Powder Technology, 30 (1981) 217 - 223

The Use of Entropy in the Description of Granular Materials

KEN-ICHI KANATANI

Department of Computer Science, Gunma University, Kiryu, Gunma 376 (Japan)

(Received January 6, 1981; in revised form April 27, 1981)

SUMMARY

This is a critical review of the existing theories on characterizing the random packing of granular materials by means of entropy maximization. It is pointed out that the state density function, which specifies the limit-taking process and determines the asymptotic form of entropy, must be introduced in order to assure the invariance to parameter transformations. The entropy for the particle size distribution is first formulated, and an application to the void fraction distribution is illustrated. These examples show the relationship between the present entropy and those in information theory, statistical mechanics and thermodynamics. Then, a new viewpoint is presented; entropy is regarded as a measure of the 'distance' between statistical distributions. It is concluded that in this new sense the entropy provides a useful tool for the description of random quantities involved in granular materials.

1. INTRODUCTION

There have been several attempts to characterize the random packing of granular materials as a state of maximum entropy analogous to statistical mechanics (*e.g.* [1 - 5]). This approach has given the misleading impression that physically reasonable statistical distributions are derived by maximizing the entropy *irrespective of the physical laws governing the microscopic interactions among the particles*. In this paper we critically review the existing theories of entropy maximization and point out that the entropy alone does not suffice to determine the actual distributions of random quantities involved in granular materials. A close examination reveals that a crucial assumption on the randomness is tacitly posed in all the existing theories of this type and that there exists no logical justification for it. As a matter of fact, this difficulty is traced back to the mathematical foundation of statistical mechanics (*e.g.*, see [6]). The crucial point is the limit-taking process involved in the formulation. Usually, entropy of this type is defined with regard to discrete random variables. Then, an asymptotic form is obtained in the limit of continuous variables. The subtle point is the fact that the resulting asymptotic form depends largely on the way of limit taking. There are a number of possible limit-taking processes, and the entropy takes different forms accordingly. Hence, the form of entropy thus obtained is not invariant to parameter transformations. In order to assure the invariance, we must introduce a state density function which specifies the limit-taking process. The form of the state density function is not determined by the statistical consideration alone. It must be determined by an *a posteriori* criterion. This fact is illustrated in detail in this paper.

On the other hand, a new viewpoint is proposed for the use of entropy in describing the packing structure of granular materials; we regard entropy as a quantity measuring a kind of *distance* between two statistical distributions. This viewpoint is extensively discussed in Kullback [7] in relation to the mathematical theory of statistics, and it is in this new sense that the entropy is most useful in the description of random quantities.

Quantities involved in the random packing of granular materials are various types of random variables, and they are characterized by their distribution functions, *e.g.* the distributions of the particle size, the void fraction, the number of contact points, the contact

forces, etc. Entropy is defined for each of these quantities. In this paper, the particle size distribution is considered first. Then, we apply the formulation to the void fraction distribution, which most of the theories of this type have dealt with. Other quantities can be treated in a similar manner. (For theoretical treatment of the contact force distribution, see Kanatani [8]. See also Kanatani [9] for the mechanics of model granular materials.) We follow Jaynes [10 - 12] for the formulation and show the relationship between the present entropy and those appearing in information theory, statistical mechanics and thermodynamics. It then becomes clear that the entropy actually measures the distance between distributions, as stated above.

2. ENTROPY AND CANONICAL DISTRIBUTIONS

Here, we consider the size distribution of particles. Suppose that the particles are all spheres and that the radius a ranges in the interval $a_{\min} \leqslant a \leqslant a_{\max}$. The number N of the particles is assumed to be sufficiently large. Now, divide the interval into n consecutive sub-intervals $[a_{\min}, a']$, $[a', a'']$, ..., $[a''', a_{\max}]$. Let $a_1, a_2, \ldots, a_n$ be the midpoints of these sub-intervals in ascending order, and let the actual values of the radius be quantized to discrete values by identifying all the values in one sub-interval with its midpoint value. Let N_i be the number of the particles whose radius is quantized to a_i, and put

$$p_i = N_i/N \tag{1}$$

which represents the fraction of the particles classified into the ith subinterval.

