AN ENTROPY MODEL FOR SHEAR DEFORMATION OF GRANULAR MATERIALS

Kan-ichi Kanatani
Department of Computer Science, Gunma University, Kiryu, Gunma, Japan

(revised manuscript received October 5,1979)

ABSTRACT

A statistical mechanical analogy for characterization of granular materials is discussed by using such notions as the state of the material, the density of states, entropy, canonical distribution and the partition function. The transition law of states during shear deformations of the material is microscopically investigated in the case of two-dimensional model granular materials. The assumption of entropy growth is shown to characterize the dilatancy of the material. A rough proof is given by assuming the measure preserving property of the transition and showing its ergodicity.

1. INTRODUCTION

There have been two different approaches to the modeling of granular materials: the macroscopic or continuum approach and the microscopic or particulate approach. The continuum approach is quite adaptable to experiments where macroscopic quantities such as the stress and the strain are the main interest. Drucker & Prager [1] proposed a plasticity theory of soils based on the so called associated flow rule, which has been often criticized for disagreement with experimental observations. Spencer [2] introduced the idea that deformation occurs by shear on certain critical planes and developed the so called double-slip theory. Goodman & Cowin [3] assumed that the stress depends on the gradient of the solid volume fraction of the material and developed a general continuum theory. Kanatani [4] combined the particulate approach with the continuum approach, analyzing interparticle friction and collisions in the flow of granular materials and taking statistical average to obtain constitutive equations of an equivalent continuum. He also showed that his theory is compatible with the plasticity theory of Drucker & Prager [1] if the associated flow rule is interpreted in a wide sense and that the new interpretation resolves all the inconsistencies of the Drucker-Prager theory [5].

In the particulate approach, one considers an assembly of particles idealized, say, as rigid spheres and attempts to deduce mechanical laws governing the assembly. This idea can be traced back to Reynolds [6]. Since mechanical properties of a particle assembly are generally very sensitive to the spatial configuration of the particles, one must adopt statistical methods to obtain more or less realistic models. Newland & Allely [7] considered particles on a potential slip plane and tried to relate the statistics of the interparticle contacts to macroscopic deformations. Rowe [8] studied mechanical properties of regularly packed rods and spheres and inferred mechanical characteristics of their random assemblies. Trying to give a theoretical basis to Rowe's inference, Horne [9] described macroscopic deformations by superposing statistics of two particles. However, many characteristics of granular materials including the so called dilatancy are related to many-particle interac-
tions, and hence it seems that they cannot be characterized by the two-particle statistics alone.

In this paper, we take the particulate approach to study the incipient deformations of granular materials, which are excluded by the continuum theory of Kanatani [4,5]. We extend Rowe's idea to random assemblies of particles and describe the assembly in terms of statistical mechanical notions such as the state of the material, the density of states, entropy, canonical distribution and the partition function. Then, we investigate the transition law of states during shear deformations in the case of two-dimensional model granular materials consisting of cylindrical rods of uniform size. We show that the assumption of entropy growth characterizes the dilatancy of the material. We give a rough proof, assuming the measure preserving property of the transition and showing its ergodicity.

2. MICROSTATES AND ENTROPY

Consider a particle assembly of bulk volume \( V \). We assume that the particles are rigid spheres of volume \( v_0 \). The number \( N \) of the particles in the sample is assumed to be large, and the particles are assumed to be packed randomly. Microscopically, however, the assembly is considered to consist of small cells in which particles are packed regularly in a certain arrangement. (See FIG.1.) Let \( E \) be the local void fraction in the cells, and let \( E_1, E_2, \ldots, E_n \) be the values \( E \) can assume in the sample. Let \( V_i \) be the total volume of those cells whose void fraction equals \( E_i \), and let \( N_i \) be the number of particles in these cells. Apparently, \( V_1 + V_2 + \ldots + V_n = V \) and \( N_1 + N_2 + \ldots + N_n = N \). Put \( p_i = V_i/V \). We say that the sample consists of microstates, or simply states, 1, 2, \ldots, \( n \) with respective probabilities \( p_1, p_2, \ldots, p_n \). Let \( \bar{E} \) be the average void fraction of the overall sample. By definition,

\[
\bar{E} = \frac{V - \sum_{i=1}^{n} p_i V_i}{V} = \sum_{i=1}^{n} \frac{p_i V_i}{V} = \sum_{i=1}^{n} E_i p_i .
\]

(1)

Hence, the average void fraction \( \bar{E} \) equals the expectation value of \( E \) with respect to the probabilities \( p_i \).

