Work and energy equations and the principle of generalized effective stress for unsaturated soils

C. G. Zhao[1,*,†,‡], Y. Liu[1] and F. P. Gao[2,§]

[1]*Department of Geotechnique, School of Civil Engineering, Beijing Jiaotong University, Beijing 100044, China*
[2]*Institute of Mechanics, Chinese Academy of Sciences, Beijing, China*

SUMMARY

The effective stress principle has been efficiently applied to saturated soils in the soil mechanics and geotechnical engineering practice; however, its applicability to unsaturated soils is still under debate. The appropriate selection of stress state variables is essential for the construction of constitutive models for unsaturated soils. Owing to the complexity of unsaturated soils, it is difficult to determine the deformation and strength behaviors of unsaturated soils uniquely with the previous single-effective-stress variable theory and two-effective-stress-variable theory in all the situations. In this paper, based on the porous media theory, the specific expression of work is proposed, and the effective stress of unsaturated soils conjugated with the displacement of the soil skeleton is further derived. In the derived work and energy balance equations, the energy dissipation in unsaturated soils is taken into account. According to the derived work and energy balance equations, all of the three generalized stresses and the conjugated strains have effects on the deformation of unsaturated soils. For considering these effects, a principle of generalized effective stress to describe the behaviors of unsaturated soils is proposed. The proposed principle of generalized effective stress may reduce to the previous effective stress theory of single-stress variable or the two-stress variables under certain conditions. This principle provides a helpful reference for the development of constitutive models for unsaturated soils. Copyright © 2009 John Wiley & Sons, Ltd.

*Correspondence to: C. G. Zhao, Department of Geotechnique, School of Civil Engineering and Architecture, Beijing Jiaotong University, Beijing 100044, China.
†E-mail: cgzhao@center.njtu.edu.cn
‡Professor and Director.
§Associate Professor.

1. INTRODUCTION

In recent decades, there has been significant development in theoretical soil mechanics. However, most of the proposed theories, for example, the soil seepage and deformation models, are based on intuitions, experimental observations, empirical and macroscopic phenomenological approaches, and different fundamental premises and assumptions, which are lack of a unified theoretical basis. The porous media theory (also known as the theory of mixture) in combination with the thermodynamics and internal variable theory may provide an efficient method to uniformly describe the complicated interactions among different phases of soil under the action of external or environmental loads, so that the soil seepage and deformation could be described in a unified framework. It can serve as the rational and scientific basis for the establishment of the constitutive relationships for the porous media, especially for unsaturated soils. For porous media theory turning into the unified theoretical basis of soil mechanics, there are a lot of problems to be solved, this paper is our efforts on these problems.

Porous media theory plays an important role in many branches of engineering, including material science, the petroleum industry, chemical engineering, and soil mechanics, as well as biomechanics. The porous media theory applied in soil mechanics has been explained by its historical progression, and its current state has been thoroughly described by de Boer [1]. Based on the theory of porous media and thermodynamics, the theoretical models for the behavior of unsaturated soils have been investigated intensively by a few researchers, such as Bowen [2, 3], Lewis and Schrefler [4], Schrefler [5], Hassanizadeh and Gray [6, 7], Gray and Hassanizadeh [8], Eringen [9], Hutter *et al.* [10], Schrefler [11], Laloui *et al.* [12], Borja [13], Sheng *et al.* [14], Wei and Dewoolkar [15], Gray and Schrefler [16], Li [17, 18], Gray *et al.* [19], etc.

The selection of stress state variables is the key for establishing constitutive models of unsaturated soils. To interpret the mechanical behaviors of unsaturated soils, various forms of stress variables have been proposed in the literature. Bishop [20] proposed the most well-known effective stress of unsaturated soils using a single-stress variable, that is

$$\boldsymbol{\sigma}' = (\boldsymbol{\sigma} - P_g \boldsymbol{\delta}) + \chi s \boldsymbol{\delta} \tag{1}$$

Where $\boldsymbol{\sigma}'$ is the effective stress, P_g is the pore gas pressure, $\boldsymbol{\delta}$ is expressed in a matrix format as $\boldsymbol{\delta} = [\delta_{ij}]$, δ_{ij} is the Kronecker's delta, s is the matric suction and χ is called the effective stress parameter, which varies from 0 (dried) to 1.0 (fully saturated). Aitchison and Donald [21], Jennings [22], Bishop and Donald [23], Bishop and Blight [24], Blight [25] also gave some expressions of effective stress similar to Equation (1). The purpose of defining Equation (1) is to use the single 'effective stress' variable to determine the deformation and strength behaviors of unsaturated soils as what used for saturated soils. However, Jennings and Burland [26] questioned the validity of the effective stress principle in unsaturated soils and argued that it is not able to explain the collapse phenomenon upon wetting. They conducted a series of consolidation tests on different unsaturated soils, and found that upon wetting, that is, reducing suction, all samples collapsed at the end. However, the calculation results based on the effective stress equation (1) show that soil samples are supposed to swell instead of collapse as shown by experiment data. Additionally, there is not a unique relationship existing between the degree of saturation (S_r) and the parameter χ. Burland [27] further discussed that, from a microscopic viewpoint, there was internal contradiction in the effective stress expression with the single stress variable. Similar arguments were also put forward by Aitchison [28], Matyas and Radhakrishna [29], Brackley [30], Fredlund and Morgenstern [31], and Wheeler and Karube [32], etc. But latter researcher shown that the collapsing behavior in reality

is the macroscopic evidence of a structural instability of the soil skeleton (or plastic deformation), and it is independent of the stress variables adopted. Collapse upon wetting cannot be interpreted as an effect of a mere increase and decrease of soil skeleton stress and it is also as for the suction-induced cementing effect. In fact, the single variable of effective stress being a sort of soil skeleton stress, it is difficult at the same time to consider independently the suction-induced hardening or describe the hydraulic hysteresis behaviors of unsaturated soils. Generally speaking, the concept of single variable of effective stress in unsaturated soils is borrowed from the effective stress principle in saturated soils. It is a macroscopic, intuitive and empirical expression without a clear physical mechanism and basis. However, since it is simple and similar to the effective stress in saturated soils, it is easy for engineers to understand and apply in existing finite element programs. From the viewpoint of engineering application, it has certain practical significance to study the single variable effective stress principle and its engineering applications. Hence, Khalili [33] upheld that the single stress variable could describe strength and deformation behaviors of unsaturated soils, also gave some new explanations and proofs.

To overcome the limitations of the single variable effective stress theory, Coleman [34], Bishop and Blight [24], Blight [35] proposed to use two independent stress variables (net stress and matric suction) to describe the strength and deformation behaviors of unsaturated soils. Fredlund and Morgenstern [31] verified the validity of using two independent stress variables by null tests. From then on, research using two independent stress variables has got a rapid development and became the mainstream in the research. However, the theoretical basis of two-stress variables is not clear either. Moreover, the two-stress variable theory cannot well describe the complicated behaviors of unsaturated soils. For example, Wheeler *et al.* [36] pointed out that the properties of unsaturated soils were affected by not only the net stress and matric suction, but also some other factors such as the degree of saturation. Because even if the net stresses, suction and specific volume were same for the two soil samples, they would be expected to show different behaviors and have different effective stresses, due to the difference in the degree of saturation and the inter-particle contact forces transmitted through the soil skeleton. As a result, two independent stress variables are not sufficient to determine the deformation and strength behaviors of unsaturated soils uniquely. Later some researchers, such as Sheng *et al.* [14], Li [17, 18], Sun *et al.* [37], etc., established constitutive models for unsaturated soils in terms of the degree of saturation in addition to two-stress variables. Their works imply that the two independent stress variables of effective stress theory for unsaturated soils need further development. The historical developments of effective stress had been reviewed with objective of determining a proper stress framework for constitutive modeling of unsaturated soils by Nuth and Laloui [38].

