] dV \quad (2.263)$$

and

$$\{\dot{F}_i^e\} = [K^e] \cdot \{\dot{u}^e\} = \int_{V^e} [B]^T \cdot [\underline{\underline{L}}] \cdot [B] \cdot \{\dot{u}^e\} dV = \int_{V^e} [B]^T \cdot \{\dot{\underline{\underline{\sigma}}}\} \quad (2.264)$$

The global vector of external and internal forces as well as the stiffness matrix are then obtained by assembly [HUG87]:

$$\{\dot{F}_i\} = \mathcal{A} \{\dot{F}_i^e\} \quad (2.265)$$

$$\{\dot{F}_e\} = \mathcal{A} \{\dot{F}_e^e\} \quad (2.266)$$

$$[K] = \mathcal{A}[K^e] \quad (2.267)$$

where  $\mathcal{A}$  represents the assembly operation. One wants to minimize  $\mathcal{B}(\dot{\underline{u}})$ , i.e., to find  $\{\dot{\underline{u}}\}$  such that:

$$\frac{d\mathcal{B}}{d\{\dot{\underline{u}}\}} = \{0\} \quad (2.268)$$

where  $\{\dot{\underline{u}}\}$  is the vector containing all the unknowns. The system to solve can be reformulated as:

$$[K].\{\dot{\underline{u}}\} = \{\dot{F}_e\} \quad (2.269)$$

We will call  $\{F_e\}$  the vectors of the external forces corresponding to the loads applied on the structure,  $\{F_i\}$  the vector of the internal forces and  $[K]$  the global stiffness matrix. Until now, the problem has been written using a rate formulation but it will be solved in an incremental form expressed as:

$$\{\Delta F_i\} = [K_e].\{\Delta u\} = \{\Delta F_e\} \quad (2.270)$$

in the small strain framework. The small strain case corresponds to the situation where the displacements are small compared to the size ( $L$ ) of the structures  $u \ll L$  (small displacements) and where the strains are small  $\varepsilon \ll 1$ . The change from the rate to the incremental formulation is the basis of the finite element treatment of the small strain case. The possible nonlinearities in the resolution of the problem result from the nonlinearity of the material behavior only.

One notices that the internal forces can also be calculated as follows:

$$\{F_i^e\} = \int_{V_e} [B]^T . \{\underline{\sigma}\} dV \quad (2.271)$$

$\{F_i\}$ ,  $\{F_e\}$  and  $[K]$  are calculated element by element and then assembled. The internal forces of an element are calculated by Gauss integration:

$$\{F_i^e\} = \int_{V_e} [B]^T . \{\underline{\sigma}\} dV = \sum_g [B_g]^T . \{\underline{\sigma}_g\} (J_g w_g) \quad (2.272)$$

The elementary stiffness matrix is assessed in the following way:

$$[K^e] = \int_{V_e} [B]^T . [\underline{\underline{L}}] . [B] dV = \sum_g [B_g]^T . [\underline{\underline{L}}] . [B_g] (J_g w_g) \quad (2.273)$$

Once the vectors  $\{F_i\}$  and  $\{F_e\}$  have been assembled, their size,  $n_d$ , corresponds to the number of degrees of freedom of the problem. The matrix  $[K]$  has a size  $n_d \times n_d$ .  $\{F_i\}_k$  is the reaction associated with the degree of freedom  $k$ .

### 2.8.6. Another presentation of the finite element discretization

One can get the principle of the finite-element discretization from the principle of virtual power. The discretization of the power of external forces leads naturally to (2.262) for each element. The discretized displacement field verifies the boundary conditions in displacement: it is then kinematically admissible. However, the stress field associated with the displacement field is not necessarily statically admissible. The problem boils down to finding the displacement field  $\underline{u}$  (discretized by  $\{u\}$ ) that allows the PVP to be satisfied: the associated stress field will then be statically admissible. For any virtual field  $\underline{\dot{u}}$ , discretized by  $\{\dot{u}\}$ , the power of internal stress is given by:

$$\begin{aligned}
 w_i &= \int_{\Omega} \underline{\sigma}(\underline{u}) : \underline{\dot{\epsilon}}(\underline{\dot{u}}) dV \\
 &= \sum_e \int_{V_e} \{\underline{\sigma}(\{u^e\})\} \cdot [B] \cdot \{\dot{u}^e\} dV \\
 &= \sum_e \left( \int_{V_e} [B]^T \cdot \{\underline{\sigma}(\{u^e\})\} dV \right) \cdot \{\dot{u}^e\} \\
 &\equiv \{F_i(\{u\})\} \cdot \{\dot{u}\}
 \end{aligned} \tag{2.274}$$

