ON BASIC FEATURES OF CONSTITUTIVE MODELS FOR GEOMATERIALS

IVO HERLE

Institute of Geotechnical Engineering,
Technische Universität Dresden,
01062 Dresden, Germany,
e-mail: ivo.herle@tu-dresden.de

[Received 08 August 2007. Accepted 25 February 2008]

ABSTRACT. Constitutive models are inevitable for numerical calculations of the mechanical behaviour. A large amount of proposed models makes it difficult to judge their suitability for particular applications. In the paper, several fundamental properties of the soil behaviour (nonlinearity of stress-strain curves and stress envelopes, proportional stress and strain paths, irreversibility and deformation history) are discussed with respect to commonly used models.

KEY WORDS: constitutive models, soil behaviour, nonlinearity, irreversibility, proportional paths, deformation history.

1. Introduction

Constitutive models are inevitable for all numerical calculations of boundary value problems. Concerning geotechnical engineering, a large variety of constitutive models have been proposed, from very simple to extremely complex ones. This is not surprising if we take into account many peculiarities of the soil and rock mechanical behaviour. However, Gudehus states in 1985 [23]: “Soil seems to defeat its investigators again and again.” Thirteen years later this state has not improved substantially and Savage [57] concludes: ”One of the central problems of research in granular materials is the development of continuum constitutive equations”. Both citations are still valid today and there is not a consent on a perfect complete constitutive model for any soil type. Nevertheless, there are some basic features of the soil behaviour which should be considered in the most cases. This paper provides a short overview of these features.
2. Classification of constitutive models

It is usual to classify constitutive models with respect to their mathematical ingredients. Many comprehensive publications are available, e.g. [28, 12, 11], where a rational, mostly inductive development of models is summarized and commented. The emphasis is put on properties of equations and their hierarchy. An example of such a classification is in Table 1.

This approach is certainly useful for model developers but rather confusing for the wider professional public. Engineers as users of the constitutive models need to select proper models for a particular task. They do not want and do not need to have a thorough background on the formulation of intricate yield surfaces and complicated hardening rules. On the other hand, they want and need to be acquainted with scopes and limitations of the models with respect to the relevant effects of the soil behaviour characteristic for particular boundary value problems. Thus, not a traditional classification but a model evaluation, like presented e.g. by [69] or [46], appears more appropriate for the users of constitutive models in the geotechnical community.

Development and testing of constitutive models is based on laboratory experiments of soil specimens (element tests). Thus, also the checking of soil models should be done against some basic soil features observed in the element tests like

- nonlinearity;
- irreversibility;
- asymptotic behaviour in proportional stress and strain paths;
- failure criterion;
- stress-volumetric coupling;
- deformation history;
• anisotropy;
• time-dependence (rate-effects);
• …

The above listed features are mostly interrelated and cannot be, in general, treated independently. However, depending on the problem to be solved, some features have a greater impact on the focused quantities than others. For example, pore pressure calculations are certainly more influenced by the stress-volumetric coupling than by anisotropy. It will remain a task for an engineer, to recognize relevant soil features for a particular boundary value problem and to neglect the secondary effects. Based on this identification, a suitable constitutive model can be selected.

It is more important to assure correct trends predicted by a constitutive model in different representations than to attempt a perfect fit of a single experimental curve. One should be aware that it is not possible to obtain absolutely identical curves from experiments under the same control conditions if using different apparatuses [56]. Moreover, sophisticated laboratory techniques reveal that even apparently homogeneous deformations of boundaries may lead to complex inhomogeneous patterns inside the specimens [14]. These go far beyond standard patterns of strain localization, are hidden inside the specimen and cannot be captured by any constitutive model. Taking into account the large scatter in experimental results from different apparatuses and a frequent inhomogeneous deformation, which is not visible at the surface of the tested specimens, one should keep a reasonable distance when comparing numerical and experimental results.