The validity of the theory of the two-stress variables has also been questioned. Tarantino and Mongiovi [39] pointed out that the null tests executed by Fredlund and Morgenstern [31] could verify that two independent stress variables would be sufficient to describe fully the stress state of unsaturated soils. On the other hand, it is indicated that the axis-translation technique is no longer valid when gas phase is occluded [40]. In other words, the so-called axis-translation caused by the change of occluded pore gas pressure will affect the properties of unsaturated soils. For that reason, there will be some problems if two independent stress variables (net stress and suction) are used and variation of pore gas pressure is neglected. Although the correctness of two independent stress variables and the validity of axis-translation technique under continuous gas phase condition had been preliminarily verified by lab tests by Tarantino and Mongiovi [39], they doubted the correctness of the two independent stress variables in occluded gas cases, and suggested that the further research and more experimental evidences are needed.

The energy principle in porous media theory can be used to describe the interaction and the transformation of different components in a system. It is therefore reasonable to use the energy principle to describe the interaction and mutual transformation caused by the general stresses and conjugated strains. The deformation work in the energy equation includes stress of each phase of unsaturated soils and its conjugated strain (note: their product is the deformation work), the study of the deformation work and energy balance equation is beneficial to understand the behaviors of unsaturated soils. It has attracted much attention from many researchers, for example Houlsby [41, 42], Schrefler [11], Jommi [43], Vaunat *et al.* [44], Nuth and Laloui [38], etc. The energy balance equations and the second law of thermodynamics serve as the basis and starting point for the development of the constitutive model of unsaturated soils by using the theory of porous media. However, it is not well solved and needs further study on how to use geotechnical symbol and terminology to derive the work and energy balance equations and the relationship between them with the consideration of energy dissipation based on porous media theory. Houlsby [42] and Li [17] derived an explicit expression of input work and energy balance using geotechnical symbols and terminologies. Nevertheless, their explicit expressions of input work and energy balance were not rigorously derived from porous media theory, and just as mentioned by Nuth & Laloui [38, p. 784] that 'considering that the full implications of Houlsby's assumptions for deriving Equation (28) (the expression of input work) remained to be investigated'.

A form for the solid pressure was introduced by Lewis and Schrefler [4] and Schrefler [5], where Biot's theory was extended to two-phase flow in a deformable porous medium. If the effective stress parameter χ is set to equal the degree of saturation, S_r, in Equation (1), a generalization of Bishop's effective stress can be obtained. The generalized Bishop's effective stress has a desirable feature that it can be conveniently used in modeling a mixed saturated–unsaturated system. This effective stress expression is continuous (no jump condition exists), when the state of the material transits from the partially saturated zone to the fully saturated zone. Owing to this feature, the generalized Bishop's effective stress is currently considered as the computationally most advantageous one (Nuth and Laloui [38]). Recently, Gray and Schrefler [16, 19] have formulated an effective stress tensor, starting from a more rigorous approach, which considers not only the bulk phases but also the interfaces and their thermodynamic properties at micro- and macro-levels. The main features of this effective stress tensor are: (1) in the formulation the pressures are macroscopic averages over surfaces rather than phases and (2) the present formulation is thermodynamically consistent.

Based on the theory of porous media and thermodynamics, in this paper the work and energy balance equations for unsaturated soils by using geotechnical terminologies are derived, and the premises and assumptions are clearly stated. Then an effective stress is suggested and a principle of generalized effective stress for unsaturated soils is proposed.

2. BALANCE EQUATIONS

Soils are treated as a superposition of three phases of the continuum materials (i.e. solid grain, pore water, gas) and each phase is assumed to continuously occupy the whole space of soil. Under this treatment or assumption, each point in the soil space is simultaneously occupied by the three different continuum materials. Of course, this cannot be physically correct. In this section, the balance equations for a single-phase medium are discussed firstly, followed with those for the multiphase medium.

2.1. Balance equations for each phase of multiphase porous media

As pointed out by de Boer [1], the concept of volume fraction is of significance in porous media theory, thus the mixture theory with the introduction of volume fraction is called as porous media theory. The purpose for introducing the volume fraction is to describe the local structural variation of porous media, which could make the stress of each phase or the corresponding strain be independent state variables instead of microscopic quantities in unsaturated soil. For the three-phase unsaturated soils, the volume fraction is defined as

$$n^\alpha = \frac{V_\alpha}{V} \quad (\alpha = s, w, g) \tag{2}$$

where n^α is the volume fraction for α phase ($\alpha = s, w, g$; s, w and g represent solid grain, pore water and gas, respectively); V_α is the volume of α phase in a representative volume element (RVE), and V is the total volume of the RVE; the relationship between V and V_α is

$$V = V_s + V_w + V_g \tag{3}$$

Dividing Equation (3) by V results in

$$1 = \frac{V_s}{V} + \frac{V_w}{V} + \frac{V_g}{V} = n^s + n^w + n^g = \sum_{\alpha = s, w, g} n^\alpha \tag{4}$$

The average mass density of α phase in the RVE is defined as

$$\rho^\alpha = n^\alpha \rho_\alpha \cdot (\alpha = s, w, g) \tag{5}$$

where ρ_α is the intrinsic mass density of α phase

$$\rho = \rho^s + \rho^w + \rho^g \tag{6}$$

Applying the balance laws of mass, momentum and energy to a single-phase, that is, α phase of unsaturated soils, the following equations can be Obtained [1, 45]:

$$\frac{d\rho^\alpha}{dt} + \rho^\alpha \operatorname{div} \mathbf{v}_\alpha = \hat{\rho}^\alpha \tag{7}$$

$$\operatorname{div} \boldsymbol{\sigma}^\alpha + \rho^\alpha \mathbf{b}^\alpha + \hat{\mathbf{p}}^\alpha = \rho^\alpha \frac{d^\alpha \mathbf{v}_\alpha}{dt} + \hat{\rho}^\alpha \mathbf{v}_\alpha \tag{8}$$

$$\rho^\alpha \frac{d^\alpha \varepsilon^\alpha}{dt} = \operatorname{tr}(\boldsymbol{\sigma}^\alpha \cdot \mathbf{D}^\alpha) - \operatorname{div} \mathbf{q}^\alpha + \rho^\alpha r^\alpha + \hat{e}^\alpha - \hat{\mathbf{p}}^\alpha \cdot \mathbf{v}_\alpha - \hat{\rho}^\alpha \left(\varepsilon^\alpha - \frac{1}{2} \mathbf{v}_\alpha \cdot \mathbf{v}_\alpha \right) \tag{9}$$

where $\hat{\rho}^\alpha$ is the mass supply to α phase from other phases; $\hat{\mathbf{p}}^\alpha$ is the body force acting on α phase by other phases; $\hat{\rho}^\alpha \mathbf{v}_\alpha$, the second item on the right side of Equation (8), represents the momentum supply from other phases due to mass supply $\hat{\rho}^\alpha$; $\mathbf{v}_\alpha$ in Equations (7)–(9) is the velocity (vector) of α phase; $\boldsymbol{\sigma}^\alpha$ is the Cauchy stress of α phase ($\boldsymbol{\sigma}^\alpha = (\boldsymbol{\sigma}^\alpha)^T$), $\mathbf{D}^\alpha = \frac{1}{2}(\mathbf{L}^\alpha + \mathbf{L}^{\alpha T}) = \operatorname{sym}(\operatorname{grad} \mathbf{v}_\alpha)$, $\mathbf{L} = \operatorname{grad} \mathbf{v}_\alpha$; $\mathbf{b}^\alpha$ is the intrinsic body force per unit mass; ε^α is the internal energy of α phase per unit mass, $\mathbf{q}^\alpha$ is the heat flux vector of α phase; r^α is the energy supply (due to radiation) item for α phase; $\hat{e}^\alpha$ is energy supply from other phases. The last two items on the right-hand side of Equation (9) represent the energy supply due to exchanges of mass and momentum between α and other phases.