With each degree of freedom is associated an internal reaction. All of them are gathered in the vector  $\{F_i(\{u\})\}$ , a definition of which is given by the previous equation. It must be verified for any kinematically admissible field  $\{\dot{u}\}$ . It follows that  $\{F_i\}$  is equal to:

$$\{F_i\} = \mathcal{A}(\{F_i^e\}) \quad \text{with} \quad \{F_i^e\} = \int_{V_e} [B]^T \cdot \{\underline{\sigma}(\{u^e\})\} dV \tag{2.275}$$

The problem to solve with respect to  $\{u\}$  is then:

$$\{F_i(\{u\})\} = \{F_e\} \tag{2.276}$$

It is a nonlinear system that can be solved by Newton's method. It is suitable to calculate the Jacobian matrix:

$$[K] = \frac{\partial \{F_i(\{u\})\}}{\partial \{u\}} = \mathcal{A} \frac{\partial \{F_i^e(\{u^e\})\}}{\partial \{u^e\}} \equiv \mathcal{A}[K^e] \tag{2.277}$$

Let us calculate  $[K^e]$ :

$$\begin{aligned}
 [K^e] &= \frac{\partial \{F_i^e(\{u^e\})\}}{\partial \{u^e\}} \\
 &= \int_{V_e} [B]^T \cdot \frac{\partial \{\underline{\sigma}\}}{\partial \{\underline{\epsilon}\}} \cdot \frac{\partial \{\underline{\epsilon}\}}{\partial \{u^e\}} = \int_{V_e} [B]^T \cdot [\underline{\underline{L}}_c] \cdot [B] dV
 \end{aligned} \tag{2.278}$$

We find, of course, the same result as in (2.263). However, it appears very clearly that the stiffness matrix  $[K]$  is only used for the iterative resolution of (2.276). One can

then use an approximation of it, either by an approximation of  $\underline{L}_c$ , or by using the BFGS method to solve (2.276).

*Note:* In (2.276),  $\{F_e\}$  can, in some cases, depend on  $\{u\}$ . Corrective terms must then be applied to the matrix  $[K]$ .

### 2.8.7. Assembly through example

In order to better understand the assembly operation, let us consider a simple example. A mesh containing three linear elements A, B, C and seven nodes 1...7 is represented in Fig. 2.18. It is a planar mechanical problem; there are 14 degrees of freedom ( $7 \times 2$ ) corresponding to the node displacements. In each element, the nodes have a local numbering written in *italic*. The vector of the unknown  $\{u\}$  is then:

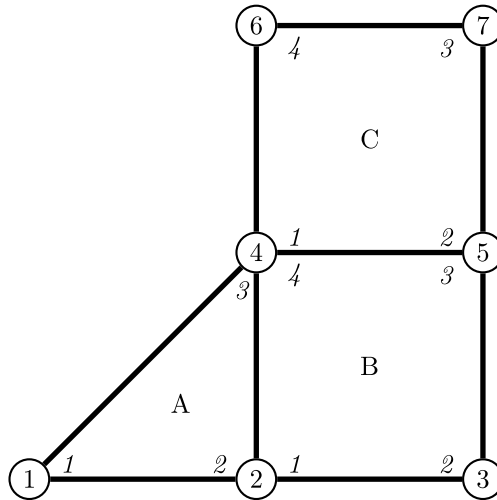
$$\{u\} = \{u_x^1, u_y^1, u_x^2, u_y^2, u_x^3, u_y^3, u_x^4, u_y^4, u_x^5, u_y^5, u_x^6, u_y^6, u_x^7, u_y^7\} \quad (2.279)$$

For the element A, the vector of the local unknown  $\{u^A\}$  is:

$$\{u^A\} = \{u_x^{A1}, u_y^{A1}, u_x^{A2}, u_y^{A2}, u_x^{A3}, u_y^{A3}\} = \{u_x^1, u_y^1, u_x^2, u_y^2, u_x^4, u_y^4\} \quad (2.280)$$

