PERGAMON

International Journal of Solids and Structures 40 (2003) 3717–3735

INTERNATIONAL JOURNAL OF SOLIDS and STRUCTURES

www.elsevier.com/locate/ijsolstr

Sufficient conditions for existence of physically significant solutions in limiting equilibrium slope stability analysis

R. Baker *

Department of Geotechnology, Faculty of Civil Engineering, Technion—Israel Institute of Technology, Technion City, Haifa 32000, Israel

Received 14 November 2001; received in revised form 8 January 2003

Abstract

Engineering assessment of slope stability is usually performed using limiting equilibrium analysis. This framework includes a process of minimization which identifies the critical slip surface and its associated minimal safety factor. The approach makes sense only if a minimum safety factor exists, i.e. if there is a slip surface for which the safety factor is smaller than safety factors associated with all other slip surfaces. The present work establishes conditions which guarantee that slope stability problems have a physically significant minimum. The question of existence of a minimum is relevant to all slope stability formulations which satisfy equilibrium conditions without a priori assumptions with respect to the shape of potential slip surfaces. The main purpose of the present work is to "legitimatize" the approximate, but practically useful, limiting equilibrium technique by placing it on secure foundations.

The present work shows that the restrictions required in order to ensure the existence of a minimum include three, well motivated, physical elements: (a) Two integral inequality constraints restricting legitimate forms of slip surfaces, and normal stress functions acting on them. These constraints represent the obvious observation that under usual conditions slopes fail by moving down and away from the main body of the slope. (b) The strength model (Mohr envelope), should imply a finite tensile strength. (c) A "cracking criterion" which specifies the consequences (crack formation) occurring when the soil's tensile strength is fully mobilized.

Keywords: Limiting equilibrium; Slope stability; Variational methods; Existence of minimum

1. Introduction

Engineering assessment of earth slope stability is usually performed using limiting equilibrium analysis. This framework includes a process of minimization which identifies the critical slip surface $y_c(x)$, and its associated minimal safety factor F_s. The approach makes sense only if a minimum safety factor exists, i.e. if there is a slip surface for which the safety factor is smaller than safety factors associated with all other slip surfaces. Classical presentations of slope stability analysis (e.g. Morgenstern and Price, 1965; Janbu, 1973; and many others); ignore this question, trusting essentially that the existence of a minimum safety factor is guaranteed by the physical nature of the problem.

* Fax: +972-482-93323.
E-mail address: baker@techunix.technion.ac.il (R. Baker).

0020-7683/03/$ - see front matter
doi:10.1016/S0020-7683(03)00075-1

Baker and Garber (1977, 1978) presented a variational approach to slope stability analysis, which generalized the classical formulations. This formulation has a number of fundamental advantages:

1. Use of variational calculus provides a consistent and practical procedure for identification of the critical pair $\{y_c(x), F_s\}$. Classical procedures rely on trial and error techniques in order to identify the critical conditions, and such techniques are not practical when the family of potential slip surfaces is not restricted, a priori, to some simple form (e.g. straight line, circle, spiral, etc.). It can be argued that classical formulations do not solve the slope stability problem, merely define it.
2. Unlike other presentations, the variational approach does not employ statical assumptions. These assumptions are replaced by minimization of the safety functional with respect to the normal stress distribution $\sigma(x)$ acting along potential slip surfaces $y(x)$. This minimization process results automatically with the most conservative "statical assumption" that is consistent with general principles of limit equilibrium analysis.
3. Application of the variational approach yields general results which are consistent with well-known plasticity solutions (e.g. in a linear and homogeneous setting critical slip surfaces may be either straight lines or a log-spirals, and the normal stress distribution along potential slip surfaces satisfies Koiter's differential equation). In a general, non-homogeneous case, or when a non-linear failure criterion is used, the variational approach yields two coupled first order differential equations controlling the functions $\{y(x), \sigma(x)\}$. These differential equations provide a natural generalization of classical plasticity results.
4. Numerical results based on this approach were presented by Garber and Baker (1977), Baker (1981) and Baker and Leshchinsky (1987). For homogeneous slopes, a linear failure criterion, and sufficient tensile strength to prevent the formation of tension cracks, the results of Baker (1981) are practically identical to Taylor's (1937), stability chart. The variational approach is, however, powerful enough to allow the analysis of such diverse phenomena as formation of tension cracks (Baker, 1981), effects of non-linear failure criteria (Baker and Frydman, 1983), and three-dimensional effects (Leshchinsky et al., 1985; Baker and Leshchinsky, 1987).

The variational formulation was criticized by De Josseline De Jong (1980, 1981), who argued that this analysis results with a stationary value, which has an indefinite character rather than a minimum. Consequently, he concluded that the variational formulation is, in principle, meaningless, despite its apparent advantages. This conclusion was supported by Castilo and Luceno (1980, 1982). Their argument was based on a number of counter examples showing that for an arbitrary, but given, slip surface; it is possible to establish a normal stress function which yields safety factors that are smaller than the "minimum" obtained by the variational analysis.

In the present work we incorporate some additional physical restrictions into the basic limiting equilibrium framework, and verify that those restrictions guarantee that the slope stability problem has a well-defined solution (minimum). These restrictions are implied, without being explicitly stated, in all practical applications of this methodology, and under usual circumstances they do not change the solution of the problem (they are non-active constraints).

2. The slope stability problem

2.1. Conventions and definitions

Basic elements of a simple slope stability problem are introduced in Fig. 1, which shows a homogeneous slope with zero pore pressures, constant unit weight γ, and a straight face. H and β are the slope's height

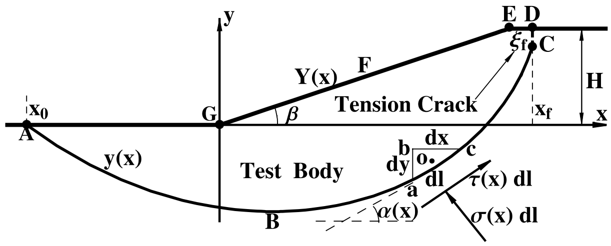

Fig. 1. Basic elements of a schematic slope stability problem.

and inclination respectively. The problem is formulated in a Cartesian coordinate system (x, y) with the y coordinate increasing upwards (against the direction of gravity), the positive x coordinate pointing into the main body of the slope, and the origin located at the toe of the slope (point G). A basic element in limiting equilibrium analysis is a potential test body ABCDEFGA (Fig. 1), bounded by a slip surface $y(x)$ (curve ABC), a vertical face of an end crack CD, and the part DEFGA of the function $Y(x)$ defining the slope's surface. The definition of a test body implies that $x_f > x_0$, $\xi(x) \equiv Y(x) - y(x) \geqslant 0$, and the equality sign may be realized only at the end points $\{x_0, x_f\}$. Under usual conditions $\xi(x_0) = 0$ and $\xi(x_f) \equiv \xi_f \geqslant 0$ where ξ_f is the depth of the end crack. $\alpha(x)$ is the inclination of the slip surface at x, and $\{\sigma(x), \tau(x)\}$ represent the distributions of normal and shear stresses along $y(x)$.

The following sign convention is used for stress components $\{\sigma(x), \tau(x)\}$ acting along $y(x)$:

(a) Positive normal stresses represent compression.
(b) Positive shear stresses produce a counter clockwise rotation about any point O located inside the differential element abc in Fig. 1. This differential element is located inside the test body, and it is attached to the potential slip surface $y(x)$. This is the conventional soil mechanics sign convention used with Mohr circles.

The sign convention for force components and moments is:

(a) A positive force component points to the positive direction of the corresponding coordinate.
(b) All moments are taken about the toe of the slope (point G). Moments are positive if they produce rotation from the positive direction of the x coordinate towards the positive y coordinate.

As their name implies, limiting equilibrium procedures utilize essentially two basic elements:

(i) Equilibrium conditions for a test body.
(ii) A limiting relation between normal and shear stress acting along a given slip surface. This relation introduces also the notion of safety factor with respect to shear strength.

These classical elements are discussed in the following subsections.

