A CONVEXITY - MINIMIZATION OF FUNCTION SUBJECT TO CONSTRAINTS

In this Appendix, we will discuss certain mathematical tools that are employed in the thermodynamic treatment of evolution laws. When establishing the evolution laws, use is made of properties relating to *convex functions*. In plasticity, the evolution laws may be established by invoking the postulate of maximum dissipation. However, this postulate is to be fulfilled on condition that the stress state is located inside or on the yield surface. This constraint condition complicates the necessary mathematical tools considerably, in particular because the constraint is given in terms of an inequality. These topics are dealt with in the mathematical literature concerned with *nonlinear optimization* and for further information, we may refer, for instance, to Luenberger (1984) and Strang (1986).

A.1 Convex function

In the discussion of evolution laws that fulfill the dissipation inequality, extensive use is made of convex functions and we will therefore provide a strict mathematical definition of convexity of a function.

Considering a one-dimensional problem with the variable given by y, a *convex* function is illustrated in Fig. A.1a) whereas a *concave* function is shown in Fig. A.1b). Referring to Fig. A.2, it appears that for a one-dimensional problem, convexity may be expressed as

$$f((1-\alpha)y^{(1)} + \alpha y^{(2)}) \leq (1-\alpha)f(y^{(1)}) + \alpha f(y^{(2)})$$

where $y^{(1)}$ and $y^{(2)}$ are two arbitrary points and $0 \leq \alpha \leq 1$.

Let us generalize these results and consider a function $f(y_i)$ of N variables,

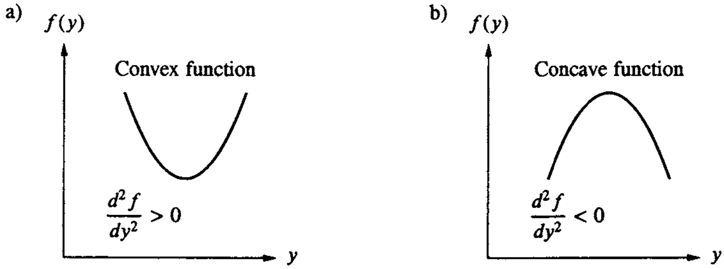

Figure A.1: One-dimensional illustration of: a) convex function and b) concave function.

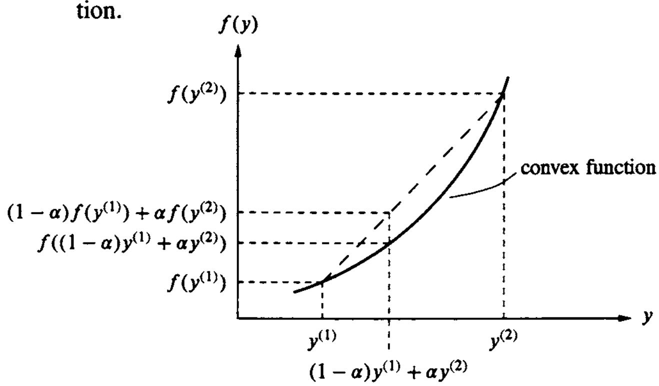

Figure A.2: Illustration of convexity.

i.e. $i = 1, 2, \cdots, N$. By definition, we then have

$$\boxed{\begin{array}{l} \textit{The function } f(y_i) \textit{ is convex, if, for any two} \\ \textit{points } y_i^{(1)} \textit{ and } y_i^{(2)}, \textit{ we have} \\ f((1-\alpha)y_i^{(1)} + \alpha y_i^{(2)}) \leq (1-\alpha)f(y_i^{(1)}) + \alpha f(y_i^{(2)}) \\ \textit{where } 0 \leq \alpha \leq 1 \end{array}} \tag{A.1}$$

The function f is said to be *strictly convex*, if the strict inequality holds whenever $y_i^{(1)} \neq y_i^{(2)}$ and $0 < \alpha < 1$.