Consider the following problem: Prescribe the values of \( p \) such that the packing is completely random subject to the constraint that the average void fraction be \( \bar{E} \). Let us hypothetically divide the space occupied by the sample into \( M \) cells of equal volume and try to assign \( n \) states to these cells. Let \( N_i \) be the number of cells to which the \( i \)-th state is assigned. Then,

\[
\sum_{i=1}^{n} N_i = N ,
\]

\[
\bar{E} = \sum_{i=1}^{n} E_i p_i .
\]

(2)

The number of all the possible ways of such assignment is

\[
W = M! / N_1! N_2! \ldots N_n !.
\]

(3)

Take logarithm of this expression. Assuming that \( N \) and \( M \)'s are sufficiently large, we can apply the Stirling approximation formula to it to obtain

\[
\log W = - N \sum_{i=1}^{n} (N_i / M) \log (N_i / M) .
\]

(4)
If each assignment is equivalent a priori, the "most probable" way of assignment is the one that makes expression (4) maximum subject to constraints (2). Since \( p_i = N_i/M \), we can say that the most probable \( p_i \) are obtained by maximizing

\[
H_n = - \sum_{i=1}^{N} p_i \log p_i ,
\]

subject to the constraints

\[
\sum_{i=1}^{N} p_i = 1 , \quad \sum_{i=1}^{N} E_i p_i = \bar{E} .
\]

Since \( H_n \) is Shannon's entropy in information theory [10], we call it the entropy of the sample. The principle of entropy maximization for prescription of probabilities to a completely random system subject to constraints was first proposed by Jaynes [11,12,13], and it found various applications in characterization of granular materials, e.g. the coordination number (Brown [14]). Statistical concepts similar to the above have been also introduced to determine constitutive equations of sand (Mogami [15], Shimbo [16]).

3. DENSITY OF STATES AND CANONICAL DISTRIBUTION

Consider a continuous version of \( H_n \) entropy (5). As the sample size increases, the number of those values \( E \) can be assumed to increase, and in the limit we can define the probability density \( p(E) \) such that the probability of the states with void fraction in the interval \( [E, E+dE] \) is \( p(E)dE \). As Jaynes pointed out [12], the counterpart of (5) is not \(- \int p(E) \log p(E) dE \) because it is not invariant to the choice of parameters (e.g. void fraction, void ratio, solid volume fraction, solid volume ratio). We must define the density of states \( \Omega(E) \) such that the number of the states with void fraction in the interval \( [E, E+dE] \) divided by the total number of states in the sample is \( \Omega(E)dE \). Then, \( H_n \) has the following asymptotic form:

\[
H_n = - \int p(E) \log \frac{p(E)}{\Omega(E)} dE - \log n .
\]

The last term diverges as \( n \to \infty \) but is independent of \( p(E) \). Hence, we can define entropy as

\[
H = - \int p(E) \log \left( \frac{p(E)}{\Omega(E)} \right) dE .
\]

The integration is taken over the domain of \( E \) in which \( p(E) \) and \( \Omega(E) \) are defined. Apparently, this expression is invariant to the choice of parameters. This agrees with Kullback's information in statistical information theory [16] and is known to represent in a certain sense the "distance" between the two distributions \( p(E) \) and \( \Omega(E) \).