2.2. Equilibrium equations

Let $\{\mathbf{T}, \mathbf{N}\}$ be the resultants of the shear and normal stress distributions acting along a particular slip surface $y(x)$. $\{\mathbf{M}_T, \mathbf{M}_N\}$ are the moments of $\mathbf{T}$ and $\mathbf{N}$. $\{\mathbf{T}_H, \mathbf{T}_V\}$ and $\{\mathbf{N}_H, \mathbf{N}_V\}$ are the horizontal and vertical components of $\{\mathbf{T}, \mathbf{N}\}$. $\mathbf{W}$ is the weight of the test body and $\mathbf{M}_W$ is the moment of this weight. $y'(x)$ is the derivative of $y(x)$. Explicit expressions for the above-defined quantities are summarized in Eqs. (1).

$$\mathbf{T}_\mathrm{H} = \int_{x_0}^{x_\mathrm{f}} \tau(x)\,\mathrm{d}x; \quad \mathbf{T}_\mathrm{V} = \int_{x_0}^{x_\mathrm{f}} \tau(x)y'(x)\,\mathrm{d}x; \quad \mathbf{M}_\mathrm{T} = \int_{x_0}^{x_\mathrm{f}} \tau(x)[y'(x)x - y(x)]\,\mathrm{d}x \tag{1.1}$$

$$\mathbf{N}_\mathrm{H} = \int_{x_0}^{x_\mathrm{f}} \sigma(x)y'(x)\,\mathrm{d}x; \quad \mathbf{N}_\mathrm{V} = \int_{x_0}^{x_\mathrm{f}} \sigma(x)\,\mathrm{d}x; \quad \mathbf{M}_\mathrm{N} = \int_{x_0}^{x_\mathrm{f}} \sigma(x)[x + y'(x)y(x)]\,\mathrm{d}x \tag{1.2}$$

$$\mathbf{W} = \gamma\int_{x_0}^{x_\mathrm{f}} \xi(x)\,\mathrm{d}x; \quad \mathbf{M}_\mathrm{W} = \gamma\int_{x_0}^{x_\mathrm{f}} x\xi(x)\,\mathrm{d}x \tag{1.3}$$

Equilibrium conditions for the test body shown in Fig. 1 are:

$$\mathbf{T}_\mathrm{H} = \mathbf{N}_\mathrm{H}; \quad \mathbf{T}_\mathrm{V} = \mathbf{W} - \mathbf{N}_\mathrm{V}; \quad \mathbf{M}_\mathrm{T} = \mathbf{M}_\mathrm{W} - \mathbf{M}_\mathrm{N} \tag{2}$$

Eqs. (1) and (2) are valid only if $y(x)$ is a uni-valued function of x. This restriction excludes from consideration slip surfaces resulting with formation of overhanging cliffs. Formation of overhanging cliffs requires considerable tensile strength which most soils do not possess, and such failure mechanisms are not considered in the present work. Restricting attention to uni-valued slip surfaces implies that $-\pi/2 \leqslant \alpha(x) \leqslant \pi/2$; the limiting relation $\alpha \to -\pi/2$ can be realized only at x_0, and $\alpha \to \pi/2$ can occur only at x_f. The vertical face CD of the end crack (Fig. 1) violates the requirement that $y(x)$ is uni-valued, and therefore this face cannot be considered as a part of the slip surface. The unknown stress functions $\{\sigma(x), \tau(x)\}$ are defined only along $y(x)$, and the face CD of the end crack is assumed to be stress free. In principle it is possible to impose any stress distribution along CD (e.g. the effect of water standing in the crack), but these stresses must be known a priori, and they are not included in the unknown functions $\{\sigma(x), \tau(x)\}$.

2.3. The limiting equilibrium hypothesis and its complementary interpretation

The equilibrium conditions in Eq. (2) include three unknown functions $\{y(x), \sigma(x), \tau(x)\}$. All limiting equilibrium procedures eliminate the function $\tau(x)$ by the assumption:

$$\tau(x) \equiv \frac{S(x)}{F} = \frac{S[\sigma(x)]}{F} \tag{3}$$

In this equation F is an unknown number called the safety factor with respect to shear strength, $S(\sigma)$ is a strength function specifying the dependence of shear strength on normal (in general effective) stresses, and $S(x) \equiv S[\sigma(x)]$ represents the distribution of available shear strength along $y(x)$. Physically $S(\sigma)$ can be identified with the upper branch of a (generally non-linear) Mohr envelope. This identification implies that $S(\sigma)$ is a non-negative function satisfying $S(\sigma) \geqslant 0$. The safety factor is a measure of stability, with $F = 1$ implying a state of failure in which existing shear stresses are equal to the available shear strength. Most practical limiting equilibrium procedures utilize the linear Mohr–Coulomb failure criterion, $S(\sigma) \equiv C + \psi\sigma$ where $\psi = \tan(\phi)$, and $\{C, \phi\}$ are the cohesion intercept and angle of internal friction respectively. This criterion implies the well-known restriction $\sigma(x) \geqslant -t^*$ where $t^* = C/\psi$. In the legitimate range of σ values ($\sigma \geqslant -t^*$), both $S(\sigma)$ and $S(x) = S[\sigma(x)]$ are obviously non-negative.

In addition to its role as a definition of the safety factor with respect to strength, Eq. (3) implies that the distributions of shear stress $\tau(x)$ and available shear strength $S(x)$ are geometrically similar, being "scaled versions" of one another. In particular, combining the observation that $S(x)$ is non-negative and the fact that F is a constant independent of x, Eq. (3) implies that $\tau(x)$ has a constant sign along the entire slip surface. For positive F values Eq. (3) results with $\tau(x) \geqslant 0$. In that case the sign convention with respect to shear stresses implies that shear stresses along $y(x)$ are directed from A to C (Fig. 1). The assumption that $\tau(x)$ and $S(x)$ are geometrically similar is physically justified only at failure when $\tau(x) = S(x)$ and $F = 1$.

Nevertheless, limiting equilibrium procedures which are applied to both stable and unstable situations are all based on Eq. (3). The significance of this observation may be seen by rewriting Eq. (3) in the form:

$$S_{\mathrm{m}}(\sigma) \equiv S(\sigma)/F \tag{4.1}$$

$$\tau(x) = S_{\mathrm{m}}[\sigma(x)] \tag{4.2}$$

Eq. (4.1) defines a "mobilized" strength function $S_{\mathrm{m}}(\sigma) = C_{\mathrm{m}} + \sigma\psi_{\mathrm{m}}$, in which $C_{\mathrm{m}} = C/F$ and $\psi_{\mathrm{m}} = \tan(\phi_{\mathrm{m}}) = \psi/F$ are mobilized strength parameters. It is possible to consider $S_{\mathrm{m}}(\sigma)$ as a strength function characterizing a fictitious material with reduced strength. In the framework of this interpretation Eq. (4.2) is a failure condition for this fictitious material. Eqs. (4) indicate that limiting equilibrium procedures deal actually with the state of failure of a fictitious material characterized by the mobilized strength function $S_{\mathrm{m}}(\sigma)$ rather than with the real problem which is characterized by the actual strength function $S(\sigma)$, and does not necessarily corresponds to a failure state. One may legitimately wonder if such an artificial problem has relevance to real life engineering situations (the long and successful history of this methodology implies that this is so). However, this is not the issue in the present work; we accept this situation as a "given"; and will utilize its consequences. For ease of reference we call the interpretation associated with Eqs. (4) "the complementary interpretation". The following observations are relevant with respect to this interpretation:

(1) Eqs. (3) and (4) are formally equivalent, and the complementary interpretation does not change the basic assumption of limiting equilibrium.
(2) The basic assumption of limiting equilibrium (Eq. (3)), relates only two stress components $\{\sigma, \tau\}$. Knowledge of these components does not allow definition of the state of stress (Mohr circle), at different points along $y(x)$.
(3) The complementary interpretation implies that Mohr circles are tangential to the mobilized strength envelope $S_{\mathrm{m}}(\sigma)$. Combined with given values of $\{\sigma, \tau\}$, this tangency requirement is sufficient to define the complete state of stress along potential slip surfaces. As a result, the complementary interpretation allows incorporation of various physical considerations into the formal limiting equilibrium framework (e.g. Baker et al., 1993), and it will play a central role in the present work.

The notions of mobilized strength parameters and envelopes are well known; however the implications of the complementary interpretation were not fully utilized in previous formulations of limiting equilibrium slope stability analysis.

2.4. The limiting equilibrium equations

Combining Eqs. (2) and (3) results with:

$$\mathbf{T}_{\mathrm{H}} = \frac{\overline{\mathbf{T}}_{\mathrm{H}}}{F} = \mathbf{N}_{\mathrm{H}}; \quad \mathbf{T}_{\mathrm{V}} = \frac{\overline{\mathbf{T}}_{\mathrm{V}}}{F} = \mathbf{W} - \mathbf{N}_{\mathrm{V}}; \quad \mathbf{M}_{\mathrm{T}} = \frac{\overline{\mathbf{M}}_{\mathrm{T}}}{F} = \mathbf{M}_{\mathrm{W}} - \mathbf{M}_{\mathrm{N}} \tag{5.1}$$

where

$$\overline{\mathbf{T}}_{\mathrm{H}} = \int_{x_0}^{x_{\mathrm{f}}} S[\sigma(x)]\,\mathrm{d}x; \quad \overline{\mathbf{T}}_{\mathrm{V}} = \int_{x_0}^{x_{\mathrm{f}}} S[\sigma(x)]y'(x)\,\mathrm{d}x; \quad \overline{\mathbf{M}}_{\mathrm{T}} = \int_{x_0}^{x_{\mathrm{f}}} S[\sigma(x)][y'(x)x - y(x)]\,\mathrm{d}x \tag{5.2}$$

Eq. (5.1) combine the two basic elements utilized in limiting equilibrium slope stability analysis (equilibrium equations for a test body, and definition of safety factor), and they will be referred to as the basic equations of limiting equilibrium. These equations depend on two unknown functions $\{y(x), \sigma(x)\}$ and the unknown constant F. Recalling that strength functions are non-negative, the first of Eq. (5.2) implies $\overline{\mathbf{T}}_{\mathrm{H}} \geqslant 0$.