3. Nonlinearity of the stress-strain curve

Nonlinearity of the stress-strain response is perhaps the most important feature of the soil behaviour. Still, it is not straightforward how to handle it. We can namely observe two opposite tendencies: compressive stress paths are characterized by increasing stiffness for constant strain rate whereas shear response is characterized by decreasing stiffness (Fig. 1).

This pattern can be modelled by different equations for each type of the behaviour. Terzaghi and later Ohde [51] proposed a pressure-dependent stiffness for oedometric compression and Kondner [42] used a hyperbolic relationship between shear stress and shear strain. However, the fast development of numerical methods needed a general treatment of both paths. This was accomplished by Duncan and Chang [17] who unified both equations for a FE implementation. Their approach became popular and can still be found
Ivo Herle

Fig. 1. Different nonlinear responses for compression and shear

in many software packages. Although single curves can be modelled rather realistically, we must realize that there is not a thorough theory behind this approach. Perhaps the most remarkable deficiency is the lack of a consistency condition which results in a sharp jump in stiffness between loading and unloading along a neutral path. Corresponding numerical problems have been reported [63]. Moreover, the volumetric response is not related to the stress-strain curves and cannot take into account coupling between shear stress and volume changes and vice versa.

In spite of the well-know deficiencies one can find further proposals which are mere extensions of this concept. Especially, the recent trend in the description of the behaviour at small strains has brought several contributions in this direction, see e.g. [36] or [53]. They are based on a close fit of experimentally measured shear modulus by a single curve. A scalar relationship is then used to model the incremental stiffness in FE calculations. However, experimental paths may be very different from in situ conditions and a concept for an extrapolation behind the experimental paths is missing.

Nonlinear “increase” or “decay” (also called “degradation”) of incremental stiffness can be modelled equally well by the majority of more sophisticated and complex models, at least for standard laboratory paths. Looking at a simulation of a stress-strain curve of a drained triaxial compression test on a remoulded sample we cannot distinguish between e.g. the Kondner model, the Cam Clay model, a kinematic hardening model or a hypoplastic model. Therefore, we should always consider more aspects simultaneously [46]. Cam Clay models [55, 58] can be excellent for soft soils under drained triaxial paths but they overpredict the stiffness in undrained conditions [52]. Hypoplastic models [39, 70, 24] are excellent in asymptotic states and capture realistically
the influence of mean pressure and relative density. Still, the initial stiffness in shear is often underestimated.

The nonlinear stress-strain behaviour is coupled with the volumetric behaviour and both effects cannot be separated from each other. However, quasi-elastic models usually cannot deal with it. Especially at limit stress states, the direction of elastic strain increments is completely unrealistic. An insufficient modelling of volumetric effects may be less important e.g. for excavation or footing problems [71], but it becomes critical in problems with a high degree of confinement like piles, tunnels or anchors [50, 33, 32]. In case of undrained conditions, the volumetric constraint is maximal and effective stress calculations with quasi-elastic models are impossible due to unrealistic pore pressures.

Fig. 2. Distribution of settlements beneath an embankment. "Plastic" denotes Mohr-Coulomb model and "power-law" stands for the stress-dependent elasticity.

Quasi-elastic models may be inappropriate even in tasks with apparently well-defined paths. Consider e.g. a settlement calculation of an embankment on soft subsoil [31]. Taking into account the plane strain conditions and the symmetry of the embankment, many practitioners use an one-dimensional approach ("oedometric conditions") with load exerted by the embankment weight. Without hardening plasticity or hypoplasticity, only elastic strains develop, yielding a single value of the maximal settlement in the embankment symmetry axis. Discretizing a 2D domain, one obtains additionally a distribution of the settlements beneath the embankment (Fig. 2). Further improvement can be the incorporation of a limit stress condition, e.g. within a nonlinear elastic-perfectly plastic model. The differences in results are enormous and
in case of the plastic calculations, the maximum settlement is reached outside
the symmetry axis. Thus, modelling of the nonlinear behaviour even for purely
deformation problems must be always embedded into a realistic description of
the limit stress states. Similar observations can be also done for calculations
of tunnels [8] and excavations [68].