Now, let us consider the *most probable* distribution of p_i. Consider an arbitrary assignment of N particles to the n sub-intervals. If N_i is the number of particles assigned to the ith sub-interval, then the number of possible assignments which lead to this configuration is

$$W = N!/N_1!N_2!\ldots N_n! \tag{2}$$

Taking the logarithm of this expression, and assuming that N and each N_i are sufficiently large, we apply the Stirling approximation formula to obtain

$$\log W = -N \sum_{i=1}^{n} \left(\frac{N_i}{N}\right) \log\left(\frac{N_i}{N}\right). \tag{3}$$

If each assignment is equivalent *a priori*, the most probable assignment is the one that maximizes this expression. Hence, it is concluded that the most probable p_i is obtained by maximizing the entropy:

$$H_n = -\sum_{i=1}^{n} p_i \log p_i \tag{4}$$

subject to

$$\sum_{i=1}^{n} p_i = 1 \tag{5}$$

and other necessary constraints. For example, if the average radius is specified to be $\bar{a}$, then we must add the constraint

$$\sum_{i=1}^{n} a_i p_i = \bar{a} \tag{6}$$

Now, let us consider an asymptotic form of the entropy by taking the limit in such a way that $N \to \infty$, $n \to \infty$ and the length of each subinterval goes to zero with the boundary values $a_{\min}$ and $a_{\max}$ fixed. The crucial point is that we *must not* replace expression (4) by

$$-\int p(a) \log p(a)\, da \tag{7}$$

because this form is *not invariant to parameter transformations.* For example, consider a monotone function $b(a)$ of a. The probability density $p(b)$ of b is defined by $p(a)\, da = p(b)\, db$. Hence, $p(a) = p(b)\, db/da$. Substitution of this in (7) yields

$$-\int p(b) \log p(b)\, db - \int p(b) \log(db/da)\, db \tag{8}$$

which does not coincide with $\int p/b) \log p(b)\, db$ because of the second term. There are in general a number of choices of parameters for a single object. For example, the parameters equivalent to the radius of a sphere are the diameter $2a$, the projected area πa^2, the volume $4\pi a^3/3$ and so on. We must have the same form of entropy for each of them, and one form must be derived from another by the parameter transformation alone.

Now, we carefully derive a correct asymptotic form of (4). The continuous probability density $p(a)$ is defined in such a way that the fraction of the number of particles in the differential interval $[a, a + da]$ is equal to $p(a)\, da$

Next, the state density function $\Omega(a)$ is defined in such a way that the fraction of the number of midpoints of the sub-intervals in the differential interval $[a, a + da]$ is equal to $\Omega(a)\,da$ in the limit. Then, we obtain

$$-\int p(a) \log(p(a)/\Omega(a))\, da - \log n \qquad (9)$$

The last term diverges as $n \to \infty$, but it is independent of $p(a)$ and hence it can be dropped for maximization. Thus, we arrive at a continuous version of entropy of the form

$$H = -\int p(a) \log(p(a)/\Omega(a))\, da \qquad (10)$$

The integration is carried out over $a_{\min} \leqslant a \leqslant a_{\max}$. Henceforth, it is understood that the integration is carried out over the range where the state density function takes on positive values. It is easily checked that this form is in fact invariant to parameter transformations, since the state density function $\Omega(b)$ for $b = b(a)$ of the previous example is given by $\Omega(a)\,da = \Omega(b)\,db$ and hence $\Omega(b) = \Omega(a)\,da/db$.