Experience with conventional limit equilibrium analysis shows that there exist many different triplets $\{y(x), \sigma(x), F\}$ satisfying Eq. (5.1). Let $\{\tilde{y}(x), \tilde{\sigma}(x)\}$ be a pair of functions satisfying those equations for some value of F. We call such a pair a potential (or legitimate), failure mechanism. The basic slope stability problem is to find the critical pair $\{\tilde{y}_c(x), \tilde{\sigma}_c(x)\}$ which is associated with the minimum value, F_s, of F (assuming that this problem has a solution, i.e. that a minimum exists). Stated differently, the equations of limiting equilibrium associate an F value with each pair of legitimate functions $\{\tilde{y}(x), \tilde{\sigma}(x)\}$; i.e. they define a functional relation of the type $F = \widehat{F}[\tilde{y}(x), \tilde{\sigma}(x)]$. We call $\widehat{F}[\tilde{y}(x), \tilde{\sigma}(x)]$ the safety functional. It is not possible to establish an explicit expression for this functional, however, for the present purpose, the important point is only that this functional exists.

The discussion up to this point involved classical elements only, common to all existing limiting equilibrium procedures. Conventional procedures introduce also various geometrical and statical assumptions that are not relevant for the present purpose. In the following sections we introduce two additional elements (the classification inequalities, and the cracking hypothesis), which are assumed, but not explicitly stated, in all applications of limiting equilibrium slope stability analysis.

3. The classification inequalities

Consider the simple unloaded slope stability problem shown in Fig. 2a. It is natural to assume that at failure (of the artificial material defined by mobilized strength function $S_m(\sigma)$), a certain test body moves down (in the direction of gravity), and to the left (away from the main body of the slope). Eqs. (5) show that the forces $\{\mathbf{T}_H, \mathbf{T}_V\}$ are proportional to the mobilized strength $S_m(x)$. Strength is mobilized in response to movement. Therefore the forces $\{\mathbf{T}_H, \mathbf{T}_V\}$ should be directed counter to directions of the expected movements as shown in Fig. 2a. The present sign convention for force components implies therefore that $\{\mathbf{T}_H, \mathbf{T}_V\}$ should be non-negative, and it is possible to write:

$$\mathbf{T}_H = \int_{x_0}^{x_f} S_m[\sigma(x)]\,dx \geqslant 0; \quad \mathbf{T}_V = \int_{x_0}^{x_f} S_m[\sigma(x)]y'(x)\,dx \geqslant 0 \tag{6}$$

The following comments are appropriate with respect to Eq. (6):

(a) Strictly speaking the considerations leading to Eq. (6) are valid only at failure, while limiting equilibrium analysis is applied to both stable and unstable configurations. However, the complementary interpretation implies that limiting equilibrium analysis actually deals with a state of failure of an artificial material, and within the framework of this interpretation the above inequalities are valid for all F values.

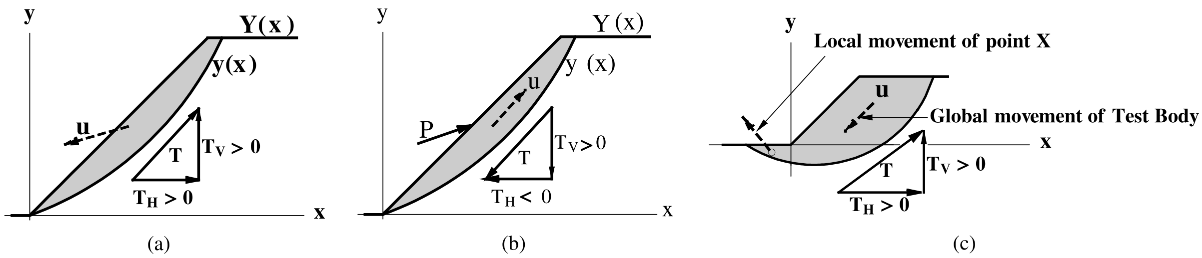

Fig. 2. Modes of failure: (a) active problem; (b) passive problem and (c) deep slip surface.

(b) Eq. (6) involve integrals of $S_m(x)$, i.e. these equations represent global restrictions. The global nature of these relations implies that the assumptions with respect to directions of movements refer to global or "average" movements of the test body. Individual points in this body may move in different directions (e.g. point X in Fig. 2c probably moves upward, opposite to the direction of the global vertical movement).
(c) We are not aware of a single practical solution of a slope stability problem violating inequalities (6). In the present work those inequalities were motivated by considerations related to movements. However, in general, limiting equilibrium analysis does not deal with movements or deformations, and from now on we will consider Eq. (6) as postulates, justified only by their consequences.

Applying a sufficiently large external force **P** to the face of the slope, it is possible to create a "passive" condition in which failure occurs with a test body moving upwards and into the slope (Fig. 2b). In passive problems the resultants $\{\mathbf{T}_H, \mathbf{T}_V\}$ must point downwards and outwards. In terms of the present sign convention, such forces are negative, and passive problems are characterized by $\{\mathbf{T}_H \leqslant 0, \mathbf{T}_V \leqslant 0\}$. Conventional slope stability problems have an "active" character of the type shown in Fig. 2a. Both active and passive problems are obviously legitimate, but since a test body cannot move up and down at the same time, these two problems can not be considered simultaneously. Thus, inequalities specifying the sign of the terms $\{\mathbf{T}_H, \mathbf{T}_V\}$ define the type of problem under consideration, i.e. these inequalities classify stability problems, and they will be called "the classification inequalities". Mixed problems, in which $\{\mathbf{T}_H \leqslant 0, \mathbf{T}_V \geqslant 0\}$ or $\{\mathbf{T}_H \geqslant 0, \mathbf{T}_V \leqslant 0\}$ are in principle possible, but their physical significance is not obvious, and they will not be considered in the present work.

So far we have discussed active and passive stability problems; however the same consideration can be applied to each legitimate pair $\{\tilde{y}(x), \tilde{\sigma}(x)\}$ which satisfies the equations of limiting equilibrium (5). A given pair $\{\tilde{y}(x), \tilde{\sigma}(x)\}$ determines a unique value for $\{\mathbf{T}_H, \mathbf{T}_V\}$, and according to the sign of these quantities, such a pair describes active, passive, or mixed failure mechanism. In general, the class of legitimate functions $\{\tilde{y}(x), \tilde{\sigma}(x)\}$ contains members associated with all four possible combinations for the signs of $\{\mathbf{T}_H, \mathbf{T}_V\}$. However, when solving an active problem one needs to consider only active failure mechanisms, and Eq. (6) become integral inequality constraints that exclude from consideration failure mechanisms which are not relevant to solution of an active problem.

The present work deals mainly with conventional, "active" slope stability problems. It is instructive however to discuss, briefly, some characteristics of the passive case. The first of Eq. (5.1) can be written as $F = \overline{\mathbf{T}}_H/\mathbf{T}_H$. When discussing Eqs. (5) it was established that $\overline{\mathbf{T}}_H \geqslant 0$. As a result, the only way to obtain $\mathbf{T}_H \leqslant 0$ (as required by the definition of passive problems), is to admit a negative safety factor. This apparently strange result is a consequence of the presently adopted sign convention with respect to forces and shear stresses. When solving a passive problem it is obviously more convenient to change the sign convention in order to deal with positive safety factors. However, at this stage we are dealing with the general structure of limiting equilibrium problems, and in order to investigate this structure it is necessary to adhere to a single sign convention under all situations. The sign of the safety factor is related to the direction of shear stresses along the slip surface; for negative F values Eq. (3) delivers negative shear stresses, and the sign convention implies that these stresses are directed from C to A (Fig. 1) as must be the case in a passive problem. A large negative value of F means a high level of safety for a passive failure mechanism. Consequently, identification of the critical condition for a passive problem requires maximization (rather than minimization), of the safety functional. This result is consistent with the situation encountered in earth pressure theory in which active and passive problems are associated with different types of extremization.

In order to illustrate the significance of the classification inequalities in a simple setting, consider the problem in Fig. 3a. This figure shows a slope loaded by a single horizontal force $\mathbf{P}_H$, acting on the slope's surface. For the present purpose, it is convenient to consider a class of simple failure mechanisms with $y(x)$ taken as straight lines through the toe, and $\sigma(x)$ as the triangular distribution ABC shown at the bottom of

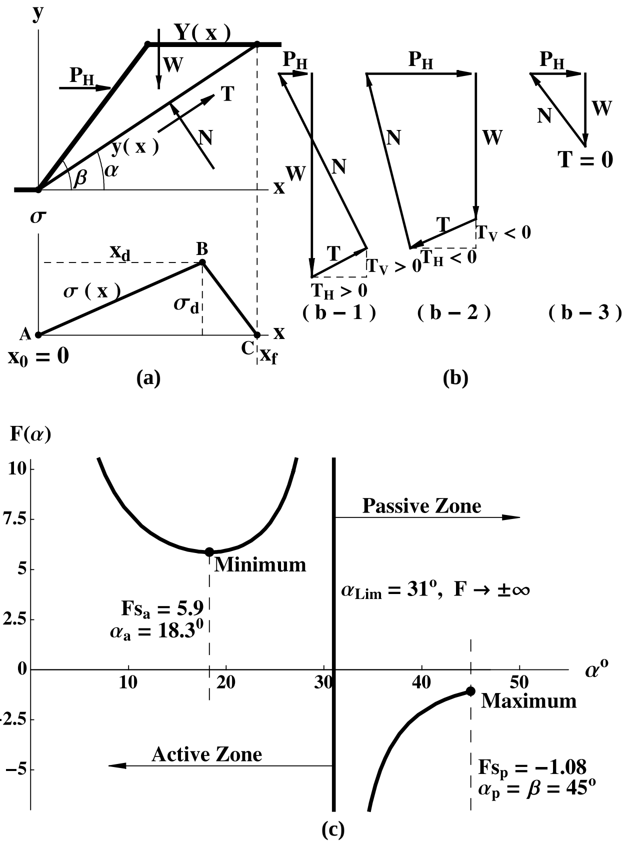

Fig. 3. Active and passive solutions: (a) failure mechanisms; (b) force polygons and (c) the stability function $F(\alpha)$.