Now, we consider the following problem: Maximize \( H \) subject to the constraints

\[
\int p(E) dE = 1 , \quad \int E p(E) dE = \bar{E} .
\]

Introducing Lagrange multipliers, we finally obtain the following form of \( p(E) \), which we call the canonical distribution.

\[
p(E) = e^{\theta E} \Omega(E)/\Omega(\theta) , \quad Z(\theta) = \int e^{\theta E} \Omega(E)dE .
\]

We call \( \mathcal{Z}(\theta) \) the partition function. The value of parameter \( \theta \) is determined so that the average void fraction be \( \bar{E} \). Since

\[
(\frac{d}{d \theta}) \log \mathcal{Z}(\theta) = \mathcal{Z}'(\theta)/\mathcal{Z}(\theta) = \int E p(E)dE = \bar{E} ,
\]

we can determine \( \theta \) by solving

\[
\bar{E} = (\frac{d}{d \theta}) \log \mathcal{Z}(\theta) .
\]

Thus, \( \theta \) is a function of \( \bar{E} \) and we call it the conjugate average void fraction. We also see

\[
(\frac{d^2}{d \theta^2}) \log \mathcal{Z}(\theta) = \mathcal{Z}''(\theta)/\mathcal{Z}(\theta) - (\mathcal{Z}'(\theta)/\mathcal{Z}(\theta))^2 = \bar{E}' - \bar{E}^2 ,
\]

which is the variance \( \nu \) of the void fraction in the sample. Hence,
\[ V = (u^2/2u^2) \log Z(0) . \]  

After application of (12), the entropy for canonical distribution (10) becomes  
\[ H(E) = - \mu \overline{E} + \log Z(0) . \]  
This is the Legendre transform of \( \log Z(0) \) with respect to \( \overline{E} \), and hence by virtue of (12)  
\[ dH = - \theta d\overline{E} . \]  

All the above formulations are parallel to those of statistical mechanics, where \( \overline{E} \) is the total energy of the system in a thermal equilibrium. The temperature \( T \) of the system is related to \( \theta \) by \( \theta = 1/(kT)^{-1} \), where \( k \) is the Boltzmann constant. (See Jaynes [11] for details.) The method of Lagrange multipliers can only give the conditions that \( H \) becomes stationary, but it can be easily shown from the convexity of the logarithmic function that the canonical distribution (10) actually maximizes the entropy. Similarly, it can be shown that the right-hand side of (12) is a monotonic function of \( \theta \), and hence \( \theta(\overline{E}) \) is a single-valued function of \( \overline{E} \) (A.V. Khinchin [18]).

4. STATE TRANSITION OF TWO-DIMENSIONAL GRANULAR MATERIALS

Consider a two-dimensional model of granular materials consisting of cylindrical rods (FIG.2). Draw straight lines passing through the centers of the particles and the contact points, and define angles \( \alpha \) and \( \beta \) as shown in FIG.2. The domains of \( \alpha \) and \( \beta \) are  
\[ \pi/6 < \alpha < \pi/3 , \quad 0 < \beta < \pi/2 , \]  
respectively, where we have temporarily excluded the special case of the closest packing \((\alpha = \pi/6, \pi/4)\). It is later considered. We assume that all the states are those thus parameterized by \((\alpha, \beta)\) or the closest packing and ignore all the other possible sparse arrangements. This assumption is not so unreasonable if we consider materials in compression. Take the \( x \)-axis and the \( y \)-axis as the principal stress axes, and let \( \sigma_1 \) and \( \sigma_2 \) be the respective principal stresses (positive for compression). Decompose the contact force into normal and tangential components, and put them to be \( n_1, n_2 \), and \( \tau \) as indicated in FIG.2. The tangential force at the two pairs of contact points (see FIG.2) has the same magnitude in consequence of the balance of moments. Let \( \vec{f} = (f_x, f_y) \) be the force per unit length exerted through line \( AA' \) in FIG.2. By definition of the principal stress axes,  
\[ f_x = \sigma_1 \sin(\alpha + \beta) , \quad f_y = \sigma_2 \cos(\alpha + \beta) . \]  
Since this force is balanced by the contact forces \( n_2 \) and \( \tau \), we get  
\[ n_2 = 2\alpha(\sigma_1 \sin(\alpha + \beta) \sin(\alpha + \beta) - \sigma_2 \cos(\alpha + \beta) \cos(\alpha + \beta)) , \]  
\[ \tau = 2\alpha(\sigma_1 \sin(\alpha + \beta) \cos(\alpha + \beta) + \sigma_2 \cos(\alpha + \beta) \sin(\alpha + \beta)) , \]  
where \( \alpha \) is the radius of the particles. Similarly, considering the force acting through line \( BB' \) in FIG.2, we get
\[ n_2 = 2a(\sigma_1 \sin(\alpha - \beta) \cos(\alpha + \beta) + \sigma_2 \cos(\alpha - \beta) \sin(\alpha + \beta)) \].