2.2. *Balance equations for multiphase porous medium*

The mechanical and thermodynamical balance equations and constitutive models of single-phase media developed from the continuum mechanics serve as the basis for the development of multi-phase porous medium theory. However, how to establish the multiphase porous medium theory based on single-phase continuum theory is still in research. Currently, the most well-accepted theory is based on the following three principles proposed by Truesdell [46] to establish the macroscopic governing equations for the porous medium as a whole:

(1) All properties of the mixture must be mathematical consequences of properties of the constituents.
(2) Hence, as to describe the motion of a constituent, we may in imagination isolate it from the rest of the mixture provided we allow properly for the actions of the other constituent upon it.
(3) The motion of the mixture is governed by the same equations as it is a single body.

The above three metaphysical principles state that the balance equations for each constituent still exist, but the actions of other constituent upon it should be taken into account, that is ,there should be an item that accounts for the effect of other constituents. The motion of a mixture as a whole is the superposition of corresponding motion of the single constituent. The governing equations of the mixture should have the same format as those of the single constituent that requires the sum of the actions from other constituents be equal to zero, to satisfy the third principle. The balance equations of each constituent are given in Section 2. These equations can be added up to derive the balance equations of the mixture as a whole. By summing up all the phases for Equations (7)–(9), the following equations can be derived:

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} + \rho \operatorname{div}\mathbf{v} = 0 \tag{10}$$

where $\rho = \sum \rho^{\alpha}$; $\mathbf{v} = \sum (\rho^{\alpha}/\rho)\mathbf{v}_{\alpha}$; $\sum \hat{\rho}^{\alpha} = 0$.

$$\rho \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \operatorname{div}\boldsymbol{\sigma} + \rho \mathbf{b} \tag{11}$$

where

$$\sum \hat{\mathbf{p}}^{\alpha} = 0, \quad \mathbf{b} = \sum \frac{\rho^{\alpha}}{\rho}\mathbf{b}^{\alpha}, \quad \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \sum \left[\frac{\rho^{\alpha}}{\rho}\frac{\mathrm{d}^{\alpha}\mathbf{v}_{\alpha}}{\mathrm{d}t} + \frac{\hat{\rho}^{\alpha}}{\rho}\mathbf{v}_{\alpha}\right], \quad \sum \boldsymbol{\sigma}^{\alpha} = \boldsymbol{\sigma}$$

$$\rho \frac{\mathrm{d}\varepsilon}{\mathrm{d}t} = W + \rho r - \operatorname{div}\mathbf{q} \tag{12}$$

where

$$W = \sum tr(\boldsymbol{\sigma}^{\alpha} \cdot \mathbf{D}^{\alpha}); r = \sum \frac{\rho^{\alpha}}{\rho}r^{\alpha}; \mathbf{q} = \sum \mathbf{q}^{\alpha}; \ \sum \hat{e}^{\alpha} = 0; \frac{\mathrm{d}\varepsilon}{\mathrm{d}t} = \sum \left[\frac{\rho^{\alpha}}{\rho}\frac{\mathrm{d}^{\alpha}\varepsilon^{\alpha}}{\mathrm{d}t} + \frac{\hat{\rho}^{\alpha}}{\rho}\varepsilon^{\alpha} - \frac{1}{2}\hat{\rho}^{\alpha}\mathbf{v}_{\alpha}\cdot\mathbf{v}_{\alpha}\right]$$

The above balance equations lay the foundation for the development of the macroscopic governing equations for describing the physical behaviors of a mixture.

In fact, a more stringent and precise approach should take account of the interfaces [16, 19]. But interfacial properties are generally difficult (if not impossible) to determine, and thus one can expect that the application would be difficult. In an alternative way, which is adopted in this paper, the interfaces are not explicitly considered as the physical components [6, 8]. In this approach, the volume fractions of bulk components are introduced as independent state variables, and as such, the pressure difference between two bulk phases will not vanish. Moreover, with introducing internal variables, this approach can also be used to address some important dissipative phenomena associated with interfaces.

3. EXPRESSION OF WORK AND FREE ENERGY FOR UNSATURATED SOILS

3.1. Expression of work in unsaturated soils

The key step to develop the constitutive equations of unsaturated soils using the thermodynamical approach is to construct two thermodynamic potential functions, the free energy and dissipation functions. Based on these two potential functions, the corresponding incremental constitutive equations can be derived according to Ziegler [47, 48] and Houlsby and Puzrin [49]. The two thermodynamic potential functions are derived from work and energy balance equations. The work expression in unsaturated soils should be first discussed.

Hansen [50] pointed out that the various constrained equations derived from the three metaphysical principles proposed by Trusdell [46] are not unified mathematically. In addition, it is quite difficult to give reasonable explanations in physics for the quantity determination of some variables in the constrained equations. Therefore, it is necessary to do some further research on the definition of the variables of a mixture as a whole as well as the corresponding constrained equations. For example, in theory of mixture the velocity defined in the second term of Equation (10) is a mass-weighted average quantity. But in soil mechanics or hydrology, the velocity of soil skeleton $\mathbf{v}_s$ is generally used as a reference configuration to construct seepage or other equations. The equations based on the theory of mixture have different formats with the ones developed in some engineering fields, such as soil mechanics and hydrology, due to the selection of different basic kinematical variables. In this section, the work expression and some physical parameters of unsaturated soils are derived using the variables and parameters commonly used in soil mechanics. Based on the work expression, the energy equation and the effective stress of unsaturated soil will be discussed.

The basic assumptions are made for unsaturated soils

(1) The three phases, solid, water and gas, have same temperature at the same point in unsaturated soil.
(2) The transfer of mass, entropy and internal energy among the three phases is ignored.
(3) There are obvious interfaces among the three phases, and they are immiscible with each other.
(4) The velocities of seepage and airflow are sufficiently small such that the diffusion effects on internal energy, stress, heat and entropy are all negligible.

The velocity of α fluid phase relative to the solid phase is defined as:

$$\bar{\mathbf{v}}_\alpha = \mathbf{v}_\alpha - \mathbf{v}_s \quad (\alpha = w, g) \tag{13}$$

The $\mathbf{v}_\alpha$ in Equation (13) is the velocity of α fluid phase (true velocity of the fluid), and the volume fraction in unsaturated soil can be calculated by

$$n^s = 1 - n; \quad n^w = nS_r; \quad n^g = n(1 - S_r) \tag{14}$$

where n is the porosity of soil and S_r is the degree of saturation.