Fig. 3a. The functions $\sigma(x)$ are defined in terms of two undetermined parameters σ_d and x_d representing the magnitude and location of the maximum normal stress acting on the slip surface. For each inclination α of the slip surface it is possible to solve the three equations of limiting equilibrium (5), and establish the unknown values $\{F, \sigma_d, x_d\}$. As a result, each value of α in the range $\beta \geqslant \alpha \geqslant 0$ defines a pair of legitimate functions $\{\tilde{y}(x), \tilde{\sigma}(x)\}$ satisfying the equations of limiting equilibrium, and the safety functional $\widehat{F}[\tilde{y}(x), \tilde{\sigma}(x)]$ degenerates into a one dimensional function $F(\alpha)$. The class of failure mechanisms obtained this way is probably not critical, but it is legitimate, being included in $\{\tilde{y}(x), \tilde{\sigma}(x)\}$.

Solving the equations of limiting equilibrium for F using α as a parameter it is not difficult to verify that the function $F(\alpha)$ is given by:

$$F(\alpha) = \frac{Cx_f(\alpha)(1 + \tan^2(\alpha)) + \psi[\mathbf{W}(\alpha) + \mathbf{P}_H \tan(\alpha)]}{\mathbf{W}(\alpha)\tan(\alpha) - \mathbf{P}_H} \tag{7.1}$$

where

$$x_f(\alpha) = \frac{H}{\tan(\alpha)} \quad \text{and} \quad \mathbf{W}(\alpha) = \frac{\gamma H^2}{2}\left(\frac{1}{\tan(\alpha)} - \frac{1}{\tan(\beta)}\right) \tag{7.2}$$

Fig. 3c is a plot of $F(\alpha)$ evaluated for $C = 10$ kPa, $\phi = 30°$, $\gamma = 20$ kN/m^3, $H = 10$ m, $\beta = 45°$ and a horizontal external line force, $\mathbf{P}_H$ of 400 kN/m applied at $H/3$. It is possible to verify that for this input information each value of α results with a legitimate normal stress distribution, satisfying $\sigma_d \geqslant 0$ and $x_f \geqslant x_d \geqslant 0$. For the present purpose, the significant feature of Fig. 3c is the discontinuous nature of the function $F(\alpha)$, which at $\alpha \cong 31°$ approaches $\pm\infty$. F is positive in the range $0 \leqslant \alpha \leqslant 31°$, and the force polygon (b1) in Fig. 3b shows that both $\mathbf{T}_H$ and $\mathbf{T}_V$ are positive. Consequently, α values in the range $0 \leqslant \alpha \leqslant 31°$ are associated with active failure mechanisms. The function $F(\alpha)$ delivers negative F values in the range $31° \leqslant \alpha \leqslant \beta$, and the force polygon (b2) shows that this range corresponds to passive failure. The value $\alpha \cong 31°$, at which the function $F(\alpha)$ is discontinuous, is the boundary between active and passive failure mechanisms. The force polygon (b3) shows that at this boundary $\mathbf{T}$ (and therefore also $\{\mathbf{T}_H, \mathbf{T}_V\}$), are equal to zero, i.e. the discontinuity in Fig. 3c is associated with the boundary of the classification inequalities (6). The discontinuity at $\alpha \cong 31°$ is a consequence of the fact that safety factors are ratios of "resisting" to "driving" forces, and such ratios tend to $\pm\infty$ when "driving" forces are equal to zero. The notions of "resisting" and "driving" forces cannot be quantified (every force has both driving and stabilizing effects), and these notions are used here only in order to explain the form of the function $F(\alpha)$.

Failure mechanisms defined in Fig. 3a are legitimate, (i.e. they belong to the class of legitimate functions $\{\tilde{y}(x), \tilde{\sigma}(x)\}$ satisfying the equations of limiting equilibrium (5)). Consequently, this figure shows that by itself the safety functional is unbounded (ranging from plus to minus infinity), and it does not have a global minimum or maximum. However; restricting attention to active stability problems, Fig. 3c shows that the function $F(\alpha)$ has a well-defined local minimum, occurring at a stationary point of this function. Similarly, restricting attention to passive problems, the function $F(\alpha)$ has a maximum (in the limited class of failure mechanisms defined in Fig. 3a this maximum occurs at the boundary $\alpha = \beta$, not at a stationary point, but this is not essential). Fig. 3 illustrates why it is necessary to use the classification inequalities in order to distinguish between active and passive stability problems. Without these inequalities, even the degenerate form $F(\alpha)$ of the safety functional $\widehat{F}[\tilde{y}(x), \tilde{\sigma}(x)]$ is unbounded; it does not have a minimum; so formally the limiting equilibrium problem does not have a solution. This discontinuity of $\widehat{F}[\tilde{y}(x), \tilde{\sigma}(x)]$ was not considered in previous formulations of the problem, and it is responsible for most of the confusion related to the conceptual validity of the variational approach.

The discontinuity of $\widehat{F}[\tilde{y}(x), \tilde{\sigma}(x)]$ is not important for practical applications of the limiting equilibrium approach since it is associated with $F \to \pm\infty$, which corresponds to a state of absolute safety. However, the indefinite results obtained by De Josseline De Jong (1980) are probably related to the fact that without the restrictions provided by the classification inequalities, the safety functional $\widehat{F}[\tilde{y}(x), \tilde{\sigma}(x)]$ is unbounded, and formally, the slope stability problem does not have a solution (minimum). Fig. 3c shows that the classification inequalities are essential for a proper formal definition of the slope stability problem, but they have no effect on the solution point (minimum), of this problem (they define merely the region in which the solution exists), i.e. the classification inequalities are non-active constraints. In the simplified setting considered in Fig. 3 the classification inequalities resulted in the restriction $0 \leqslant \alpha \leqslant 31°$ defining a family of legitimate slip surfaces for the active stability problem. In general however these inequalities constitute a restriction on admissible pairs $\{y(x), \sigma(x)\}$, not only $y(x)$.

In the following we restrict attention to conventional (active), stability problems. Considering this class of problems and combining the equation of horizontal equilibrium (first of Eq. (5.1)) with Eq. (6.1) implies $F \geqslant 0$. Consequently, for active problems the classification inequalities induce a zero lower bound on F values.

4. Tensile strength and the cracking hypothesis

The Mohr–Coulomb failure criterion implies the well-known restriction $\sigma \geqslant -t^*$ where $t^* = C/\tan(\phi)$. It is natural to interpret t^* as the tensile strength implied by this criterion. The definition of mobilized strength envelopes implies that $t^* = C/\psi = C_m/\psi_m$, i.e. all the "fictitious materials" associated with

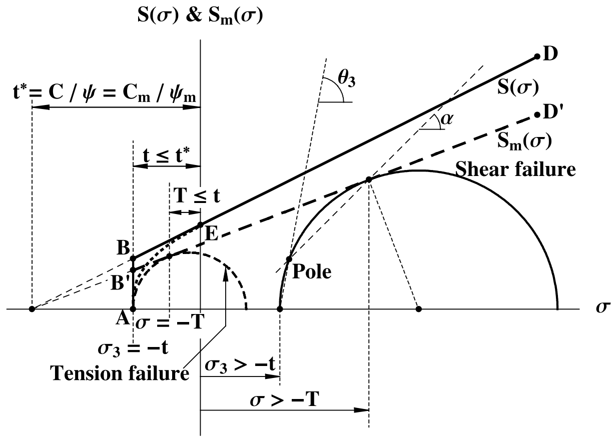

Fig. 4. State of stress along potential slip surfaces.

different F values have the same tensile strength t^*. The actual tensile strength "t" of real soils is usually much smaller then the t^* value implied by the Mohr–Coulomb strength function, and in the tensile stresses zone strength functions (Mohr envelopes) of real soils are highly non-linear having the character of the doted line AE in Fig. 4. The restricted Mohr-Coulomb strength function (line ABED in Fig. 4), provides a first order approximation of this non-linear behavior. This strength function is defined as $S(\sigma) = C + \sigma\psi$, $\sigma \geqslant -t$, where $0 \leqslant t \leqslant t^*$.