If we put
\[ p = (\sigma_1 + \sigma_2)/2 \], \[ q = (\sigma_1 - \sigma_2)/2 \],
which represent compressive and shearing forces respectively, we can rewrite (19), (20) and (21) as
\[ \tau = 2a(-pcos2\alpha + qcos2\beta) \],
\[ n_1 = 2a(psin2\alpha - qsin2\beta) \],
\[ n_2 = 2a(psin2\alpha + qsin2\beta) \].

Now, we consider the slip of the particles. We say that the slip is (+1), if it occurs at the contact points with normal force \( n_1 \) and if \( \tau \) is positive. If \( \tau \) is negative, the slip is said to be (-1). The other possible slips (+2) and (-2) are similarly defined. Let \( \phi \) be the friction angle of the particles. The Coulomb condition that these slips occur are as follows:

\[
\begin{align*}
(+1): & \quad \tau > n_1 \tan \phi \quad \tau > q\cos(2\beta - \phi) > p\cos(2\alpha - \phi), \\
(-1): & \quad -\tau > n_1 \tan \phi \quad -\tau > q\cos(2\beta + \phi) < p\cos(2\alpha + \phi), \\
(+2): & \quad \tau > n_2 \tan \phi \quad \tau > q\cos(2\beta + \phi) > p\cos(2\alpha - \phi), \\
(-2): & \quad -\tau > n_2 \tan \phi \quad -\tau > q\cos(2\beta - \phi) < p\cos(2\alpha + \phi).
\end{align*}
\]

Consider such loadings that the sample is initially in uniform compression \((q = 0)\) and gradually \(q/p\) increases. The possible slip is the one for which the Coulomb condition is first satisfied. If \(\phi < \pi/4\) or \(q/p\) is small and \(\alpha\) is nearly \(\pi/6\) or \(\pi/3\) \((\pi/6 < \alpha < \pi/4 - \phi/2, \pi/4 + \phi/2 < \alpha < \pi/3)\), then such slips occur that \(\alpha\) becomes \(\pi/6\) or \(\pi/3\) (the closest packing). Otherwise, the possible slip is shown to be (+1) if \(\alpha > \beta\) and (-1) if \(\alpha < \beta\).

Finally, consider the case of the closest packing. As is shown in FIG.3, the state specification by \((\alpha, \beta)\) is degenerate, and the same state can be specified differently. Let \(\sim\) designate the relation that both expressions specify the same state. We can easily see that
\[
\begin{align*}
(\pi/6, \beta) & \sim (\pi/6, \beta \pm \pi/3), \\
(\pi/3, \beta) & \sim (\pi/3, \beta \pm \pi/3), \\
(\pi/6, \beta) & \sim (\pi/3, \beta \pm \pi/6).
\end{align*}
\]