Volume fraction can be used to describe the internal structure change of porous medium and forms the basis for analyzing the dissipation mechanism and corresponding macroscopic irreversible deformation. Taking soil skeleton as a reference configuration, the mass balance law, Equation (9), for a single-phase can be expressed as

$$\frac{\mathrm{d}\rho^\alpha}{\mathrm{d}t} + \rho^\alpha \operatorname{div} \mathbf{v}_s = -\rho^\alpha \operatorname{div}(\mathbf{v}_\alpha - \mathbf{v}_s) = -\rho^\alpha \operatorname{div}(\bar{\mathbf{v}}_\alpha) \tag{15}$$

The bulk modulus of the α phase can be defined as $K_\alpha = \rho_\alpha \mathrm{d}P_\alpha / \mathrm{d}\rho_\alpha$, where P_α is the pore fluid pressure of α phase. The mass balance equation expressed by the volume fraction and pressure P_α is

$$\frac{\mathrm{d}\rho^\alpha}{\mathrm{d}t} = \frac{\mathrm{d}(n^\alpha \rho_\alpha)}{\mathrm{d}t} = n^\alpha \frac{\mathrm{d}\rho_\alpha}{\mathrm{d}t} + \rho_\alpha \frac{\mathrm{d}n^\alpha}{\mathrm{d}t} = n^\alpha \frac{\mathrm{d}\rho_\alpha}{\mathrm{d}P_\alpha} \frac{\mathrm{d}P_\alpha}{\mathrm{d}t} + \rho_\alpha \frac{\mathrm{d}n^\alpha}{\mathrm{d}t}$$

Substituting Equation (15) into the above equation, we get

$$\frac{n^\alpha}{K_\alpha} \frac{\mathrm{d}P_\alpha}{\mathrm{d}t} + \frac{\mathrm{d}n^\alpha}{\mathrm{d}t} + n^\alpha \operatorname{div} \mathbf{v}_s = -n^\alpha \operatorname{div} \bar{\mathbf{v}}_\alpha \quad (\alpha = s, w, g) \tag{16}$$

The total deformation work for unsaturated soils as a whole with the velocity of the soil skeleton as a reference configuration is derived as follows. Equation (12) is used as the general energy balance equation, and the first item on the right side of the equation is the deformation work W for unsaturated soils as a whole. According to the fourth equation in Equation (12), the following equation holds.

$$\begin{aligned}
W &= \sum \operatorname{tr}(\boldsymbol{\sigma}^\alpha \cdot \mathbf{D}^\alpha) = \sum \operatorname{tr}(\boldsymbol{\sigma}^\alpha \cdot \operatorname{grad} \mathbf{v}_\alpha) \\
&= \operatorname{tr}(\boldsymbol{\sigma}^s \cdot \operatorname{grad} \mathbf{v}_s) + \operatorname{tr}(\boldsymbol{\sigma}^w \cdot \operatorname{grad} \mathbf{v}_w) + \operatorname{tr}(\boldsymbol{\sigma}^g \cdot \operatorname{grad} \mathbf{v}_g) \\
&= \operatorname{tr}[(\boldsymbol{\sigma} - \boldsymbol{\sigma}^w - \boldsymbol{\sigma}^g) \cdot \operatorname{grad} \mathbf{v}_s] + \operatorname{tr}(\boldsymbol{\sigma}^w \cdot \operatorname{grad} \mathbf{v}_w) + \operatorname{tr}(\boldsymbol{\sigma}^g \cdot \operatorname{grad} \mathbf{v}_g) \\
&= \operatorname{tr}(\boldsymbol{\sigma} \cdot \operatorname{grad} \mathbf{v}_s) + \operatorname{tr}[\boldsymbol{\sigma}^w \cdot \operatorname{grad}(\mathbf{v}_w - \mathbf{v}_s)] + \operatorname{tr}[\boldsymbol{\sigma}^g \cdot \operatorname{grad}(\mathbf{v}_g - \mathbf{v}_s)] \\
&= \operatorname{tr}(\boldsymbol{\sigma} \cdot \operatorname{grad} \mathbf{v}_s) + \operatorname{tr}(\boldsymbol{\sigma}^w \cdot \operatorname{grad} \bar{\mathbf{v}}_w) + \operatorname{tr}(\boldsymbol{\sigma}^g \cdot \operatorname{grad} \bar{\mathbf{v}}_g)
\end{aligned} \tag{17}$$

The first line deduced in Equation (17) can refer to Hutter [40, p. 74, note19]. The relationship between the fluid stress $\boldsymbol{\sigma}^\alpha$ and corresponding fluid pressure is

$$\boldsymbol{\sigma}^\alpha = n^\alpha P_\alpha \boldsymbol{\delta} \quad (\alpha = w, g) \tag{18}$$

P_α is a scalar representing the pore fluid pressure. Note: according to the convention in soil mechanics, compression is considered as positive and tension as negative, thus a minus sign should be added to the right side of Equation (18). Substituting Equation (18) into Equation (17) yields

$$W = \operatorname{tr}(\boldsymbol{\sigma} \cdot \operatorname{grad} \mathbf{v}_s) + n^w P_w \operatorname{tr}(\operatorname{grad} \bar{\mathbf{v}}_w \cdot \delta) + n^g P_g \operatorname{tr}(\operatorname{grad} \bar{\mathbf{v}}_g \cdot \delta) \tag{19}$$

As $\mathrm{tr}(\mathrm{grad}\,\bar{\mathbf{v}}_\alpha\cdot\boldsymbol{\delta})=\mathrm{div}\,\bar{\mathbf{v}}_\alpha$, $(\alpha=w,g)$, Equation (19) can be rewritten as

$$W=\mathrm{tr}(\boldsymbol{\sigma}\cdot\mathrm{grad}\,\mathbf{v}_s)+n^w P_w\,\mathrm{div}\,\bar{\mathbf{v}}_w+n^g P_g\,\mathrm{div}\,\bar{\mathbf{v}}_g \tag{20}$$

Substituting Equation (16) into Equation (20) yields

$$
\begin{aligned}
W &=\sum\mathrm{tr}(\boldsymbol{\sigma}^\alpha\cdot\mathbf{D}^\alpha)\\
&=\mathrm{tr}(\boldsymbol{\sigma}\cdot\mathrm{grad}\,\mathbf{v}_s)-\left[P_w\frac{n^w}{K_w}\frac{\mathrm{d}P_w}{\mathrm{d}t}+P_w\frac{\mathrm{d}n^w}{\mathrm{d}t}+n^w P_w\,\mathrm{div}\,\mathbf{v}_s\right]\\
&\quad-\left[P_g\frac{n^g}{K_g}\frac{\mathrm{d}P_g}{\mathrm{d}t}+P_g\frac{\mathrm{d}n^g}{\mathrm{d}t}+n^g P_g\,\mathrm{div}\,\mathbf{v}_s\right]\\
&=\underbrace{\mathrm{tr}(\boldsymbol{\sigma}\cdot\mathrm{grad}\,\mathbf{v}_s)}_{\text{I}}-\underbrace{\left(P_w\frac{n^w}{K_w}\frac{\mathrm{d}P_w}{\mathrm{d}t}+P_g\frac{n^g}{K_g}\frac{\mathrm{d}P_g}{\mathrm{d}t}\right)}_{\text{II}}\\
&\quad-\underbrace{\left(P_w\frac{\mathrm{d}n^w}{\mathrm{d}t}+P_g\frac{\mathrm{d}n^g}{\mathrm{d}t}\right)}_{\text{III}}-\underbrace{(n^w P_w+n^g P_g)\,\mathrm{div}\,\mathbf{v}_s}_{\text{IV}}
\end{aligned}
\tag{21}
$$

Equation (21) consists of four parts: I, II, III and IV. The four parts can be simplified and explained below.