It should be noted that we do not have to consider the state density function if we assume that the midpoints $a_1, a_2, \ldots, a_n$ are always equidistant in the process of the limit $n \to \infty$. The tricky point is that the asymptotic form becomes different if they are not equidistant. However, there is no reason to assume that they are equidistant. In fact, we may as well take the limit such that $(a_1)^2$, $(a_2)^2, \ldots, (a_n)^2$ are kept equidistant, or such that $1/a_n, 1/a_{n-1}, \ldots, 1/a_1$ are equidistant. Therefore, we must necessarily introduce the state density function as defined above. The derivative db/da in (8) reflects the fact that if the values of one parameter are equidistant, *the transformed values in general are not.*

The constraints (5) and (6) become

$$\int p(a)\, da = 1 \qquad (11)$$

$$\int a p(a)\, da = \bar{a} \qquad (12)$$

It can be shown by the method of Lagrange multipliers that maximization of the entropy (10) subject to the constraints (11) and (12) yields

$$p(a) = e^{\theta a}\Omega(a)/Z(a) \qquad (13)$$

where θ is the Lagrange multiplier associated with the constraint (12) and $Z(\theta)$ is the normalization factor to satisfy the constraint (11), *i.e.*,

$$Z(\theta) = \int e^{\theta a}\Omega(a)\, da \qquad (14)$$

We call the distribution (13) the canonical distribution and $Z(\theta)$ the partition function after statistical mechanics. Since

$$\frac{d}{d\theta} \log Z(\theta) = \frac{Z'(\theta)}{Z(\theta)} = \int a p(a)\, da \qquad (15)$$

the value of θ is determined from $\bar{a}$ by solving

$$\bar{a} = (d/d\theta) \log Z(\theta) \qquad (16)$$

We call the parameter $\theta(\bar{a})$ so determined from $\bar{a}$ the conjugate average radius. We can also see from

$$\frac{d^2}{d\theta^2} \log Z(\theta) = \frac{Z'(\theta)}{Z(\theta)} - \left(\frac{Z'(\theta)}{Z(\theta)}\right)^2 \qquad (17)$$

that the variance of the radius is given by

$$\overline{(a - \bar{a})^2} = (d^2/d\theta^2) \log Z(\theta) \qquad (18)$$

The value of the entropy (10) for the canonical distribution (13) is

$$H(\bar{a}) = -\theta \bar{a} + \log Z(\theta) \qquad (19)$$

which is, by virtue of eqn. (16), the Legendre transform of $\log Z(\theta)$ with respect to $\bar{a}$. Hence, we obtain the differential equality

$$dH = -\theta d\bar{a}, \quad i.e. \quad \theta = -dH/d\bar{a} \qquad (20)$$

The method of Lagrange multipliers only gives the condition that the entropy becomes stationary, but it can be shown from the convexity of the logarithmic function that the canonical distribution (13) indeed maximizes the entropy. Similarly, it can be shown that the right-hand side of eqn. (16) is a monotone function of θ and hence $\theta(\bar{a})$ is defined as a single-valued function of $\bar{a}$ (*e.g.*, see [6]).

All the above formulations parallel those of statistical mechanics. To see it, we only have to replace a by E, the energy of a gas in thermal equilibrium. The temperature T is related to θ by $\theta = -(kT)^{-1}$, where k is the Boltzmann constant. The thermodynamic entropy S is related to H by $S = kH$. Equation (13) is the Gibbs canonical distribution, and eqns. (16) and (18) are well-known identities of the partition function. Equation (20) in this case is rewritten as

$$d\bar{E} = -T\, dS, \quad i.e. \quad T^{-1} = dS/d\bar{E} \qquad (21)$$

If the increment $d\bar{E}$ of the total energy consists of the increment dV of the internal energy and the increment pdU of the work done by the external pressure, eqn. (21) is expressed as

$$dU = T\,dS - p\,dV \tag{22}$$

which is the well-known 'law' of thermodynamics.

Thus, we have completely solved the problem of finding the most probable distribution of particle size in a granular assembly, but are we to be satisfied? Of course not, for the state density function is yet to be given. The problem is that there exists no theoretical method to find it. As long as it is indeterminate, the whole theory remains formal and in a sense trivial. In statistical mechanics, the state density function $\Omega(E)$ is defined so that the number of eigenstates of the Hamiltonian with energy levels in the interval $[E, E + dE]$ is equal to $\Omega(E)\,dE$. This is equivalent to saying that *all the configurations of the eigenstates that realize the total energy E are mutually equivalent a priori.* This is one of the most crucial assumptions in statistical mechanics (*e.g.*, see [6]).