The vector of the internal forces associated with the element A is:

$$\{F_i^A\} = \{F_x^{A1}, F_y^{A1}, F_x^{A2}, F_y^{A2}, F_x^{A3}, F_y^{A3}\} \quad (2.281)$$



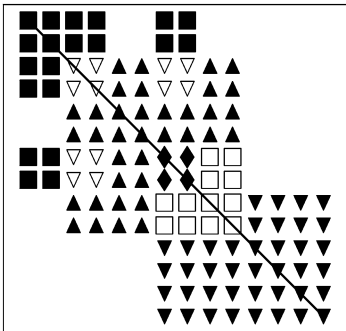
**Figure 2.18.** Example of meshes (mesh: A, B and C) showing the global numbering of the nodes and the local one (in *italic*)



The elementary stiffness matrix  $[K^A]$  is a  $6 \times 6$  full matrix. The assembly of the vectors  $\{F_i^A\}$ ,  $\{F_i^B\}$  and  $\{F_i^C\}$  gives the vector  $\{F_i\}$  such that:

$$\{F_i\} = \left\{ \begin{array}{l} F_x^1 = F_x^{A1} \\ F_y^1 = F_y^{A1} \\ F_x^2 = F_x^{A2} + F_x^{B1} \\ F_y^2 = F_y^{A2} + F_y^{B1} \\ F_x^3 = F_x^{B2} \\ F_y^3 = F_y^{B2} \\ F_x^4 = F_x^{A3} + F_x^{B4} + F_x^{C1} \\ F_y^4 = F_y^{A3} + F_y^{B4} + F_y^{C1} \\ F_x^5 = F_x^{B3} + F_x^{C2} \\ F_y^5 = F_y^{B3} + F_y^{C2} \\ F_x^6 = F_x^{C4} \\ F_y^6 = F_y^{C4} \\ F_x^7 = F_x^{C3} \\ F_y^7 = F_y^{C3} \end{array} \right\} \quad (2.282)$$

The assembly of the elementary matrices  $[K^A]$ ,  $[K^B]$  and  $[K^C]$  gives a matrix whose non-null terms correspond to couples of columns and lines whose number represents nodes belonging to the same element, or nodes that have been connected by the user:



$[K] =$

- A
- ▲ B
- ▼ C
- ▽ A and B
- △ A and C
- B and C
- ◆ A, B and C

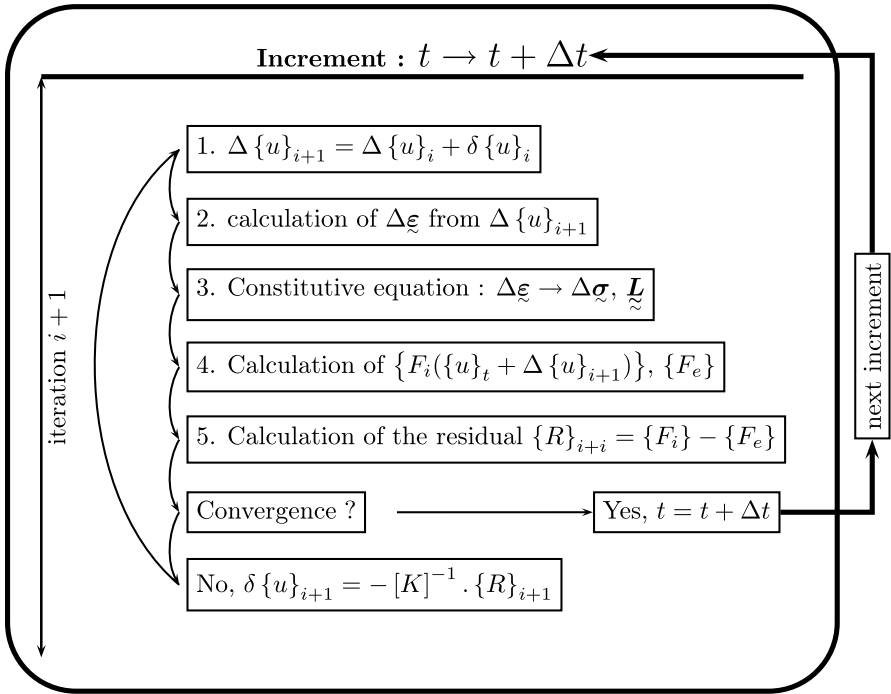
(2.283)