Keeping parts I and II constant and substituting Equation (14) into part III yield

$$\text{III}=P_w\frac{\mathrm{d}(nS_r)}{\mathrm{d}t}+P_g\frac{\mathrm{d}[n(1-S_r)]}{\mathrm{d}t}=\underbrace{n(P_w-P_g)\frac{\mathrm{d}S_r}{\mathrm{d}t}}_{\text{IIIa}}+\underbrace{[S_r P_w+(1-S_r)P_g]\frac{\mathrm{d}n}{\mathrm{d}t}}_{\text{IIIb}}$$

Substituting Equation (14) and (16) (when $\alpha=\mathrm{s}$) into part IV yields

$$
\begin{aligned}
\text{IV}&=n[S_r P_w+(1-S_r)P_g]\,\mathrm{div}\,\mathbf{v}_s=(1-n_s)[S_r P_w+(1-S_r)P_g]\,\mathrm{div}\,\mathbf{v}_s\\
&=[S_r P_w+(1-S_r)P_g]\,\mathrm{div}\,\mathbf{v}_s-n_s[S_r P_w+(1-S_r)P_g]\,\mathrm{div}\,\mathbf{v}_s\\
&=[S_r P_w+(1-S_r)P_g]\,\mathrm{div}\,\mathbf{v}_s-[S_r P_w+(1-S_r)P_g]\left(-\frac{n^s}{K_s}\frac{\mathrm{d}P_s}{\mathrm{d}t}+\frac{\mathrm{d}n^s}{\mathrm{d}t}\right)\\
&=\underbrace{[S_r P_w+(1-S_r)P_g]\,\mathrm{div}\,\mathbf{v}_s}_{\text{IVa}}+\underbrace{[S_r P_w+(1-S_r)P_g]\frac{n^s}{K_s}\frac{\mathrm{d}P_s}{\mathrm{d}t}}_{\text{IVb}}-\underbrace{[S_r P_w+(1-S_r)P_g]\frac{\mathrm{d}(1-n)}{\mathrm{d}t}}_{\text{IVc}}
\end{aligned}
$$

Combining part IVa with part I as well as part IVb with part II and canceling part IIIb and part IVc, Equation (21) can be formatted as

$$W=\left\{\underbrace{\mathrm{tr}(\boldsymbol{\sigma}\cdot\mathrm{grad}\,\mathbf{v}_s)}_{\text{I}}-\underbrace{[S_r P_w+(1-S_r)P_g]\,\mathrm{div}\,\mathbf{v}_s}_{\text{IVa}}\right\}-\underbrace{n(P_w-P_g)\frac{\mathrm{d}S_r}{\mathrm{d}t}}_{\text{IIIa}}$$

$$-\left\{\underbrace{\left[P_w\frac{n^w}{K_w}\frac{\mathrm{d}P_w}{\mathrm{d}t}+P_g\frac{n^g}{K_g}\frac{\mathrm{d}P_g}{\mathrm{d}t}\right]}_{\text{II}}+\underbrace{[S_rP_w+(1-S_r)P_g]\frac{n^s}{K_s}\frac{\mathrm{d}P_s}{\mathrm{d}t}}_{\text{IVb}}\right\}$$

$$=\mathrm{tr}\{\{\boldsymbol{\sigma}-[S_rP_w+(1-S_r)P_g]\boldsymbol{\delta}\}\cdot\mathrm{grad}\,\mathbf{v}_s\}-n(P_w-P_g)\frac{\mathrm{d}S_r}{\mathrm{d}t}$$

$$-\left\{P_w\frac{n^w}{K_w}\frac{\mathrm{d}P_w}{\mathrm{d}t}+P_g\frac{n^g}{K_g}\frac{\mathrm{d}P_g}{\mathrm{d}t}+[S_rP_w+(1-S_r)P_g]\frac{n^s}{K_s}\frac{\mathrm{d}P_s}{\mathrm{d}t}\right\} \tag{22}$$

The stress in the first and second parenthesis on the right side of equation (22) can be, respectively, defined as

$$\boldsymbol{\sigma}'=\boldsymbol{\sigma}-[S_rP_w+(1-S_r)P_g]\boldsymbol{\delta} \tag{23}$$

$$-n(P_w-P_g)\frac{\mathrm{d}S_r}{\mathrm{d}t}=sn\frac{\mathrm{d}S_r}{\mathrm{d}t} \tag{24}$$

The work in the first parenthesis on the right side of Equation (22) is the product of the stress $\boldsymbol{\sigma}'$ and the deformation rate of soil skeleton. It is the work done by the driving stress, $\boldsymbol{\sigma}'$ for the deformation of the soil skeleton. The driving stress $\boldsymbol{\sigma}'$ is called as the average soil skeleton stress, which is undertaken and transmitted by the soil skeleton. In the second parenthesis $s(=P_g-P_w)$ is the matric suction in unsaturated soils. It is a generalized driving force, and the volume fraction for the liquid $n^w=nS_r$ is its corresponding generalized flow. The item in the second parenthesis represents the work done by matric suction. The third parenthesis in Equation (22) represents the work done by the variation of the pressure of each phase. By assuming that the solid and water are incompressible, the term in third parenthesis in Equation (22) reduces to

$$\left[-\frac{P_g}{K_g}n(1-S_r)\frac{\mathrm{d}P_g}{\mathrm{d}t}\right]$$

representing the work done by the pressure variation of the gas, it can be deduced as:

$$-n(1-S_r)\frac{P_g}{K_g}\frac{\mathrm{d}P_g}{\mathrm{d}t}=-n^g\frac{P_g}{\rho_g}\frac{\mathrm{d}\rho_g}{\mathrm{d}t}=-n^gP_g\frac{\dot{\rho}_g}{\rho_g}=n^gP_g\left(\mathrm{div}\,\mathbf{v}_g+\frac{\dot{n}^g}{n^g}\right) \tag{25}$$

According to the mass balance equation (7) and Equation (5), when $\alpha=\mathrm{a}$, the following two equations can be obtained:

$$\dot{\rho}^g+\rho^g\,\mathrm{div}\,\mathbf{v}_g=0\Rightarrow\frac{\dot{\rho}^g}{\rho^g}=-\mathrm{div}\,\mathbf{v}_g$$

$$\frac{\dot{\rho}^g}{\rho^g}=\frac{\dfrac{\mathrm{d}}{\mathrm{d}t}(n^g\rho_g)}{n^g\rho_g}=\frac{\dot{n}^g\rho_g+n^g\dot{\rho}_g}{n^g\rho_g}=\frac{\dot{n}^g}{n^g}+\frac{\dot{\rho}_g}{\rho_g}$$

Then, we have

$$\frac{\dot{\rho}_g}{\rho_g}=\frac{\dot{\rho}^g}{\rho^g}-\frac{\dot{n}^g}{n^g}=-\mathrm{div}\,\mathbf{v}_g-\frac{\dot{n}^g}{n^g}$$

If $\dot{n}^g / n^g$ is small enough to be ignored, Equation (25) can be rewritten as follows:

$$-n(1-S_r)\frac{P_g}{K_g}\frac{\mathrm{d}P_g}{\mathrm{d}t} = P_g(n^g \operatorname{div} \mathbf{v}_g) \tag{26}$$

3.2. Relationships between free energy and work

Ignoring the effect of temperature variation, an isothermal energy balance equation for unsaturated soils can be derived from the energy balance equation (12). According to the aforementioned second assumption, the transfer of mass, momentum and energy among the three phases can be neglected, that is, $\hat{\rho}_\alpha$, $\hat{\mathbf{p}}_\alpha$ and $\hat{e}_\alpha$ vanish. Then Equation (12) reduces to

$$\rho^\alpha \frac{\mathrm{d}^\alpha \varepsilon^\alpha}{\mathrm{d}t} = \operatorname{tr}(\boldsymbol{\sigma}^\alpha \cdot \mathbf{D}^\alpha) - \operatorname{div}\mathbf{q}^\alpha + \rho^\alpha r^\alpha \tag{27}$$

According to de Boer [1], the second law of thermodynamics can be expressed as:

$$\sum_\alpha \left[\rho^\alpha \frac{\mathrm{d}^\alpha \eta^\alpha}{\mathrm{d}t} + \hat{\rho}^\alpha \eta^\alpha - \frac{1}{\theta}\rho^\alpha r^\alpha + \operatorname{div}\left(\frac{\mathbf{q}^\alpha}{\theta}\right) \right] \geqslant 0$$

It can also be reformatted as:

$$\sum_\alpha \left[\theta\rho^\alpha \frac{\mathrm{d}^\alpha \eta^\alpha}{\mathrm{d}t} - \rho^\alpha r^\alpha + \operatorname{div}\mathbf{q}^\alpha - \frac{1}{\theta}\mathbf{q}^\alpha \operatorname{grad}\theta \right] \geqslant 0 \tag{28}$$

where η^α is the specific entropy for the α phase and θ is the absolute temperature ($\theta \geqslant 0$).