4. Limit stress envelope

Nonlinearity does not involve only the stress-strain behaviour. The
locus of limit stresses is usually nonlinear as well (Fig. 3 (left)), although many
rather sophisticated models assume linearity. Only in case of hard grains in a
relatively narrow pressure range or for remoulded clayey soils the envelope can
be linearized.

Carefully conducted laboratory experiments reveal a pronounced in-
crease of the maximum stress ratio with decreasing mean pressure not only for
sand and other granular materials [41] but for clayey soils as well [2]. Friction
angles above $50^\circ$ are not exceptions for sandy soils in the low pressure range.

The nonlinearity of limit stresses has a decisive influence e.g. on slope
stability calculations [2]. A linear extrapolation of the limit stress envelope
can produce a serious overestimation of the safety factor $F$. For a simple slope
geometry, a Mohr-Coulomb envelope may yield $F = 1.5$ whereas a nonlinear
envelope $F < 1$ [4].

Further trouble can arise from the formulation of the limit stress con-
dition in stress invariants (Fig. 3 (right)). If the third stress invariant is not
taken into account, the cross-section of the resulting failure criterion in the de-

---

Fig. 3. A linear approximation inside the experimental range usually neglects the
nonlinearity of the limit stress envelope (left). Deviatoric section of different limit
stress surfaces (right)

---

Ivo Herle
viatoric plane is of a Drucker-Prager type. E.g. the family of Cam Clay models is formulated in $p'$ and $q$ invariants with the critical stress ratio $q/p' = M$. This is unrealistic not only in 3D calculations. Even in axisymmetric or plane-strain problems, the large difference between the shear strength in compression and extension can become dangerous. Using $M = 1.2$, we get

$$\sin \varphi_c = \frac{3M}{6 + M} \leadsto \varphi_c = 30^\circ \text{ in compression, and}$$

$$\sin \varphi_e = \frac{3M}{6 - M} \leadsto \varphi_e = 48.6^\circ \text{ in extension}$$

for the friction angle within the Mohr-Coulomb criterion.

Many constitutive models incorporate void ratio in order to simulate asymptotic critical states with vanishing volumetric strain rates. A unique relationship between mean stress and void ratio in the critical state is presupposed. Although this assumption is pragmatic and certainly sufficient under many circumstances, advanced laboratory experiments on sand \cite{16} and clay \cite{66} show a clear dependence of the critical void ratio on the deformation (consolidation) history. Furthermore, we can also find hints that compression (oedometric and isotropic) lines in a $e$-ln $p'$ representation are not parallel with the critical state line \cite{27}. In this case the common constitutive assumption — ”undrained shear strength is proportional to the consolidation stress” — would be no longer valid.

5. Proportional stress and strain paths

Proportional stress and strain paths are obtained if the stress rate and the strain rate are proportional to the current stress $\sigma$ and strain $\varepsilon$, respectively, i.e.

$$\dot{\sigma} = \text{const} \cdot \sigma \quad \text{and} \quad \dot{\varepsilon} = \text{const} \cdot \varepsilon.$$  

Proportional paths correspond to common loading scenarios in the laboratory and also in many in situ problems \cite{26, 23}. Isotropic or oedometric compression tests may serve as examples (Fig. 4). Considering e.g. oedometric compression, the stress ratio along the corresponding proportional stress path is $K_0$ (earth pressure coefficient at rest). This value can be incorrectly predicted by several models. Not only by elasticity where $K_0 = \nu/(1 - \nu)$ is a function of Poisson ratio $\nu$. Also the Cam Clay model and its derivatives remarkably overpredict $K_0$ \cite{52, 9}. E.g. for $\varphi = 30^\circ$ the Cam-Clay model predicts $K_0 = 0.62$ whereas the empirically verified Jaky relation yields $K_0 = 0.50$. 
Fig. 4. Proportional strain paths and the corresponding stress paths