Using the complimentary interpretation it is possible to establish the complete state of stress, at each point along the slip surface as shown in Fig. 4. In particular, it is possible to establish the magnitude of the minor principal stress σ_3 and the direction of the minor principal plane θ_3 as shown in that figure. Considering the geometry of the Mohr circle in Fig. 4 it is not difficult to verify that:

$$\sigma_3 = \frac{1 - \sin(\phi_{\mathrm{m}})}{\cos^2(\phi_{\mathrm{m}})}(\sigma - C_{\mathrm{m}} \cos(\phi_{\mathrm{m}})) \tag{8}$$

The definition of tensile strength implies that all normal stresses, including σ_3, must be larger than $(-t)$. Combining this requirement with Eq. (8) results in:

$$\sigma(x) \geqslant -T \tag{9.1}$$

where

$$T = T(C, \phi, t, F) = \cos(\phi_{\mathrm{m}})\left(t\frac{\cos(\phi_{\mathrm{m}})}{1 - \sin(\phi_{\mathrm{m}})} - C_{\mathrm{m}}\right) \tag{9.2}$$

The following comments are relevant with respect to Eqs. (9):

1. A restricted Mohr–Coulomb strength function is defined in terms of three independent strength parameters $\{C, \phi, t\}$.
2. Fig. 4 provides a clear physical interpretation of $\{t, t^*$, and $T\}$; t is the soil's tensile strength, t^* is the tensile strength implied by the conventional (unrestricted), Mohr–Coulomb envelope; while $(-T)$ is the magnitude on the normal stress acting on the slip surface when $\sigma_3 = -t$. It is noted that t is a material property (a given constant), while T depends on F and the three strength parameters $\{C, \phi, t\}$.

3. The geometry of Mohr circles in Fig. 4, and the definition of mobilized strength envelopes, imply that $t^* \geqslant t \geqslant T$, and $T = t$ only if $t = t^*$. In that case Eq. (9.1) is reduced to the classical relation $\sigma \geqslant -t^* = -C/\psi$.
4. The setting $t = t^*$ is obviously illegitimate in the limiting case of $\phi = 0$; implying an infinite tensile strength, which is definitely an unreasonable proposition for a particulate media like soil. For frictionless material the inequality $\sigma(x) \geqslant -t^*$ is not restrictive, allowing normal stresses to reach minus infinity. However, when $\phi = 0$, Eqs. (9) are reduced to $\sigma(x) \geqslant -T = C/F - t$, and this bound is restrictive (finite), for all legitimate values of $\{t, C, F\}$.

In the general $\phi \neq 0$ case the setting $t = t^*$ is legitimate (although not necessarily realistic), and we will use this setting. t^* approaches infinity in the particular case of $\phi \equiv 0$, and in that case t must be considered as an independent strength parameter in order to exclude from consideration the unrealistic situation involving soils (a particulate media), with infinite tensile strength. The advantage of the general representation (9) is that it is valid for all legitimate t values regardless of the magnitude of ϕ.

The limiting case $\sigma = -T$ implies full mobilization of tensile strength, and it is necessary to specify the implications of this particular physical state. In the present work, it is assumed that satisfaction of the limiting relation $\sigma(x) = -T$ results with formation of a tension crack at x, extending from the slip surface to the surface of the slope. This assumption will be referred to as the cracking hypothesis.

The cracking hypothesis implies that satisfaction of the limiting condition $\sigma(x) = -T$ at some internal point $x_0 < x < x_f$ results with an internal tension crack L'CL as shown in Fig. 5. This internal crack separates the test body into two independent parts ABCLEGA and CDL'C which interact with each other only at a single point. The limiting equilibrium approach is based on consideration related to a single test body, and the two test bodies in Fig. 5 cannot be analyzed simultaneously (within the limiting equilibrium framework each one of these test bodies may have a different safety factor). Consequently, normal stress distributions in which the relation $\sigma(x) = -T$ is satisfied at an internal point are not legitimate and they must be excluded from consideration. It is noted that the above argument does not exclude solutions with multiple tension cracks; however, in the limiting equilibrium framework such solutions are obtained by solving a sequence of stability problems, each one of which consists of a single test body.

Based on the above considerations $\sigma = -T$ can be realized only at the end points x_0 or x_f. Restricting attention to active failure mechanisms excludes the possibility of tension cracks at x_0 resulting with $\sigma(x_0) > -T$ and $y(x_0) = Y(x_0)$ (active failure mechanisms imply that the test body moves away from the main body of the slope, and this movement results with closure of a crack located at x_0). Therefore $\sigma = -T$

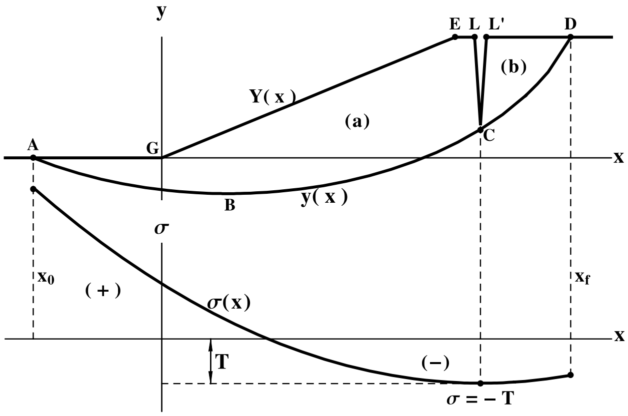

Fig. 5. Internal cracks.

can be satisfied only at x_f. Combination of the cracking hypothesis and the classification inequalities implies that in active stability problems $y(x_f) > y(x_0)$ (if $y(x_f) < y(x_0)$ then the test body will move towards the crack; resulting with a passive problem, and closure of that crack). This restriction excludes from consideration tensile cracks which are deeper than the total height of the slope. For the present purpose (derivation of sufficient conditions for existence of a minimum in slope stability problems), the most important implication of the cracking hypothesis is that it allows the condition $\sigma = -T$ to occur only at (x_f).

In order to incorporate the above considerations into the general framework, it is convenient to consider, simultaneously, the system of inequalities $y(x) \leqslant Y(x)$, $\sigma(x) \geqslant -T$, and $-\pi/2 < \alpha(x) < \pi/2$ which follow from the geometrical definition of the test body, existence of tensile strength, and the requirement that the slip surface is a uni-valued function of x. Specifying these inequalities at x_f, results with the following boundary condition which must be satisfied at the end point of the slip surface.

$$\text{At } x_f \begin{cases} y(x_f) = Y(x_f) \quad \text{and} \quad \sigma(x_f) \geqslant -T \quad \text{and} \quad \alpha(x_f) < \pi/2 \\ \text{or} \\ \sigma(x_f) = -T \quad \text{and} \quad y(x_f) \leqslant Y(x_f) \quad \text{and} \quad \alpha(x_f) < \pi/2 \\ \text{or} \\ \alpha(x_f) \to \pi/2 \quad \text{and} \quad \sigma(x_f) \geqslant -T \quad \text{and} \quad y(x_f) \leqslant Y(x_f) \end{cases} \tag{10}$$

Eq. (10) will be referred to as the cracking criterion. The first option in this equation corresponds to the conventionally considered case where there is no end crack. In order for this situation to exist, it is necessary that the normal stress at x_f is not less then the limiting value $-T$, and the slip surface does not start forming an overhanging cliff. The second and third options in Eq. (10) show that an end crack having a non-zero depth will appear if the tensile strength is fully mobilized ($\sigma = -T$) or if the slip surface approaches a vertical tangent ($\alpha_f \to \pi/2$). An earlier form of this criterion was used by Baker (1981).

The following implications of Eq. (10) are noted:

(a) Despite its complex appearance, the boundary condition given in Eq. (10) fits naturally into the general solution procedure implied by the variational slope stability analysis (Baker and Garber, 1978; Baker, 1981). In those works, it is shown that potentially critical functions $\{y(x), \sigma(x)\}$ are solutions of two first order non-linear differential equations (Euler's equations). Starting the integration process at some x_0, using the initial conditions $\{y(x_0) = Y(x_0), \sigma(x_0) = \sigma_0 > -T\}$, the integration process is continued until one of the three options in Eq. (10) is satisfied. In other words, Eq. (10) is a natural "stopping rule" for integration of Euler's differential equations. Satisfaction of Eq. (10) defines the location of x_f and the magnitude ξ_f of the end crack. Consequently, ξ_f cannot be assumed a priori, as is commonly done in conventional slope stability analysis. This solution process automatically eliminates from considerations normal stress distributions in which the limiting condition $\sigma = -T$ is realize at an internal point, and $\sigma = -T$ can be realized only at x_f. In other words the cracking criterion provides the formal mechanism ensuring satisfaction of the cracking hypothesis and its implications.

(b) The above considerations implies that $\sigma(x) > -T$ at all internal points, and $\sigma = -T$ can be (but not necessarily is), realized only at x_f. These are restrictions defining a class of admissible normal stress functions. Normal stress functions violating those restrictions correspond to non-physical situations, which are excluded from considerations by the cracking criterion.

(c) The cracking hypothesis excludes internal cracks by allowing the condition $\sigma = -T$ to occur only at a single point. Consider however the limiting case of surface failure in which $y(x) = Y(x)$ at every point. In that case, the depth of cracks is zero (i.e. they do not exist); the restriction to a single crack is not significant; and the cracking hypothesis losses its restrictive power. It is well known that the critical conditions in $C = 0$ materials are realized for $y(x) = Y(x)$, and this limiting case will be considered explicitly later on.

5. The general (variational), slope stability problem

The safety functional $\widehat{F}[\tilde{y}(x), \tilde{\sigma}(x)]$ is an abstract conceptual relation, without explicit representation, and it cannot be minimized directly. However, the three equations of limiting equilibrium (5.1), are linear in $(1/F)$; it is possible, therefore, to solve one of these equations for F (thus obtaining an explicit representation for F in terms of $\{y(x), \sigma(x)\}$); considering the remaining two equilibrium conditions as integral constraints. The basic idea of considering equilibrium equations as constraints in limiting equilibrium analysis is essentially due to Kopacsy (1955, 1957, 1961). However, Kopacsy tried to formulate a slope stability problem without using the notion of safety factors, and for no obvious reason he chose to minimize the weight $\mathbf{W}$ of the test body subject to the requirement that the three equilibrium equations are satisfied. Such a formulation has no obvious physical justification. The process of defining a minimization criterion by solving one equilibrium equation, treating the other two as constraints, was introduced by Baker and Garber (1977, 1978).