Combining Equations (27) and (28) results in the following inequality:

$$\sum_\alpha \left[\rho^\alpha \theta \frac{\mathrm{d}^\alpha \eta^\alpha}{\mathrm{d}t} - \rho^\alpha \frac{\mathrm{d}^\alpha \varepsilon^\alpha}{\mathrm{d}t} + \operatorname{tr}(\boldsymbol{\sigma}^\alpha \cdot \mathbf{D}^\alpha) - \frac{1}{\theta}\mathbf{q}^\alpha \operatorname{grad}\theta \right] \geqslant 0 \tag{29}$$

Based on the transformation law among the potential functions in thermodynamics, the Helmholtz free energy of unit volume is expressed as follows:

$$A^\alpha = \varepsilon^\alpha - \theta\eta^\alpha \tag{30}$$

Ignoring the effect of temperature variation with time t, the material time derivative for Equation (30) is given as:

$$\frac{\mathrm{d}^\alpha A^\alpha}{\mathrm{d}t} = \frac{\mathrm{d}^\alpha \varepsilon^\alpha}{\mathrm{d}t} - \theta\frac{\mathrm{d}^\alpha \eta^\alpha}{\mathrm{d}t}$$

Then Equation (29) can be rewritten as

$$-\frac{\mathrm{d}A}{\mathrm{d}t} + W - \frac{1}{\theta}\mathbf{q}\cdot\operatorname{grad}\theta \geqslant 0 \tag{31}$$

where

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \sum_\alpha \left(\rho^\alpha \frac{\mathrm{d}^\alpha A^\alpha}{\mathrm{d}t} \right)$$

W represents work as shown in Equation (22), and $\mathbf{q}$ represents the heat flux vector of soil mixture as shown in Equation (12).

Following Houlsby and Puzrin [49], the first two items of Equation (31) can be defined as dissipation function D:

$$D = -\frac{\mathrm{d}A}{\mathrm{d}t} + W \tag{32}$$

The third item of Equation (31) $\{-(1/\theta)\mathbf{q}\cdot\mathrm{grad}\,\theta\}$ is heat dissipation and is always non-negative since heat always flows in the direction opposite to temperature gradient. Compared with mechanical dissipation (with regard to unsaturated soil mechanics), the heat dissipation is relatively small. As such, $D\geqslant 0$ is regarded as reasonable, therefore,

$$D = -\frac{\mathrm{d}A}{\mathrm{d}t} + W \geqslant 0$$

$$-\frac{1}{\theta}\mathbf{q}\cdot\mathrm{grad}\,\theta \geqslant 0$$

The above two equations are more rigorous restrictions than the second law of thermo-dynamics and have been well accepted (Houlsby and Puzrin [49]). Equation (32) can be formatted as follows:

$$W = \frac{\mathrm{d}A}{\mathrm{d}t} + D \tag{33}$$

where W represents the work as defined by Equation (22), A represents Helmholtz free energy functions and D represents the dissipation function. Equation (33) shows that the total deformation work in the isothermal process of unsaturated soils is equal to the sum of the rate of the Helmholtz free energy and the dissipation function, which is useful for the development of constitutive models of unsaturated soils such as Li [17].

4. EFFECTIVE STRESS FOR UNSATURATED SOILS

The effective stress principle for saturated soils has brought great innovations in the soil mechanics and geotechnical engineering practice. However, its applicability to unsaturated soils is still under debate among numerous researchers. In the work by Jommi [43], Vaunat *et al.* [44], Wheeler *et al.* [36], and Nuth and Laloui [38], an effective stress was proposed on the basis of the work hypothesis. The generalized stress work-conjugated with strain of solid skeleton was defined as the effective stress for unsaturated soils (also named as average soil skeleton stress).

In accordance with the above method for the definition of effective stress, it is suggested that the effective stress for unsaturated soils is defined by Equation (23), that is, $\boldsymbol{\sigma}' = \boldsymbol{\sigma} - [S_r P_w + (1 - S_r)P_g]\boldsymbol{\delta}$, which is a generalized stress work-conjugated with the displacement of solid skeleton determined by the total deformation work in Equation (22) and is the effective stress for unsaturated soils suggested in this paper. Equation (23) indicates that the effective stress is related to the stress of each phase as well as its volume fraction implying that the effective stress is dependent on the material geometry properties. This suggested effective stress is different from Cauchy stress in a single-phase continuum mechanics in which the stress variable is assumed independent of the material geometry properties. However, from the viewpoints in the mixture theory, there exists

a natural relationship between effective stress and material geometry properties, which suggests that the stress is influenced by the variation of microscopic structure and component volume in unsaturated soils.

Compared with the single-effective-stress variable theory [20], the effective stress proposed in this paper (i.e. Equation (23)) involves the saturation degree S_r other than the parameter χ. The S_r and $(1-S_r)$ can be physically explained as the weighing parameters for water and gas pressure respectively. Therefore, the effective stress equals to the total stress minus the comprehensive pore pressure combined by the water pressure and gas pressure with the corresponding volume fraction as weighing parameters. As a kind of effective stress, the combination of these three stresses is undertaken and transmitted by the soil skeleton. When the saturation degree $S_r = 1$, the effective stress of unsaturated soil (Equation (23)) reduces to the effective stress of saturated soil:

$$\boldsymbol{\sigma}' = \boldsymbol{\sigma} - [S_r P_w + (1-S_r) P_g]\boldsymbol{\delta} = \boldsymbol{\sigma} - P_w\boldsymbol{\delta}$$

It should be pointed out that the effective stress presented in this paper is identical to the one used by Wheeler *et al.* [36], Jommi [43], etc. Their expression is based on the deformation work of unsaturated soils given by Houlsby [42]. Nevertheless, the effective stress proposed in this paper is based on the deformation work deduced in the porous media theory. It is noted that Equation (23) is similar to the effective stress expression given by Lewis and Schrefler [4] and Schrefler [5]. Clearly, Equation (23) can be considered as a simplification of the one proposed by Gray and Schrefler [16] based on a more rigorous approach. However, Equation (23) is easier to apply.

5. THE GENERALIZED EFFECTIVE STRESS PRINCIPLE FOR UNSATURATED SOILS

As aforementioned, neither single effective stress variable theory nor two-stress variable theory can well describe all the properties and behaviors of unsaturated soils. The behaviors of unsaturated soils are not only determined by the two stresses, but also affected by the degree of saturation and the variation of gas pressure etc. Jommi [43] suggested that 'for the development of constitutive models, strain variables should be taken into account too.'