It turns out to be advantageous to express this convexity property in a different form. From (A.1) we obtain for $0 < \alpha \leq 1$

$$\frac{f(y_i^{(1)} + \alpha \Delta y_i) - f(y_i^{(1)})}{\alpha} \leq f(y_i^{(2)}) - f(y_i^{(1)}) \tag{A.2}$$

where $\Delta y_i = y_i^{(2)} - y_i^{(1)}$. To evaluate the left-hand term we make a Taylor expansion about point $y_i^{(1)}$, i.e.

$$f(y_i^{(1)} + \alpha\Delta y_i) = f(y_i^{(1)}) + (\frac{\partial f}{\partial y_i})^{(1)}\alpha\Delta y_i + \mathcal{O}(\alpha^2)$$

where $(\ )^{(1)}$ means that the quantity within the parenthesis is to be evaluated at point $y_i^{(1)}$. From the expression above, it follows that

$$\frac{f(y_i^{(1)} + \alpha\Delta y_i) - f(y_i^{(1)})}{\alpha} = (\frac{\partial f}{\partial y_i})^{(1)}\Delta y_i + \mathcal{O}(\alpha)$$

Letting $\alpha \to 0$, we obtain

$$[\frac{f(y_i^{(1)} + \alpha\Delta y_i) - f(y_i^{(1)})}{\alpha}]_{\alpha\to 0} = (\frac{\partial f}{\partial y_i})^{(1)}\Delta y_i \tag{A.3}$$

It may be observed that the left-hand term in the mathematical literature is called the *directional derivative* of f when moving in the direction Δy_i; the terminology of *Gateaux* or *Frechet derivative* is also often used, see Vainberg (1964).

Use of (A.3) in (A.2) implies

$$(\frac{\partial f}{\partial y_i})^{(1)}\Delta y_i \leq f(y_i^{(2)}) - f(y_i^{(1)}) \tag{A.4}$$

We have then shown that (A.1) implies (A.4); let us next prove that (A.4) implies (A.1). For this purpose, we accept (A.4) and choose $y_i^{(1)}$ as $(1-\alpha)y_i^{(3)} + \alpha y_i^{(4)}$ and alternatively $y_i^{(2)}$ as $y_i^{(2)} = y_i^{(3)}$ or $y_i^{(2)} = y_i^{(4)}$. This leads to

$$\begin{aligned} -(\frac{\partial f}{\partial y_i})^{(1)}\alpha(y_i^{(4)} - y_i^{(3)}) &\leq f(y_i^{(3)}) - f((1-\alpha)y_i^{(3)} + \alpha y_i^{(4)}) \\ (\frac{\partial f}{\partial y_i})^{(1)}(1-\alpha)(y_i^{(4)} - y_i^{(3)}) &\leq f(y_i^{(4)}) - f((1-\alpha)y_i^{(3)} + \alpha y_i^{(4)}) \end{aligned}$$

Multiplying the first expression by $(1-\alpha)$ and the second by α and adding, give the result

$$f((1-\alpha)y_i^{(3)} + \alpha y_i^{(4)}) \leq (1-\alpha)f(y_i^{(3)}) + \alpha f(y_i^{(4)})$$

which is in complete agreement with (A.1). We have then proved that

$$\boxed{\begin{array}{l} \textit{The function } f(y_i) \textit{ is convex, if and only if,} \\ \textit{for any two points } y_i^{(1)} \textit{ and } y_i^{(2)}\textit{, we have} \\ f(y_i^{(2)}) - f(y_i^{(1)}) \geq (\frac{\partial f}{\partial y_i})^{(1)}(y_i^{(2)} - y_i^{(1)}) \end{array}} \tag{A.5}$$

In the one-dimensional case, this property is illustrated in Fig. A.3.