In the present work we use the equation of horizontal equilibrium (first of Eq. (5.1)) in order to define the minimization criterion $\widetilde{F}[y(x), \sigma(x)] \equiv \overline{\mathbf{T}}_{\mathrm{H}}/\mathbf{N}_{\mathrm{H}}$; considering the remaining two elements of (5.1) as integral constraints. This choice is arbitrary, and the same final solution would be obtained using any one of Eq. (5.1) for the definition of $\widetilde{F}[y(x), \sigma(x)]$, treating the remaining two equilibrium conditions as constraints. This is a consequence of the isoperimetric theorem of variational calculus (Petrov, 1968). Obviously $\widetilde{F}[y(x), \sigma(x)]$ and $\widehat{F}[\tilde{y}(x), \tilde{\sigma}(x)]$ are different functionals.

Combining results, the general (variational) slope stability problem is defined by the following relations:

$$F_{\mathrm{s}} = \min_{\{y(x),\sigma(x)\}} \{\widetilde{F}[y(x), \sigma(x)]\} = \widetilde{F}[y_{\mathrm{c}}(x), \sigma_{\mathrm{c}}(x)] \tag{11.1}$$

where

$$F = \widetilde{F}[y(x), \sigma(x)] \equiv \overline{\mathbf{T}}_{\mathrm{H}}[y(x), \sigma(x)]/\mathbf{N}_{\mathrm{H}}[y(x), \sigma(x)] \tag{11.2}$$

subject to satisfaction of the following system of constraints:

$$\overline{\mathbf{T}}_{\mathrm{V}} = F(\mathbf{W} - \mathbf{N}_{\mathrm{V}}); \quad \overline{\mathbf{M}}_{\mathrm{T}} = F(\mathbf{M}_{\mathrm{W}} - \mathbf{M}_{\mathrm{N}}) \tag{11.3}$$

$$\mathbf{T}_{\mathrm{H}} \geqslant 0; \quad \mathbf{T}_{\mathrm{V}} \geqslant 0 \tag{11.4}$$

At all internal points

$$y(x) < Y(x); \quad \sigma(x) > -T; \quad -\pi/2 < \alpha(x) < \pi/2 \tag{11.5}$$

Boundary conditions:

$$x_{\mathrm{f}} > x_0; \quad \sigma(x_0) > -T; \quad y(x_0) = Y(x_0); \quad y(x_{\mathrm{f}}) > y(x_0) \tag{11.6}$$

$$\text{At } x_{\mathrm{f}} \begin{cases} y(x_{\mathrm{f}}) = Y(x_{\mathrm{f}}) \quad \text{and} \quad \sigma(x_{\mathrm{f}}) \geqslant T \quad \text{and} \quad \alpha(x_{\mathrm{f}}) < \pi/2 \\ \text{or} \\ \sigma(x_{\mathrm{f}}) = T \quad \text{and} \quad y(x_{\mathrm{f}}) \leqslant Y(x_{\mathrm{f}}) \quad \text{and} \quad \alpha(x_{\mathrm{f}}) < \pi/2 \\ \text{or} \\ \alpha(x_{\mathrm{f}}) \to \pi/2 \quad \text{and} \quad \sigma(x_{\mathrm{f}}) \geqslant T \quad \text{and} \quad y(x_{\mathrm{f}}) \leqslant Y(x_{\mathrm{f}}) \end{cases} \tag{11.7}$$

It is convenient to review, briefly, the various elements in Eqs. (11). Eq. (11.1) defines the general minimization framework characterizing limiting equilibrium problems. The triplet $\{y_{\mathrm{c}}(x), \sigma_{\mathrm{c}}(x), F_{\mathrm{s}}\}$ consisting of the critical slip surface, normal stress functions and the minimal safety factor is the solution "point" of the above problem. The explicit form of the minimization criterion $\widetilde{F}[y(x), \sigma(x)]$ is given in Eq. (11.2). This form is obtained by solving the equation of horizontal equilibrium for F, and this guarantees satisfaction of horizontal equilibrium. Eq. (11.3) are integral constraints expressing the requirements of vertical and moment equilibrium respectively. Eq. (11.4) are the classification inequalities that define the conventional, active, slope stability problem. Eq. (11.5) are consistency requirements which follow from the definition of a test body, existence of tensile strength, and the requirement that the slip surface is a uni-valued function of x. The first of Eq. (11.6) excludes from consideration degenerated failure mechanisms consisting of a single point. The second and third of these equations prevent formation of tension cracks at x_0. The fourth of these equations excludes cracks which are deeper than the total height of the slope. Eq. (11.7) is the cracking criterion discussed previously. Not all of the requirements specified in Eqs. (11) are independent, but such redundancy does not affect the following arguments.

Eqs. (11) are a general definition of the slope stability problem. Those equations do not include arbitrary assumptions with respect to the functions $\{y(x), \sigma(x)\}$. All restrictions imposed on those functions follow from the classification inequalities and the cracking hypothesis which are well motivated physically. It is noted that all other existing limiting equilibrium formulations employ a priori assumptions restricting the class of functions $\{y(x), \sigma(x)\}$ considered in the minimization stage specified in Eq. (11.1). By their very nature such restrictions are un-conservative, leading to overestimation of minimal safety factors, and they should be avoided in order to guarantee safe design.

Faced with a minimization problem including inequalities, one always has two options:

(a) Incorporate these inequalities into the solution process, thus ensuring that the solution triplet $\{y_c(x), \sigma_c(x), F_s\}$ is legitimate. This approach increases the dimensionality of the minimization problem; (inequality constraints are incorporated into analysis by defining slack variables with respect to which it is necessary to minimize). Therefore this type of formulation may become quite awkward, but sometimes this is unavoidable.
(b) Solve the problem without the inequality constraints, and check if the resulting solution satisfies them automatically. If the solution satisfies constraints that were not enforced, then these constraints are "non-active". Non-active constraints have no effect on results of the analysis, and this justifies their omission in the solution stage.

Let $\{y^*(x), \sigma^*(x)\}$ be the set of all pairs $\{y(x), \sigma(x)\}$ for which $\mathbf{T}_H = \mathbf{T}_V = 0$. This set represents failure mechanisms located on the boundary (in functions space), of the classification inequalities. Recalling that in active problems the shear stresses $\tau(x)$ are positive along the entire slip surface; it is not difficult to realize that the case $\mathbf{T}_H = \mathbf{T}_V = 0$ can be realized only if $\tau(x)$ is identically equal to zero (i.e. the test body is in equilibrium without any shear stresses). Combining the basic assumption of limiting equilibrium (3), with the cracking hypothesis (which allows $S(x)$ to be zero only at x_f), imply that failure mechanisms located on the boundary $\{y^*(x), \sigma^*(x)\}$ are associated with $F = \widetilde{F}[y^*(x), \sigma^*(x)] \to \pm\infty$. Consequently, the classification inequalities are always non-active constraints, which need not be incorporated into the solution procedure. Fig. 3c illustrates this conclusion in a simple one-dimensional setting. Considering the creaking criterion (11.7) as a stopping rule for integrations of Euler's equation this criterion ensures that the inequalities in (11.5) are satisfied automatically, and they too become non-active constraints which need not be incorporated into the formal solution procedure.

6. Existence of a physically significant solution for active slope stability problems

6.1. General framework of the existence proof

The minimization problem (11) has a minimum by virtue of the fact that subject to the constraints supplied by the classification inequalities (11.4), the functional $\widetilde{F}[y(x), \sigma(x)]$ can deliver only non-negative values. This, however, is just a formality; the minimum of this problem can be realized either at a stationary point of the minimization criterion, or along the lower bound $F = 0$ provided by the classification inequalities. A result that the minimum value F_s of F is equal to zero would indicate that the problem defined by Eqs. (11) is not restrictive enough to yield significant solutions, justifying, in effect, the criticism of De Josseline De Jong (1980, 1981) and Castilo and Luceno (1980, 1982). However, in order to prove that the solution of this problem is physically significant it is sufficient to show that the lower bound $F = 0$ is inaccessible, i.e. that there is no pair of functions $\{y(x), \sigma(x)\}$, satisfying all the requirements in Eqs. (11), for which F is equal to zero. Combining the result that F cannot be zero, with the observation that $\widetilde{F}[y(x), \sigma(x)]$ is bounded from below by the value of zero, implies that the problem has a regular minimum which is realized at a stationary point.

Fig. 6 illustrates the above argument in a simple one-dimensional setting. All four functions shown in that figure are bounded from below by zero, and therefore they have minima. The minima of the functions in Fig. 6a and b occur along the lower bound $y = 0$, and their stationary points correspond to an inflection and maximum respectively. The function in Fig. 6c does not have an x value at which $y = 0$, and its minimum must occur at the stationary point. Fig. 6c shows that the requirement that there is no x value resulting with $y = 0$ is a sufficient but not necessary condition for the function $y(x)$ to have a minimum at a stationary point. Sewell (1987) presented similar arguments in a general multi-dimensional setting.