However, not the proportional paths *per se* but their asymptotic properties are of great importance. Starting from any feasible stress state, a proportional stress path is finally reached with constant strain rate (Fig. 5). Thus, proportional strain paths tend to proportional stress paths and vice versa. Consequently, the past deformation history is erased and the soil behaviour is fully determined by the actual stress and the direction of the deformation rate (if effects of density variation and the magnitude of strain rate are neglected). When straining along the proportional stress and strain paths, a SOM (swept-out-memory) state is reached [26]. This behaviour was found experimentally [21] and it can be modelled if the stress rate is a homogeneous function of the $n$-th degree of stress, i.e.

\[
\dot{\sigma} = f(\lambda \sigma, \dot{\varepsilon}, \ldots) = \lambda^n f(\sigma, \dot{\varepsilon}, \ldots).
\]

The asymptotic property of the proportional paths is of a great advan-
tage in the determination of initial states for numerical simulations of boundary value problems. Measured values of state variables are mostly burdened by various errors. Using such values for numerical predictions can be risky. It is much better to obtain the initial state from a numerical simulation of simplified geological histories composed of proportional paths. In this way, uncertainty and scatter in input data can be reduced.

Some of the proportional strain paths lead to the limit stress states. Hence, a numerical procedure with proportional strain paths of different directions from various stress states can check whether stresses do not escape from their allowable range [38]. This test may be useful especially during the development of new models (Fig. 6).

Fig. 6. Testing of constitutive equations with proportional paths (left: correct model, right: deficient model) [38]

6. Irreversibility (unloading)

Many boundary value problems are not characterized only by monotonous paths. Sharp bends or path reversals ("unloading") must be often taken into account.

Consider "unloading" into the elastic range. It is often recommended to use $\nu$ between 0.15 and 0.25 [10]. However, $\nu = 0.2$ corresponds to the ratio of stress increments

$$\frac{\sigma_3}{\sigma_1} = \frac{\nu}{1 - \nu} = 0.25.$$  

This ratio is too low as compared with experimental results (Fig. 7) and the corresponding elastic stress path (arrow in Fig. 7) approaches fast the limit
Fig. 7. Elastic unloading (arrow) in oedometric test for $\nu = 0.2$. Points: experiments with Karlsruhe sand [5, 40]; full lines: hypoplastic model [30].

Fig. 8. Stress-strain curves (left) and an unrealistic volume increase at unloading (right) in a standard triaxial test. Full line: Mohr-Coulomb model, dotted line: correct behaviour from a hypoplastic calculation.

Another aspect concerns the volumetric behaviour after a reversal of the stress (or strain) path. It was shown experimentally that a pronounced contractancy sets on after each reversal [20, 43]. However, elasticity predicts the opposite if the mean pressure decreases, see Fig. 8. A comparison of the
simulation results of a tail gap grouting of a TBM reveals a qualitatively different resulting settlement trough due to this effect [44], see Fig. 9.

![Diagram of deformation history](image)

Fig. 9. Surface deformation due to tail gap grouting during a tunnel excavation [44]. Results with the Mohr-Coulomb (left) and the hypoplastic (right) model

7. Deformation history

Material parameters are quantities which characterize intrinsic properties of a particular soil independently of deformation. On the contrary, state variables change their values in the course of straining and reflect not only the actual state but include also information on the deformation history. They should be physically meaningful and preferably also measurable, at least via a hypothetical experiment (Gedankenexperiment) [22].

The most important state variable for soils is the effective stress tensor (its counterpart — strain tensor — is usually not a state variable since its initial
value can be chosen arbitrarily if we do not have any reference configuration). Stiffness and strength of soil are influenced especially by the mean stress. The direction of strain increments depends also on the actual stress state.