We can summarize the law of state transition by FIG.4. A state \((\alpha, \beta)\) is represented as a point inside rectangle ABCD or on its circumference, where AD and BC, DG and HC, GH and AE, EF and HC, and FR and AE are identified respectively. The slip is (+1) in the left half AEHD and is (-1) in the right half EBCN. In both cases, it is clear from the definition of \(\alpha\) and \(\beta\) in FIG.2 that \(\alpha\) and \(\beta\) change their values by the same absolute amount. Hence, if the initial state corresponds to point \(P_0\) in FIG.4, it moves up along a straight line with tangent \(45^\circ\) and reaches point \(P_1\), which is identified as point \(P_2\). It then moves down to point \(P_3\), which is again identified as point \(P_2\), and repeats this cycle. The short arrows in FIG.4 indicate the possible incipient transitions. We can see that the states inside triangles ADG and FBC soon disappear because there are no trajectories of states entering these regions. After all, all the states fall into periodic cycles. If we cut out AEHD and EBCN,
and put $AE$ and $GH$ together and $EF$ and $HC$ together, we obtain two cylinders on which all the trajectories of states are closed orbits and are parallel to each other.

5. CANONICAL DISTRIBUTION AND ENTROPY OF TWO-DIMENSIONAL GRANULAR MATERIALS

We now consider the probability distribution of states. As is seen from FIG.2, the void fraction of state $(a, \theta)$ is

$$E(a) = 1 - \frac{a}{\pi/2} \sin 2a$$

which does not depend on $\theta$. Hence, we have only to consider distributions for $a$. In other words, since we are concerned with the average void fraction $\overline{E}$ as the macroscopic parameter and it is invariant to rigid rotations of the sample, we can assume that $\theta$ is distributed uniformly over all directions. First, we must determine the density of states $\Omega(E)$, which is fundamental to our formulation. However, we may instead determine $\Omega(a)$ because our formulation is invariant to the choice of parameters. Here, we assume that $\Omega(a)$ is uniform:

$$\Omega(a) = \frac{6}{\pi}$$
$$\frac{\pi/6}{\pi} < a < \frac{\pi}{3}$$

This means that we are assuming all the states with arbitrary $a$ in the interval $(\pi/6, \pi/3)$ are mutually equivalent a priori for constituting a completely random sample. In terms of $E$, we have

$$\Omega(E) = \frac{6}{(1 - E)^{15/2} - 32E + 16 - \pi^2}$$

The minimum value $E_0$ of $E$ corresponds to $a = \pi/6, \pi/3$ and the maximum value $E_1$ corresponds to $a = \pi/4$. They have the following values.

$$E_0 = 1 - \sqrt{3} \approx 0.0931\ldots$$
$$E_1 = 1 - \pi/4 = 0.2146\ldots$$

The canonical distribution is given by

$$p(a) = \frac{6}{\pi} e^{6E(a)/2Z(\theta)}$$
$$Z(\theta) = \frac{(6/\pi)^{\pi/2}}{\int_0^{\pi/3} e^{6E(a)/2} da}$$

See FIG.5. The corresponding distribution $p(E)$ for $E$ is plotted in FIG.6. The conjugate average void fraction $\theta$ is determined by solving (12) and has the values plotted in FIG.7. The entropy $H(E)$ is plotted in FIG.8. The 'most' random sample with maximum $H$ is the one with $E = E_0$, where

$$E_0 = 0.176040769\ldots$$

This value corresponds to $\theta = 0$ and in this case the canonical distribution $p(a)$ becomes uniform, i.e., $p(a) = \Omega(a)$. 
6. ENTROPY GROWTH, DILATANCY AND ERGODICITY

The internal configuration of particles will change when the material undergoes shear deformations, and its exact description is almost impossible. Here, we assume the growth of entropy in analogy with Boltzmann's H theorem in the kinetic theory of gas. A rough proof is given later in the case of two-dimensional granular materials. We assume that the distribution of states is always canonical, i.e., the material preserves its randomness during deformations. Let ε be a parameter measuring the degree of shearing. Then, dH/dε is determined only by the present canonical distribution and hence is regarded as a function of θ. If we put
\[ \frac{dH}{dc} = F(\theta) \quad (32) \]

we must demand that

\[ F(\theta) > 0 \quad (33) \]

with equality for \( \theta = 0 \), in which case \( H \) attains its maximum. By (16) we have

\[ \frac{dE}{dc} = -\frac{F(n(E))}{n(E)} \quad (34) \]