Based on the above argument, a proof that the stability problem (11) is well set, possessing a minimum which is realized at a stationary point, is reduced to a demonstration that the constraints defining this problem exclude the possibility that $F = 0$. This argument depends on two assumptions:

(a) Continuous dependence of F on failure mechanisms $\{y(x), \sigma(x)\}$.
(b) The constraints defining the problem in Eqs. (11) do not exclude all potential failure mechanisms.

Both of these assumptions are obviously satisfied for the present problem. Subject to these two conditions, a proof that F cannot be zero implies that the minimization problem (11) must have at least one stationary point (it may have more then one), and the minimum of this problem must be realized at one of the stationary points.

6.2. The basic proof

Consider first the general case in which $\phi \neq 0$. In that case the assignment $t = t^* = T = C/\psi$ is legitimate and it is convenient to start the investigation considering this case. In order to prove that safety factors cannot approach zero we will assume that F is equal to zero and show that this assumption leads to a contradiction among the elements of problem (11). Assuming $F = 0$ implies $\phi_m = \pi/2$, $\psi_m \to \infty$ and $C_m \to \infty$, but the ratio $C_m/\psi_m = C/\psi$ remains finite, resulting in the mobilized strength envelope $S_m(\sigma|F = 0)$ shown as the vertical line AB in Fig. 7a (the notation $S_m(\sigma|F = 0)$ means that the function $S_m(\sigma)$ is associated with $F = 0$). The complementary interpretation requires that Mohr circles describing the state of stress along potential slip surfaces must be tangential to $S_m(\sigma)$. Fig. 7a shows that when $S_m(\sigma)$ is the

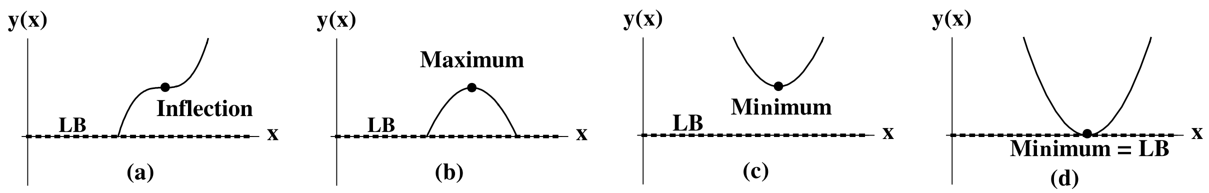

Fig. 6. Stationary values of functions with lower bounds: (a) inflection; (b) maximum; (c) minimum and (d) minimum at lower bound.

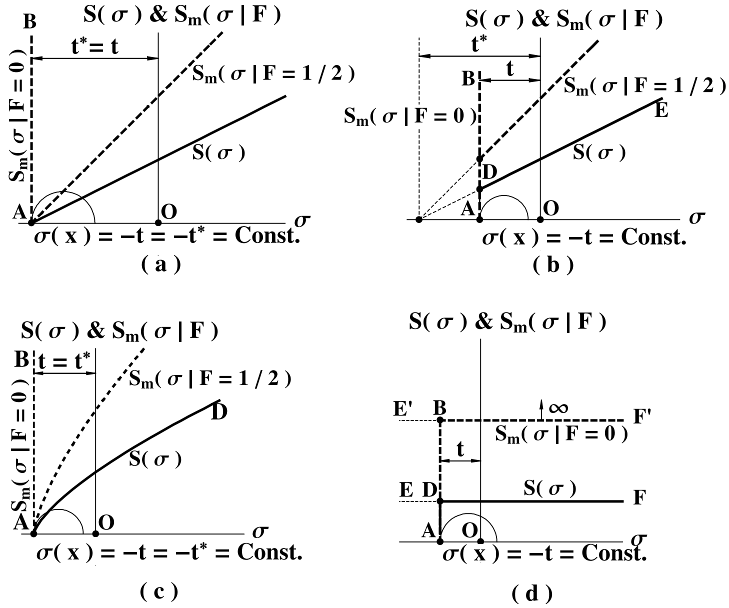

Fig. 7. Mobilized failure envelopes associated with $F = 0$. (a) Case 1: $\phi \neq 0$, $t = t^* = c/\tan(\phi)$. (b) Case 2: $\phi \neq 0$, $t < t^* = c/\tan(\phi)$. (c) Case 3: non-linear failure criterion $S(\sigma)$. (d) Case 4: constant strength function with tension cut-off.

vertical line AB the tangency requirement can be satisfied only if $\sigma(x) = -T = -t^* = -c/\psi =$ Constant at all point x along the slip surface. However the cracking hypothesis allows the relation $\sigma(x) = -T$ to occur only at a single point, resulting with a contradiction. As a result, F cannot be equal to zero, implying that problem (11) has a regular minimum occurring at a stationary point of the functional $\widetilde{F}[y(x), \sigma(x)]$. It is noted that for $F = 0$ the tangency requirement does not result with a unique state of stress (the radius of the Mohr circle in Fig. 7a is arbitrary); however we have just proved that $F = 0$ is impossible, so the non unique state of stress is not consequential.

Fig. 7b shows that the same argument holds for restricted Mohr–Coulomb envelopes with $t < t^*$. In fact the same proof is valid for any non-linear strength envelope with a finite tensile strength (Fig. 7c). This observation is practically significant since it can be shown (e.g. Jiang et al., 2003; Baker, 2002, 2003) that non-linearity of strength functions may, under certain circumstances, have a very significant effect on results of slope stability computations.

The above proof depends on the restriction (implied by the cracking hypothesis), that $\sigma(x) = -T$ can occur only at a single point. This requirement is not satisfied in the limiting case of surface failure ($y(x) = Y(x)$) in cohesion-less ($C = 0$) materials. In this limiting case $\sigma(x) = -T = -t^* = 0$ at all x values, and in order to complete the above proof it is necessary to verify that $\widetilde{F}[y(x) = Y(x), \sigma(x) = 0]$ cannot be equal to zero. Analysis of this limiting case is however trivial, leading to the result $F = \widetilde{F}[y(x) = Y(x), \sigma(x) = 0|C = 0] = \tan(\phi)/\tan(\beta)$ where β is the slope inclination. This relation is usually derived in the framework of the infinite slope approximation, but in fact it is valid also for finite slopes (e.g. Baker, 1981). The ratio $\tan(\phi)/\tan(\beta)$ approaches zero only in the limiting case of a vertical slope. This result suggests that problem (11) is not well set (does not have a minimum), in the particular case $\{C = 0, \beta = \pi/2\}$, and this observation is consistent with the well-known result that cohesion-less soils cannot support vertical slopes.

The methodology of the general proof breaks down in the limiting case of a constant strength function $S(\sigma) = C$. In this case the mobilized strength function $S_m(\sigma|F = 0)$ is a horizontal line located at infinity, and such envelope does not imply any restriction on $\sigma(x)$. The source of the difficulty associated with this case is that strength model $S(\sigma) = C$ implies an infinite tensile strength t^*. Introducing a tension cut off at some finite t value (i.e. considering the restricted envelope $S(\sigma) = C, \sigma \geqslant -t$), removes the unreasonable implication of the constant strength model, resulting with the situation shown in Fig. 7d. Applying the general argument to the mobilized strength function ABF' in Fig. 7d, shows that F cannot be zero, implying that problem (11) has a minimum also in this case.

The above considerations demonstrate the essential role played by tensile strength in the general structure of the slope stability problem. The stability of a given slope is affected by both tensile and shear strengths. Using an unrealistic model which implies an infinite tensile strength the general slope stability problem (Eq. (11)), is not well set, and it does not have a solution (minimum). However, modifying this model by the introduction of a finite tensile strength, results with a well set stability problem. The effect of tensile strength on slope stability was not sufficiently emphasized in previous presentations of the limiting equilibrium methodology. The above results show that the general constraint $\sigma(x) \geqslant -t^*$ should be considered as a physical relation representing the limited tensile strength of real soils, rather than merely as a formal consistency criterion for the Mohr–Coulomb failure criterion (which does not exist in the limiting case of $\phi = 0$ materials).

7. Summary conclusions and discussion

The results derived in the previous section show that when the strength model is associated with a finite tensile strength, safety factors cannot approach zero. A minimization problem in which the minimization criterion is bounded from below, and the lower bound cannot be realized, has a regular minimum, which is realized at a stationary point. This verifies that the general minimization problem (11) is properly set, possessing a regular minimum which is realized at a stationary point, thus establishing the main purpose of the present work.

The elements involved in the proof that the slope stability problem has a proper minimum, realized at a stationary point, are:

(A) *The complementary interpretation* which implies that Mohr circles describing the state of stress along potential slip surfaces are tangential to mobilized strength envelopes. This interpretation makes it possible to define completely the state of stress along potential slip surfaces, thus providing the basis for definition of tensile strength in a limiting equilibrium framework, and introduction of physical restrictions on normal stress functions $\sigma(x)$.

(B) *The classification inequalities* $\{\mathbf{T}_{\mathrm{H}} \geqslant 0, \mathbf{T}_{\mathrm{V}} \geqslant 0\}$ which define the active slope stability problem. These inequalities ensure that the minimization criterion is bounded from below by the value of zero, thus supplying the basic framework for the existence proof.