However, the stress tensor is insufficient in describing differences in the behaviour of loose and dense soils. Therefore, the void ratio is often used as a further state variable. It is clearly necessary to consider the actual void ratio and not only its initial value. Furthermore, the void ratio itself is of a low importance if it is not related to its physical limits. The relative density

\[ D_r = \frac{(e_{\text{max}} - e)}{(e_{\text{max}} - e_{\text{min}})} \]

appears appropriate at first glance.

It is known from laboratory experiments that dense soils under high pressures respond in a similar way like loose soils at lower pressures. In order to model such a behaviour, a relationship between the actual void ratio and a pressure-dependent parameter must be established. One possibility is the equivalent stress \( p_e \) proposed by Hvorslev \[34\]

\[ p_e = p_0 \exp \left[ \frac{1}{\lambda} (e_0 - e) \right], \]

which relates the actual void ratio \( e \) to the normal consolidation line described as

\[ e = e_0 - \lambda \ln(p/p_0). \]

Another possibility is to consider the distance \( \psi \) between \( e \) and the critical void ratio \( e_c \) at the same effective mean stress \( p \) \[7\]

\[ \psi = e - e_c. \]

Both, \( p_e \) and \( \psi \) relate the actual void ratio \( e \) to a single pressure-dependent parameter. To judge if soil should be considered as a loose or a dense one, Gudehus \[24\] proposed as a density measure the so-called pressure-dependent relative void ratio

\[ r_e = \frac{e - e_d}{e_c - e_d} \]

with \( e_d \) and \( e_c \) being the critical and minimum void ratio, respectively. Both, \( e_c \) and \( e_d \) are pressure-dependent \[6\].

A consistent treatment of the pressure and density effects is very important for the description of the asymptotic behaviour towards the critical state. The critical void ratio decreases with mean pressure and, consequently,
dilatancy and contractancy effects are always coupled with particular pressure levels. The description of critical states should belong unambiguously to the essential features of all models for soils [47].

The effective stress tensor and the void ratio sufficiently characterize the soil state in proportional (monotonous) paths with constant deformation rate. Due to the SOM property, the influence of previous deformation history is erased. The situation becomes more difficult for the behaviour after sharp bends in deformation paths. Consider isotropic compression of a soil element (Fig. 10). The different response of the curves \(a\) and \(b\) at the point \((p_1, e_1)\) can be explained and modelled by the opposite deformation direction during the monotonous loading and unloading, respectively. However, the deformation direction at the point \((p_2, e_2)\) is the same for loading and reloading. Hence, the state variables stress and void ratio do not contain enough information to model this case. A further state variable is needed for taking into account the role of recent deformation.

Fig. 10. Isotropic compression — monotonous versus cyclic behaviour (the state during reloading is not sufficiently described by stress \(p\) and void ratio \(e\))

Fig. 11. Evolution of intergranular strain \(S\) during monotonous deformation

Similar conclusions were already drawn in the 19th century for earth pressure [13]. Nevertheless, a clear picture on this topic was first enabled through the development of techniques for the measurement of small strains directly on laboratory specimens [37, 64, 67]. Experimentally observed relationships between the stiffness and the strain path direction and length after a sharp bend [3, 19, 35, 65] gave a strong impact towards new formulations of constitutive models for soils. Most models are based on the theory of elastoplas-
ticity with kinematic hardening [45, 15, 62, 18, 54]. They can reproduce many small-strain effects but their mathematical construction is complicated following the strict rules of the elastoplastic geometrical framework. Further difficulties can arise from the determination of initial values of additional state variables. These values must be known prior to the calculation of BVP although they are in fact not measurable. Therefore, it is important that these initial values can be obtained from numerical simulations of asymptotic states [22].