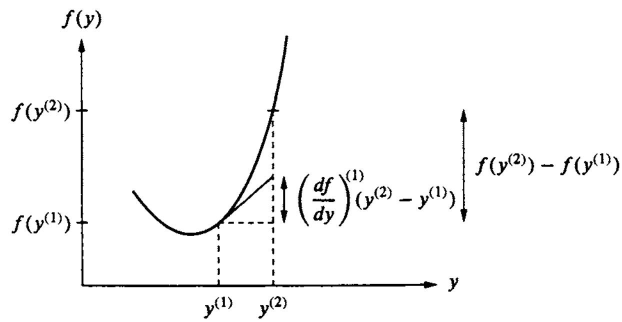

Figure A.3: Illustration of $f(y_i^{(2)}) - f(y_i^{(1)}) \geq (\frac{\partial f}{\partial y_i})^{(1)}(y_i^{(2)}) - y_i^{(1)})$.

It is also possible to express convexity in yet another form. We first make the following definition

$$\boxed{\textit{The Hessian } F_{ij} \textit{ of } f(y_i) \textit{ is defined by } F_{ij} = \frac{\partial^2 f}{\partial y_i \partial y_j}} \tag{A.6}$$

It appears that the Hessian is a symmetric matrix. Now a Taylor expansion about $y_i^{(1)}$ and use of (A.6) give

$$\begin{aligned} f(y_i^{(2)}) - f(y_i^{(1)}) &= (\frac{\partial f}{\partial y_i})^{(1)}(y_i^{(2)} - y_i^{(1)}) \\ &+ \frac{1}{2}(y_i^{(2)} - y_i^{(1)})F_{ij}((1-\theta)y_m^{(1)} + \theta y_m^{(2)})(y_j^{(2)} - y_j^{(1)}) \end{aligned} \tag{A.7}$$

For some θ-value in the range $0 \leq \theta \leq 1$, this result is exact. Use of (A.5) in (A.7) leads to

$$(y_i^{(2)} - y_i^{(1)})F_{ij}((1-\theta)y_i^{(1)} + \theta y_i^{(2)})(y_j^{(2)} - y_j^{(1)}) \geq 0$$

Since $y_i^{(2)} - y_i^{(1)}$ denotes arbitrary quantities, we conclude that F_{ij} must be positive semi-definite. On the other hand, if the Hessian F_{ij} is assumed to be positive semi-definite then (A.7) implies (A.5). We are then led to the following conclusion

$$\boxed{\begin{array}{l}\textit{The function } f(y_i) \textit{ is convex, if and only if, the} \\ \textit{Hessian } F_{ij} = \dfrac{\partial^2 f}{\partial y_i \partial y_j} \textit{ is positive semi-definite}\end{array}} \tag{A.8}$$

Referring back to the one-dimensional convex function $f(y)$ shown in Fig. A.1a), we see that convexity requires $d^2f/dy^2 \geq 0$ and (A.8) is the generalization of this result. With respect to Fig. A.1a), we may remark that the limit case $d^2f/dy^2 = 0$ corresponds to $f(y)$ being a straight line.

A.2 Unconstrained minimum

Consider the function $f(y_i)$ of N unknowns, i.e. $i = 1, 2, \cdots, N$. Assume that $f(y_i)$ at some point attains an extremum. At that point, we have per definition

$$df = \frac{\partial f}{\partial y_i} dy_i = 0 \tag{A.9}$$

In the present case, where we are considering unconstrained problems, the quantities dy_i can take any values. From (A.9) we therefore conclude that

$$\boxed{\frac{\partial f}{\partial y_i} = 0 \quad \Rightarrow \quad \textit{extremum point}} \tag{A.10}$$

This result evidently generalizes the trivial result when we have one variable, only.