The theory on unsaturated soils has been developing from simple to complex. The theory on effective stress principle for unsaturated soils should be developed in the same way. First, a single variable of the effect stress for unsaturated soils was proposed. The limitations of single stress variable were noticed later. Then the more complicated theory, that is, the two-stress variable theory, was put forward. Recently, Wheeler *et al.* [36] applied two-stress variables and an additional variable (the degree of saturation) to describe the properties of unsaturated soils. Now it has stepped into the stage that multiphase porous medium theory can be employed to comprehensively and objectively describe unsaturated soils behaviors in terms of both general stress variables of three phases of unsaturated soils and their work-conjugated strains.

According to Ziegler's [48] and Houlsby and Puzrin's [49] theory, the stresses and its corresponding elastic and plastic deformations in soils can be derived using free energy functions and dissipation functions given in balance equation (33), but their work is only based on single media or saturated soils. Li [17] gave an elastoplastic framework for three-phase unsaturated soils based on Equations (33) and (22). From Li's [17] viewpoint, Equations (33) and (22) are key equations to describe the fundamental properties and to establish elastoplastic constitutive equations for unsaturated soils. Equation (22) expresses the deformation work of unsaturated soils by the product of generalized stress and its conjugated generalized deformation. From the equality between

Equations (22) and (33), it can be seen that various stresses in unsaturated soils and corresponding deformations (including elastic and plastic deformations) are affected by all the items in the right side of equation (22), in which the first item represents the work done by the product of the effective stress undertaken by solid skeleton and the deformation of the solid skeleton, the second item represents the work done by the product of matric suction and the variation of the degree of saturation, and the third item represents the work done by the product of gas volume and variation of gas pressure.

Based on the above discussion, a principle of generalized effective stress for unsaturated soils is proposed as follows: it should be based on the deformation work equation (22) and consider the effects of all the three phases including all the generalized stress and its conjugated generalized deformation in Equation (22), as written explicitly as follows:

$$
\begin{bmatrix} \boldsymbol{\sigma}' = \boldsymbol{\sigma} - [S_r P_w + (1 - S_r) P_g] \boldsymbol{\delta} \\ s \\ P_g \end{bmatrix} \quad \text{and} \quad \begin{pmatrix} \operatorname{grad} \mathbf{v}_s \\ n \dfrac{\mathrm{d} S_r}{\mathrm{d} t} \\ n^g \left(\operatorname{div} \mathbf{v}_g + \dfrac{\dot{n}^g}{n^g} \right) \end{pmatrix} \tag{34}
$$

The single stress variable of effective stress only considers Bishop-stress and its work-conjugated strain and the effects of the work done by last two items are ignored, then the big errors may occur and the relationship between deformation and the effective stress cannot be represented uniquely. The effective stress derived in this way will certainly have limitations. And although the two-stress variables theory considers more variables, it is still difficult to establish a complete and unique constitutive relationship. Hence, Wheeler *et al.* [36] considered the effect of the degree of saturation in addition to two-stress variables. His work has already contained all the first two items in the right of Equation (22), but the effect of the third item was still not taken into account. As mentioned above, neither single stress variable nor two-stress variables theory can well describe the properties and behaviors of unsaturated soils in a scientific and rigorous way. Some other factors such as gas pressure and its volume must be taken into account according to the principle of generalized effective stress proposed in the paper, which may be of importance in case of occluded pores or in case of reservoir analysis [4].

In fact, in the unsaturated soil experiments, the pore gas pressure in a sample is actually considered to be equal to the gas pressure on the sample surface (usually equal to the atmospheric pressure). But if gas in the pore space is occluded, the gas can hardly flow out of the pore space. For this reason, when external forces act on the sample, the gas pressure in the occluded pore space will not be equal to that on the sample surface (even if it has gone through a longtime equalization process). In this case, if the pore gas pressure in the sample is assumed to be equal to that on the sample surface, the test result will depart from the fact. When external forces act on the sample, the pore gas pressure in the occluded pore space will change resulting in the variation of the bulk modulus of the gas and the behaviors of unsaturated soils. The variation of gas pressure in the occluded pore space will change the soil compressibility, and the assumption that the gas pressure variation will not change the behaviors of unsaturated soils is contradictory to the real situation. Accordingly, if the two independent stress variables are adopted and the effect of the variation of gas pressure is neglected, the same errors may also occur. Li [17] has realized the limitations of two-stress variable theory of unsaturated soils. In his paper, the effect of gas phase has been

included in the constitutive framework; however, the pore gas was treated as ideal one and the gas dissipation was not reflected.

The following question may arise: as there are already complicated theories, such as two-stress variables, for unsaturated soils, is it needed to develop a more complicated one including more stresses and more factors? The answer is that the purpose of scientific research is to provide proper theories to describe the features and laws of nature accurately and objectively. If the true features and laws are complicated in nature and affected by many factors, the theory should consider these facts and contain sufficient factors in order to reflect the complicated reality. Once a more complicated scientific theory that contains more factors for reflecting complicated reality is proposed, it is easy to be applied to engineering problem by neglecting some secondary factors, and the errors caused by neglecting the secondary factors can also be evaluated by the theory. However, a simplified theory cannot fulfill this task, because many factors have already been neglected in the premises and assumptions. This is the advantages brought by the development of more complicated and comprehensive theory.

6. CONCLUSION

Owing to the complexity of the physical mechanism and the influencing factors, the behaviors of unsaturated soils are difficult to be described uniquely in terms of the single effective stress variable or two-stress variables. On the basis of the porous media theory, the energy balance equations for three-phase unsaturated soils with energy dissipation are derived, and the total deformation work W is presented in terms of geotechnical terminologies for unsaturated soils. The effective stress for unsaturated soils is defined as $\sigma' = \sigma - [S_r P_w + (1 - S_r) P_g]\delta$, which is a generalized stress work-conjugated with the displacement of solid skeleton determined by the derived expression of total deformation work.

The generalized effective stress principle takes into account of more influential factors (including variation of gas pressure, saturation degree, etc.) than the previous single effective stress variable or two-stress variable theories. The newly proposed principle of generalized effective stress may reduce to the previous effective stress theory of single-stress variable or the two-stress variables under certain conditions. As an appropriate selection of stress state variables is crucial for the development of constitutive models, this framework of the principle of generalized effective stress may provide a helpful reference for developing more accurate and comprehensive constitutive models for unsaturated soils. Its application will be discussed in another paper.

REFERENCES

1. de Boer R. *Theory of Porous Media*: *Highlights in Historical Development and Current State*. Springer: Berlin, 2000; 345–353.
2. Bowen RM. Incompressible porous media models by use of the theory of mixtures. *International Journal of Engineering Science* 1980; **18**:1129–1148.