The various symbols ■, ▲, ▼, ▽, △, □, ◆ indicate that the corresponding terms of the matrix  $[K]$  are the sum of the terms coming from the various elementary matrices. For example:

$$K_{10,7}^T = K_{6,7}^B + K_{4,1}^C \quad (2.284)$$

### 2.8.8. Principle of resolution

It is practically impossible, even for a monotonic loading, to solve (2.276) for any loading. Because of the nonlinearities, the iterative search process of the solution does not converge. The solution is then obtained using a progressive incremental loading. At time  $t$ , the solution  $\{u\}_t$  is known. One applies then an increment  $\Delta t$ , and one searches iteratively the corresponding increment of the solution  $\Delta\{u\}$ . The chart of the algorithm is given in Fig. 2.19.



**Figure 2.19.** Iterative search of increment  $\Delta\{u\}$  for a time step  $\Delta t$

*Note:* For each increment, one searches iteratively the increment  $\Delta\{u\}$ . It is generally initialized with  $\Delta\{u\} = \{0\}$ . However, it is possible, particularly in the case of monotonic loading, to initialize  $\Delta\{u\}$  for increment  $j > 1$  to

$$\Delta\{u\}_{\text{increment}}^{\text{initial } j} = \frac{\Delta t_j}{\Delta t_{j-1}} \Delta\{u\}_{\text{increment}}^{\text{solution } j-1} \quad (2.285)$$

### 2.8.9. Mechanical behavior in the finite element method

The mechanical behavior in finite element method must then, for a strain increment  $\Delta \underline{\underline{\epsilon}}$ , give the corresponding stress increment  $\Delta \underline{\underline{\sigma}}$  along with the consistent tangent matrix of the mechanical behavior

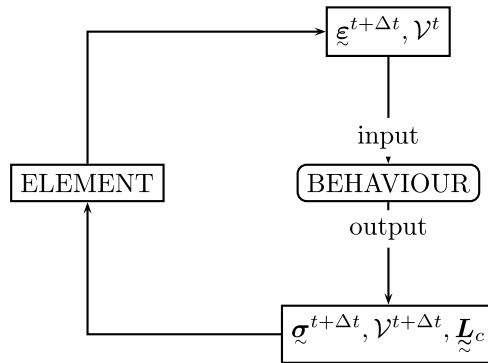
$$\underline{\underline{L}}_c = \frac{\partial \Delta \underline{\underline{\sigma}}}{\partial \Delta \underline{\underline{\epsilon}}} \quad (2.286)$$

The equation to solve is written:

$$\{F_i\} = \int_{\Omega} [B]^T \cdot \{\sigma\} dV \equiv \{F_e\} \quad (2.287)$$

Only the stress tensor  $\underline{\underline{\sigma}}$  is necessary for the equilibrium formulation. The tensor  $\underline{\underline{L}}_c$  is only used for the calculation of the elementary stiffness matrix  $[K^e]$  that is involved in the iterative solving of the previous system. It is then not necessary to exactly calculate  $\underline{\underline{L}}_c$ : an estimate can be sufficient. For example, one can use the elasticity matrix  $\underline{\underline{A}}$  for  $\underline{\underline{L}}_c$  and use a BFGS type method to correct the global matrix.

It is also appropriate not to forget that the material behavior is characterized by a set of variables ( $\mathcal{V}$ ) whose values must be calculated at the end of each increment  $\mathcal{V}^{t+\Delta t}$ . The interface between finite element and mechanical behavior can then be summed up [FOE97, BES98b]:



The software description of the mechanical behavior must not depend on the FEM. Reciprocally the element formulation/implementation must not depend on the behavior type (for example elasticity, plasticity, creep, etc.). There must just be correspondence between the element and the behavior formulation: small strain, finite strain, thermal problem, etc.

Non-Linear Mechanics of Materials

Besson, J.; Cailletaud, G.; Chaboche, J.-L.; Forest, S.

2010, XVIII, 433 p., Hardcover

ISBN: 978-90-481-3355-0