Let us next devise a method by which we can decide whether the extremum is a minimum or not. Assume that a minimum point exists and that it is located at the point y_i^*. By definition, we have

$$f(y_i) \geq f(y_i^*) \quad \text{where} \quad y_i \neq y_i^* \tag{A.11}$$

A Taylor expansion of $f(y_i)$ about the point y_i^* provides

$$f(y_i) = f(y_i^*) + (\frac{\partial f}{\partial y_i})^*(y_i - y_i^*) + \frac{1}{2}(\frac{\partial^2 f}{\partial y_i \partial y_j})^*(y_i - y_i^*)(y_j - y_j^*)$$

where higher order terms were ignored. Due to (A.10), this reduces to

$$f(y_i) = f(y_i^*) + \frac{1}{2}(y_i - y_i^*)(\frac{\partial^2 f}{\partial y_i \partial y_j})^*(y_j - y_j^*)$$

For (A.11) to be fulfilled, it follows that the Hessian $F_{ij} = \partial^2 f / \partial y_i \partial y_j$ at the minimum point must be positive semi-definite. We conclude that

$$\boxed{\begin{array}{l} \textit{Necessary and sufficient conditions for } y_i^* \textit{ being} \\ \textit{a minimum point for the function } f(y_i)\textit{:} \\ \star \quad \dfrac{\partial f}{\partial y_i} = 0 \textit{ at point } y_i^* \\ \star \quad \textit{the Hessian } F_{ij} = (\dfrac{\partial^2 f}{\partial y_i \partial y_j})^* \textit{ is positive semi-definite} \end{array}} \tag{A.12}$$

If the function $f(y_i)$ is convex, cf. (A.8), it suffices to consider the condition $\partial f / \partial y_i = 0$ in order to identify the minimum point.

A.3 Inequality constrained minimum - Kuhn-Tucker relations

The problem of finding the minimum of a function where the variables are subject to some constraints, in particularly inequalities, is relevant for a number of applications. In plasticity, for instance, we want to maximize the mechanical dissipation γ_{mech} (minimize the quantity $-\gamma_{mech}$) subject to the condition that the yield criterion $f(\sigma_{ij}, K_\alpha) \leq 0$ is fulfilled. The solution to this problem is provided by the famous *Kuhn-Tucker relations* which are treated in detail, for instance, by Luenberger (1984) and Strang (1986); here, we will merely present the results.

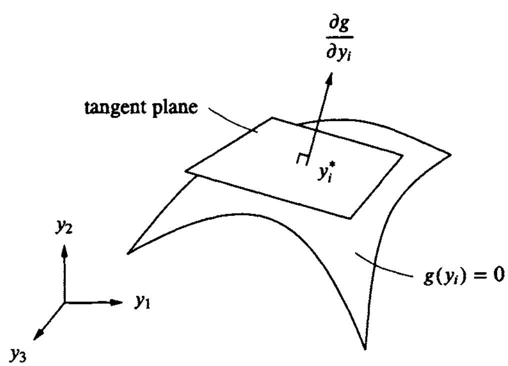

Figure A.4: Illustration of tangent plane. Three-dimensional problem with variables y_1, y_2 and y_3; one constraint equation.

The problem we are facing is that we want to minimize a function $f(y_i)$, where $i = 1, 2, \cdots, N$, subject to the inequality constraints $g^I(y_i) \leq 0$, where $I = 1, 2, \cdots, M$ and $M < N$. Considering a given constraint, it is classified as *active* if $g^I(y_i) = 0$ and *inactive* if $g^I(y_i) < 0$. A *regular point* is defined by

If at some point y_i, all the column matrices $\partial g^I/\partial y_i$ for the active constraints are linearly independent, then this point is called a regular point	(A.13)

For a given point y_i^* consider the expression

$$(\frac{\partial g^I}{\partial y_i})^*(y_i - y_i^*) = 0 \quad \text{for } I = 1, 2, ..., M \tag{A.14}$$

For each value of I, the gradient $\partial g^I/\partial y_i$ is orthogonal to the corresponding surface $g^I = 0$. It follows that the y_i−values that fulfill (A.14) define the so-