3. Bowen RM. Compressible porous media models by use of the theory of mixtures. *International Journal of Engineering Science* 1982; **20**(5):691–735.
4. Lewis RW, Schrefler BA. A finite element simulation of the subsidence of gas reservoirs undergoing a water drive. In *Finite Elements in Fluids*, Gallagher RH, Norrie DH, Oden JT, Zienkiewicz OC (eds), vol. 4. Wiley: Chichester, 1982; 179–199.
5. Schrefler BA. The finite element method in soil consolidation (with application to surface subsidence). *Ph.D. Thesis*, University College of Swansea, 1984.
6. Hassanizadeh SM, Gray WG. Mechanics and thermodynamics of multiphase flow in porous media including interphase boundaries. *Advances in Water Resources* 1990; **13**(4):169–186.
7. Hassanizadeh SM, Gray WG. Thermodynamic basis of capillary pressure in porous media. *Water Resources Research* 1993; **29**:3389–3405.
8. Gray WG, Hassanizadeh SM. Unsaturated flow theory including interfacial phenomena. *Water Resources Research* 1991; **27**:1855–1863.
9. Eringen AC. Continuum theory of swelling porous elastic soils. *International Journal of Engineering Science* 1994; **32**(7):1337–1349.
10. Hutter K, Laloui L, Vullie L. Thermodynamically based mixture models of saturated and unsaturated soils. *Mechanics of Cohesive Frictional Materials* 1999; **4**:295–338.
11. Schrefler BA. Mechanics and thermodynamics of saturated/unsaturated porous materials and quantitative solutions. *Applied Mechanics Reviews* 2002; **55**(4):351–388.
12. Laloui L, Klubertanz G, Vulliet L. Solid–liquid–air coupling in multiphase porous media. *International Journal for Numerical and Analytical Methods in Geomechanics* 2003; **27**:183–206.
13. Borja RI. Cam-clayplasticity. Part V: a mathematical framework for three-phase deformation and strain localization analyses of partially saturated porous media. *Computer Methods in Applied Mechanics and Engineering* 2004; **193**:5301–5338.
14. Sheng DC, Sloan SW, Gens A. A constitutive model for unsaturated soils: thermomechanical and algorithmic aspects. *Computational Mechanics* 2004; **33**:453–465.
15. Wei CF, Dewoolkar MM. Formulation of capillary hysteresis with internal state variables. *Water Resources Research* 2006; **42**:w07405.
16. Gray WG, Schrefler BA. Analysis of the solid phase stress tensor in multiphase porous media. *International Journal for Numerical and Analytical Methods in Geomechanics* 2007; **31**(4):541–581.
17. Li XS. Thermodynamics-based constitutive framework for unsaturated soils 1: theory. *Géotechnique* 2007; **57**(5):411–422.
18. Li XS. Thermodynamics-based constitutive framework for unsaturated soils 2: a basic triaxial model. *Géotechnique* 2007; **57**(5):423–435.
19. Gray WG, Schrefler BA, Pesavento F. The solid phase stress tensor in porous media mechanics and the Hill–Mandel condition. *Journal of the Mechanics and Physics of Solids* 2009; **57**:539–554.
20. Bishop AW. The principle of effective stress. *Teknisk Ukeblad* 1959; **106**(39):859–863.
21. Aitchison GD, Donald IB. Some preliminary studies of unsaturated soils. *Proceedings of the 2nd ANZ Conference on SMFE, Technical Publications for the New Zealand Institution of Engineers*, Wellington, New Zealand, 1956; 192–199.
22. Jennings JEB. A revised effective stress law for use in the prediction of the behaviour of unsaturated soils. *Pore Pressure and Suction in Soils*. Butterworth: London, 1961; 26–30.
23. Bishop AW, Donald IB. The experimental study of partly saturated soils in triaxial apparatus. *Proceedings of the 5th International Conference on Soil Mechanics and Foundation Engineering*. Dunod: Paris, 1961; 13–21.
24. Bishop AW, Blight GE. Some aspects of effective stress in saturated and partly saturated soils. *Géotechnique* 1963; **13**(3):177–197.
25. Blight GE. A study of effective stress for volume change. *Proceedings of Symposium on Moisture Equilibrium and Moisture Changes in Soils Beneath Covered Areas*. Butterworths: Sydney, Australia, 1965; 259–269.
26. Jennings JEB, Burland JB. Limitations to the use of effective stresses in unsaturated soils. *Géotechnique* 1962; **12**:125–144.
27. Burland JB. Some aspects of the mechanical behaviour of partly saturated soils. In *Moisture Equilibria and Moisture Changes in the Soils Beneath Areas*, Aitchison GD (ed.). Butterworth: Sydney, Australia, 1965; 270–278.
28. Aitchison GD. Soils properties, shear strength, and consolidation. *Proceedings of 6th International Conference on Soil Mechanics and Foundation Engineering*. University of Toronto Press: Toronto, 1965; 319–321.
29. Matyas EL, Radhakrishna HS. Volume change characteristics of partially saturated soils. *Géotechnique* 1968; **18**:432–448.

30. Brackley IJA. Partial collapse in unsaturated expansive clay. *Proceedings of 5th Regional Conference on Soil Mechanics and Foundation Engineering*, South Africa, 1971; 23–30.
31. Fredlund DG, Morgenstern NR. Stress state variables for unsaturated soils. *Journal of Geotechnical Engineering Division* (ASCE) 1977; **103**(GT5):447–466.
32. Wheeler SJ, Karube D. Constitutive modelling. *Proceedings of the 1st International Conference on Unsaturated Soils*, Paris, vol. 3, 1996; 1323–1356.
33. Khalili N, Geiser F, Blight GE. Effective stress in unsaturated soils: review with new evidence. *International Journal of Geomechanics* 2004; **4**(2):115–126.
34. Coleman JD. Stress–strain relations for partially saturated soils. *Géotechnique* 1962; **12**(4):348–350.
35. Blight GE. Effective stress evaluation for unsaturated soils. *Journal of the Soil Mechanics and Foundations Division* 1967; **93**(2):125–148.
36. Wheeler SJ, Sharma RS, Buisson MSR. Coupling of hydraulic hysteresis and stress–strain behaviour in unsaturated soils. *Géotechnique* 2003; **53**(1):41–54.
37. Sun DA, Sheng DC, Cui HB, Sloan SW. A density-dependent elastoplastic hydro-mechanical model for unsaturated compacted soils. *International Journal for Numerical and Analytical Methods in Geomechanics* 2007; **31**(10): 1257–1279.
38. Nuth M, Laloui L. Effective stress concept in unsaturated soils: clarification and validation of a unified framework. *International Journal for Numerical and Analytical Methods in Geomechanics* 2008; **32**(6):771–801.
39. Taramtino A, Mongiovì L. Experimental investigations on the stress variables governing unsaturated soil behaviour at medium to high degrees of saturation. In *Experimental Evidence and Theoretical Approaches in Unsaturated Soils*, Tarantino A, Mancuso C (eds). Balkema: Rotterdam, 2000; 3–19.
40. Bocking KA, Fredlund DG. Limitations of the axis translation technique. *Proceedings of 4th International Conference on Expansive Soils*, Denver, CO, 1980; 117–135.
41. Houlsby GT. The work input to a granular material. *Géotechnique* 1979; **29**(3):354–358.
42. Houlsby GT. The work input to an unsaturated granular material. *Géotechnique* 1997; **47**(1):193–196.
43. Jommi C. Remarks on the constitutive modeling of unsaturated soils. In *Experimental Evidence and Theoretical Approaches in Unsaturated Soils*, Tarantino A, Mancuso C (eds). Balkema: Rotterdam, 2000; 139–153.
44. Vaunat J, Romero E, Jommi C. An elastoplastic hydro-mechanical model for unsaturated soils. In *Experimental Evidence and Theoretical Approaches in Unsaturated Soils*. Tarantino A, Mancuso C (eds). Balkema: Rotterdam, 2000; 121–135.
45. Hutter K, Jöhnk K. *Continuum Methods of Physical Modeling*. Springer: Berlin, 2004.
46. Trusdell C. *Rational Thermodynamics* (2nd enlarged edn). Springer: New York, 1984; 221.
47. Ziegler H. *An Introduction to Thermomechanics* (2nd revised edn). North-Holland Publishing Company: Amsterdam, 1983.
48. Ziegler H, Wehrli C. The derivation of constitutive relations from the free energy and the dissipation function. *Advances in Applied Mechanics* 1987; **25**:183–238.
49. Houlsby GT, Puzrin AM. A thermomechanical framework for constitutive models for rate-independent dissipative materials. *International Journal of Plasticity* 2000; **16**:1017–1047.
50. Hansen AC. Reexamining some basic definitions of modern mixture theory. *International Journal of Engineering Science* 1989; **27**(11):1531–1544.