(C) *The cracking hypothesis* which specifies the consequences (crack formation) resulting from complete mobilization of tensile strength. Combined with the definition of a test body this hypothesis excludes normal stress functions in which the limiting condition $\sigma(x) = -T$ is satisfied at more than one point.

The following general comments are relevant with respect to the results derived in the present work:

(a) The structure of the above proof depends heavily on the cracking hypothesis. This hypothesis loses its restrictive power in cases where the strength criterion is not associated with a finite tensile strength, and in such cases it is impossible to prove that the general slope stability problem is well set. Most practical applications of limiting equilibrium slope stability analysis are associated with the linear Mohr–Coulomb criterion. This criterion defines the finite tensile strength $t^* = C/\tan(\phi)$ at all finite ϕ values, but this strength approaches infinity in the limiting case of the constant strength function $S(\sigma) = C$, which is associated with $\phi \equiv 0$. Consequently, the general slope stability problem (Eq. (11)), may be improperly set in the limiting case of the constant strength model $S(\sigma) = C$. It is impossible to state this conclusion in a more definitive form because the requirement $F \neq 0$ is a only sufficient but not necessary condition for the slope stability problem to have a minimum (e.g. Fig. 6d). Stated differently, the slope stability problem may have a minimum even in this limiting case, but the present proof methodology cannot be used in this degenerated case. It is noted that the small but finite compressibility of water guarantees that effective ϕ values can never be identically zero and the limiting case of $\phi \equiv 0$ material is not practically significant. However this limiting case is sometimes used in theoretical studies, and we found it instructive to clarify the consequences associated with using such a model in the slope stability context.

(b) The present work shows that the original formulation of the variational slope stability problem by Baker and Garber (1978) was incomplete (it did not included the restrictions A to C above). Without these constraints the formal slope stability problem is not well set, and it does not have a (global) minimum. However, the requirements in B are non-active inequalities, and the effect of C is expressed as the boundary condition (11.7) (the cracking criterion), which was used by Baker (1981). Therefore, solving the incompletely specified problem yields correct results if the starting point of the numerical minimization process is close enough to the solution point. Fig. 3c illustrates this argument in a simple one-dimensional setting.

(c) Considering a vertical cut-off in cohesive frictionless soil, De Josseline De Jong (1980) concluded that limiting equilibrium problems do not have minima (more accurately, using various formal second order variational criteria he was not able to prove that the variational solution derived by him is a minimum). It is noted that De Josseline De Jong considered the limiting case of a $\phi \equiv 0$ material without introduction of a finite tensile strength, and he did not included the classification inequalities as constraints. Based on the present perspective it is not surprising therefore that he could not prove that his solution is a minimum. In essence he considered an improperly defined problem, and verified that this problem does not have a solution. More seriously however, he "extrapolated" a result obtained for a particular singular case, and concluded that the variational approach to slope stability analysis is not valid. Such general conclusion is not implied by his results, and the present work shows that it is, in fact, incorrect. It is important to realize that there is nothing particularly "revolutionary" in the general variational formulation of slope stability analysis. This formulation utilizes classical limiting equilibrium elements (equilibrium of a test body, and definition of a safety factor), using variational calculus simply as a tool to identify the critical conditions. Therefore there cannot be a "variational fallacy" (De Josseline De Jong, 1981), and any fallacy, if one exists, must be a "limiting equilibrium fallacy". The present work verifies that the limiting equilibrium problem does have a minimum; if a minimum exists, it can be identified by variational calculus, and there is no fallacy (variational, or other).

(d) Space limitations prevent us from presenting a detailed discussion of the counter examples introduced by Castilo and Luceno (1980, 1982). It can be verified however (Baker, 2002) that all the counter examples discussed by them are associated with illegitimate normal stress functions. In particular, some counter examples violate the obvious requirement $\sigma \geqslant -T$. In one counter example they used a normal stress function satisfying $\sigma(x) = -T$ at number of internal points, thus implying formation of internal tension cracks (which are inadmissible in a limiting equilibrium framework). To a large extent the present work was motivated by an attempt to resolve the difficulties raised by those counter examples, and the writer acknowledge the important contribution of the counter examples presented by Castillo and Luceno to the present work.

(e) All limiting equilibrium procedures (including the variational formulation of Baker and Garber but also the classical procedures of Morgenstern-Price, and Janbu), address essentially the same basic problem (using different sets of "independent variables"). The transformation between the different independent variables is, in principle, trivial, and the present results show that the minimum exists for all slope stability formulations, which satisfy all equilibrium requirements. It is noted however that the present results have no implications with respect to approximate limiting equilibrium procedures such as Fellenius (1936) or Bishop (1955), which fail to satisfy all equilibrium conditions.

(f) The variational slope stability analysis presented by Baker and Garber (1978) was based on the method of Lagrange's undetermined multipliers. This method is an efficient procedure for identifying stationary points of constrained optimization problems, but it does not provide a convenient framework for establishing the character (minimum, maximum, inflection, etc.) of these points. This feature of the Lagrange method made it convenient to treat separately the questions of "existence of solution" (the present work, which does not utilize the method of Lagrange's multipliers), and "derivation of solution" (Baker and Garber, 1978; Baker, 1981; which is based on the Lagrange method). The results of the present work show that considering an active slope stability problem, one of the stationary points identified by the Lagrange procedure is a minimum, and this provides the missing link to previous presentations of the approach.

References

Baker, R., 1981. Tensile strength, tension cracks and stability of slopes. Soils and Foundations 21 (2), 1–17.

Baker, R., 2002. The general slope stability problem. Report published by the National Building Research Institute, Haifa, Israel.

Baker, R., 2003. Inter-relations between experimental and computational aspects of slope stability analysis. International Journal for Numerical and Analytical Methods in Geomechanics 27, 379–401.

Baker, R., Frydman, S., 1983. Upper bound limit analysis of soil with non-linear failure criterion. Soils and Foundations 23 (4), 34–42.

Baker, R., Garber, M., 1977. Variational approach to slope stability. In: Proceedings of the Ninth International Conference on Soil Mechanics and Foundation Engineering, Tokyo, Japan, vol. 2, pp. 65–68.

Baker, R., Garber, M., 1978. Theoretical analysis of the stability of slopes. Geotechnique 28 (4), 395–411.

Baker, R., Leshchinsky, D., 1987. Stability of conical heaps. Soils and Foundations 27 (4), 99–110.

Baker, R., Frydman, S., Talesnick, M., 1993. Slope stability analysis for undrained loading conditions. International Journal for Numerical and Analytical Methods in Geomechanics 17, 15–43.

Bishop, A.W., 1955. The use of the slip circle in the stability analysis of slopes. Geotechnique 5 (1), 7–17.

Castilo, E., Luceno, A., 1980. Evaluation of variational methods in slope analysis. Proceedings of International Symposium on Landslides, New-Delhi, India, vol. 1, pp. 255–258.

Castilo, E., Luceno, A., 1982. A critical analysis of some variational methods in slope stability analysis. International Journal for Numerical and Analytical Methods in Geomechanics 6, 195–209.

De Josseline De Jong, G., 1980. Application of calculus of variation to the vertical cut-off in cohesive frictionless soil. Geotechnique 30 (1), 1–16.

De Josseline De Jong, G., 1981. Variational fallacy. Geotechnique 31 (4), 289–290.

Fellenius, W., 1936. Calculations of the stability of earth dams. Transactions of the 2nd Congress on Large Dams, Washington, DC, vol. 4, p. 445.

Garber, M., Baker, R., 1977. Bearing capacity by variational method. Journal of the Geotechnical Engineering Division—Proceedings of ASCE 103 (GT 11), 1209–1225.

Janbu, N., 1973. Slope stability computations. In: Hirshfield, R.C., Poulos, S.J. (Eds.), Embankment Dam Engineering. Cassagrande Volume. John Wiley and Sons, pp. 47–86.

Jiang, C.J., Baker, R., Yamagami, T., 2003. The effect of strength envelope nonlinearity on slope stability computations. Canadian Geotechnical Journal 40 (2), 308–325.

Kopacsy, J., 1955. Uber die brudkflachen und bruckspannungen in der erdbauter. In: Szechy, K. (Ed.), Gedenkbuck fur Prof. Dr. J. Jaky., Budapest, Hungary, pp. 81–99.

Kopacsy, J., 1957. Three dimensional stress distribution and slip surface in earth work at rupture. Proceedings of the Forth International Conference on Soil Mechanics and Foundation Engineering, London, England, vol. 1, pp. 339–342.

Kopacsy, J., 1961. Distribution des contraintes a la rupture forme de la surface de glissement et hauteur theorique des talus. Proceedings of the 5th International Conference on Soil Mechanics and Foundation Engineering, Paris, France, vol. 2, pp. 641–650.

Leshchinsky, D., Baker, R., Silver, M.L., 1985. Three dimensional analysis of slope stability. International Journal for Numerical and Analytical Methods in Geomechanics 9, 199–223.

Morgenstern, N.R., Price, V.E., 1965. The analysis of the stability of general slip surface. Geotechnique 15 (4), 289–290.

Petrov, I., 1968. Variational Methods in Optimum Control Theory. Academic Press, New York and London.

Sewell, M.J., 1987. Maximum and Minimum Principles. Cambridge University Press, Cambridge, New York, New Rochelle, Melbourne, Sydney.

Taylor, D.W., 1937. Stability of earth slopes. Journal of Boston Society of Civil Engineering XXIV (3), 197–246.