To illustrate the kinematic hardening in the stress space, Simpson [61, 59] used the idea of a man connected to several bricks with strings of different lengths. A free movement of a man corresponds to the elastic behaviour with a high stiffness. Pulling the bricks behind results in plastic behaviour. Thus, positions of the bricks in the strain space are considered as memory (state) variables.

Niemunis and Herle [48] proposed the so-called intergranular strain tensor $S$, which is a state variable storing the recent deformation history. Although they applied this approach to hypoplasticity, the concept of intergranular strain is independent of any particular constitutive framework. Simplifying the representation into one dimension, we can write

$$\dot{S} = \left(1 - \frac{|S|}{R}\right) \epsilon, \quad \text{for } \epsilon S > 0,$$

$$\dot{S} = \epsilon, \quad \text{for } \epsilon S \leq 0,$$

with $S$ being a state variable (intergranular strain) and $R$ a material parameter related to the elastic strain range.

If $S < R$, the stiffness is increased

$$\dot{\sigma} = \frac{1}{R} [m(R - |S|) + |S|] M \epsilon, \quad \text{for } \epsilon S > 0,$$

$$\dot{\sigma} = m M \epsilon, \quad \text{for } \epsilon S \leq 0,$$

with $M$ being the tangential (hypoplastic) stiffness $M = f(\sigma, \epsilon, \dot{\epsilon} / |\dot{\epsilon}|)$ and $m$ a further material parameter controlling the stiffness increase after a reversal of the strain path.

The outlined evolution equation for $S$ possesses asymptotic properties as $S \to R$ for monotonous deformation (Fig. 11). In case of cyclic deformation, an asymptotic stabilization of $S$ is obtained as well (Fig. 12).

Capturing the recent deformation history corresponds to the modelling of anisotropy. It is not always necessary to distinguish between inherent and induced anisotropy. The first case is sometimes obtained during the deposition
Fig. 12. Evolution of intergranular strain $S$ during cyclic deformation

process which also belongs to the deformation history and should be modelled, at least qualitatively, with proportional deformation paths.

8. Final remarks

Constitutive models for soils are usually complex equations intended for predictions of the soil behaviour. Besides objectivity (independence of reference system and units), they should also include several essential features observed in laboratory tests of soil samples: nonlinearity, limit stress states, proportionality of stress and strain paths, irreversibility and realistic selection of state variables. Constitutive models should be useful not only to their designers, i.e. they must be also user-invariant [23] and different users should be able to obtain same solutions.

It cannot be expected that the end-users will write computer codes for novel and promising models or for checking the model properties. Nevertheless, the vast majority of newly developed models come from academia, which should be able to offer free source codes of the models in form of subroutines with a brief description for a download in internet. Such codes can be copyrighted, e.g. by the GNU general public licence, in order to prevent commercial misuse by third parties. A project towards a better communication on constitutive models has started recently [25].

Many important features of the soil behaviour and their constitutive modelling have not been discussed, among them rate effects, rotation of principal stress axes, cementation and degradation of grains or particle breakage. It remains an open question which trend is more suitable for such extensions: (a)
to develop simple models for specific tasks (the way preferred by practice\textsuperscript{1}), or (b) to formulate general models with comprehensive features (the way often preferred by academia).

Ambition to discover new theories belongs certainly to the main motivation for research. However, “one should consider what benefits or advantages any new proposals might offer beyond those of existing approaches” [57]. This is certainly more than urgent in case of the constitutive models for soils. It should be equally worth to put more effort into verification (or rather falsification in the Popper’s sense) and validation of existing models. After all, the reliability of computational results belongs to the main tasks of the present research in the computational mechanics [49].

REFERENCES


\textsuperscript{1}“Researchers are encouraged to present simplified versions of their findings, with the range of applicability limited as necessary, to enable designers to carry out relatively simple, but useful analyses” [60]


On Basic Features of Constitutive Models for Geomaterials