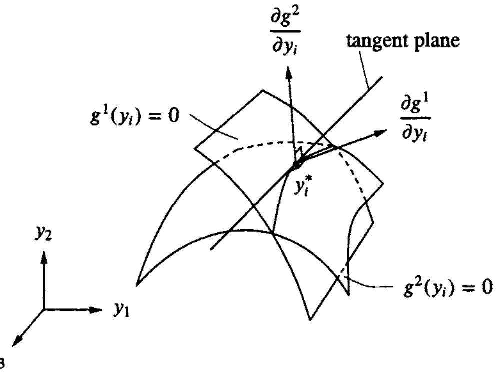

Figure A.5: Illustration of tangent plane. Three-dimensional problem with variables y_1, y_2 and y_3; two constraint equations.

called *tangent plane* at the position y_i^*. These tangent planes are illustrated in Figs. A.4 and A.5.

We then have

Problem: Determine minimum point y_i^* of the function $f(y_i)$, where $i = 1, 2, \cdots, N$, subject to the constraints $g^I(y_i) \leq 0$, where $I = 1, 2, \cdots, M$ and $M < N$.
Necessary conditions = Kuhn-Tucker relations:
The solution y_i^* must be a regular point. Moreover

$$\left(\frac{\partial f}{\partial y_i}\right)^* + \sum_{I=1}^{M} \lambda^I \left(\frac{\partial g^I}{\partial y_i}\right)^* = 0 \tag{A.15}$$

$\lambda^I \geq 0$ *for active constraints*
$\lambda^I = 0$ *for inactive constraints*
$\lambda^{(I)} g^{(I)} = 0$ *(no summation)*

which were derived by Kuhn and Tucker (1951). The result above can be interpreted in the following elegant manner. Define the Lagrange function $\mathcal{L}(y_i, \lambda^I)$ by

$$\mathcal{L}(y_i, \lambda^I) = f(y_i) + \sum_{I=1}^{M} \lambda^I g^I(y_i) \tag{A.16}$$

and determine the extremum point for $\mathcal{L}(y_i, \lambda^I)$ with respect to y_i as if the variables y_i are unconstrained. We then obtain

$$\frac{\partial \mathcal{L}}{\partial y_i} = \frac{\partial f}{\partial y_i} + \sum_{I=1}^{M} \lambda^I \frac{\partial g^I}{\partial y_i} = 0 \tag{A.17}$$

which is exactly the result we want.

The results above are necessary conditions for a minimum and we also have the results

> *Necessary and sufficient conditions:*
> *In addition to (A.15), we must have*
> $$(y_i - y_i^*)L_{ij}^*(y_j - y_j^*) \geq 0$$
> *where*
> $$L_{ij}^* = (\frac{\partial^2 f}{\partial y_i \partial y_j})^* + \lambda^I (\frac{\partial^2 g^I}{\partial y_i \partial y_j})^*$$
> *and y_i is located on the tangent plane defined by*
> $$(\frac{\partial g^I}{\partial y_i})^*(y_i - y_i^*) = 0 \quad \textit{for all active constraints}$$

This means that the quantity L_{ij}^* is positive semi-definite.

These conditions can be strengthened slightly. Since $\lambda^I \geq 0$ then, if the functions $f(y_i)$ and $g^I(y_i)$ are convex functions, i.e. the Hessians $F_{ij} = \partial^2 f/\partial y_i \partial y_j$ and $G_{ij}^I = \partial^2 g^I/\partial y_i \partial y_j$ are both positive semi-definite, this implies that L_{ij}^* is always positive semi-definite. We conclude that

> *If the functions $f(y_i)$ and $g^I(y_i)$ are convex functions,*
> *the condition $(y_i - y_i^*)L_{ij}^*(y_j - y_j^*) \geq 0$* (A.18)
> *is fulfilled*