The right-hand side is positive for \( \theta < 0 \) and negative for \( \theta > 0 \). In other words, the average void fraction \( E \) increases if \( E < E_c \) and decreases if \( E_c < E \). This phenomenon corresponds to the so-called dilatancy of granular materials. Solutions of (34) are given in the form

\[ c(E) = -\int_{E_t}^{E} \frac{\theta(\xi)n(\xi)}{P(\theta(\xi))} d\xi \quad (35) \]

where \( E_t \) is the initial average void fraction. An example of the solutions for \( F(\theta) = K\theta^2 \) (\( K = \text{const.} \)) is shown in FIG.9. Note that we have neglected the possible incipient transitions. If they are also taken into consideration, the average void fraction should first decrease near \( c = 0 \) as is usually observed. As the deformation proceeds, the average void fraction approaches a constant value \( E_c \). This stage is regarded as the flow regime Kanatani [4,5] studied.

Now, we give a theoretical basis to the above description in the two-dimensional case, assuming the measure preserving property of the transition and showing its ergodicity. As was shown in Sect.5, the trajectories of states are closed orbits on two cylinders. We assume that the transition velocity is constant for each orbit and is continuous over neighboring orbits. Since all the orbits are parallel to each other, if we neglect the 'noise' caused by the interactions through the boundaries of state cells in the sample, the transition on the state space, i.e., the two cylinders, is measure preserving. In other words, if the points in a region \( D \) on the cylinders move after some time and form a new region \( D' \), the measure (i.e., the area) of \( D' \) is equal to that of \( D \). The entropy growth assumption is nothing but to say that the distribution on the state space approaches uniform distribution as the deformation proceeds. We now prove the ergodicity of the transition, i.e., we show that any initial distribution \( \mu(a) \) converges to uniform distribution over the state space in the weak sense. Take a new inclined axis for \( a \) such that each orbit is represented by \( \theta = \)}
const. Extend the interval $[\pi/6, \pi/3]$ of $\alpha$ to $(-\infty, \infty)$ and periodically extend the initial distribution $p(\alpha)$. Since $p(\alpha)$ has period $\pi/6$, it can be expressed by the Fourier series

$$p(\alpha) = p_0 + \sum_k p_k e^{-i2\pi k \alpha}.$$  

The distribution at time $t$ is given by

$$p(\alpha, \theta, t) = p(\alpha - \nu(\theta) t),$$

where $\nu(\theta)$ is the velocity of state transition along orbit $\theta$. By symmetry, we consider the domain $0 < \theta < \pi/6$ only. Let $f(\theta)$ be an arbitrary continuous function over that domain.

Then, we can see that

$$\int_0^{\pi/6} f(\theta) p(\alpha, \theta, t) \, d\theta = \int_0^{\pi/6} p_0 f(\theta) \, d\theta + \sum_k p_k \int_0^{\pi/6} e^{-i2\pi k \theta} f(\theta) \, d\theta$$

$$= \int_0^{\pi/6} p_0 f(\theta) \, d\theta + \sum_k p_k \int_0^{\pi/6} e^{-i2\pi k \theta} f(\theta) \, d\theta$$

where the integrations are in the sense of Lebesgue. If $d\theta/\nu$ is continuous, the last term converges to $0$ as $t \to \infty$ by the Riemann-Lebesgue theorem. Thus, $p(\alpha, \theta, t)$ converges to $p_0$ as $t \to \infty$ in the weak sense, or in other words as a distribution of Schwartz. Even if the 'noise' is present, it seems probable that the convergence is not prevented and may be even accelerated.

7. CONCLUDING REMARKS

We have discussed a statistical theory of granular materials, using statistical mechanical notions. Although our analysis is restricted to the two-dimensional model granular materials, the formulation gives a conceptual basis to construct constitutive theories of granular materials in various applications.

REFERENCES


6. REYNOLDS, O., On the dilatancy of media composed of rigid particles in contact, Phil. Mag.: Ser. 5, 20, 469, 1885.