Most of the existing theories of entropy maximization applied to granular materials failed to introduce the state density function. Hence, they have given the misleading impression that physically reasonable distributions are derived by entropy maximization without any regard to the physical laws governing the microscopic interactions. A close examination reveals that the state density function is tacitly assumed to be the uniform distribution with respect to some particular form of the parameter in all the theories. For example, Mogami [1], Brown [3] and Shahinpoor [4] implicitly assumed that the state density for the void ratio is uniform. This means that the possible values of the void ratio in a sample are kept equidistant in the process of the limit taking. Jowitt and Munro [3] also assumed tacitly that particular 28 distinct packings are equivalent, *i.e.*, the state weighting factors (instead of the state density function in this discrete case) are all equal. It has now become apparent that the assumption for the state density should be elucidated at first. The plausibility of the assumption is checked only by the predictions the theory makes.

Meanwhile, there is a way to circumvent such clumsy hypotheses. Note that eqn. (4) coincides with Shannon's entropy [13] in information theory, and it measures ambiguity or complexity of the distribution *on the assumption that all the sub-intervals are equivalent a priori.* In the limit of $n \to \infty$, we have obtained eqn. (10), which coincides with Kullback's information [7] in statistical information theory. It can be shown that H is always nonpositive and that if no additional constraints exist, H takes on its maximum value 0 only when $p(a) = \Omega(a)$ (see [7]). This implies that $-H$ is a quantity that measures a kind of distance between the two distributions $p(a)$ and $\Omega(a)$. This distance is directed, *i.e.* nonsymmetric, and its properties are fully discussed by Kullback [7]. From this viewpoint the canonical distribution (13) is interpreted as the nearest distribution to $\Omega(a)$ subject to the constraint (12). In view of this it is now apparent that the maximization of entropy is utilized only when a certain standard distribution $\Omega(a)$ is specified. This distribution is regarded as the distribution of *a priori* equivalent states, *i.e.*, the distribution *most probably expected when no additional constraints (except the normalization) exist.* Now, what we can do is first to define the most random or the most probable distribution $\Omega(a)$ and then to measure the amount of difference between the actual distribution and the reference distribution by the use of the entropy. Indeed, we cannot define the amount of randomness unless we know the reference distribution, since even a uniform distribution of some parameter is changed to a non-uniform distribution by a parameter transformation.

3. ENTROPY OF TWO-DIMENSIONAL MODEL GRANULAR MATERIALS

In order to illustrate the viewpoint in the previous section, let us consider the void distribution of granular materials, which most of the existing theories of entropy maximization have dealt with. Here, we follow the model of Kanatani [5]. Consider a two-dimensional model granular material consisting of cylindrical rods of equal size. Let V be the bulk volume (per unit length of the cylinders) and N be the number of the

particles (*i.e.* the cylinders). The packing is assumed to be macroscopically random. Microscopically, however, the assembly is considered to consist of small cells in which the particles are packed regularly (Fig. 1). (If the packing is so random that each cell contains only one particle, the cells are called the 'Voronoi cells' [14].) Consider n sub-intervals of the void fraction as in Section 2. Let $E_1, E_2, \ldots, E_n$ be the values of the mid-points of the sub-intervals, and identify all the values in the sub-intervals with their midpoint values as before. Let V_i be the total volume of those cells whose void fraction is identified with E_i, and let N_i be the number of the particles in these cells. Then

$$\sum_{i=1}^{n} V_i = V, \qquad \sum_{i=1}^{n} N_i = N \tag{23}$$

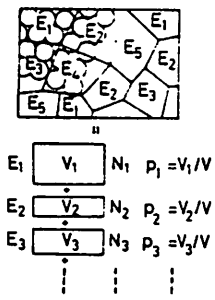

Fig. 1. Microstates and their probabilities.

Put

$$p_i = V_i/V \tag{24}$$

Then, the sample is regarded as a composite consisting of microstates 1, 2, ..., n with respective probabilities $p_1, p_2, \ldots, p_n$. (Or, equivalently, they can be regarded as the probabilities of the Voronoi cells with respective void fractions.) Let $\bar{E}$ be the average void fraction of the overall sample. By definition, $\bar{E} = (V - vN)/V$, where v is the volume of one particle. From eqns. (23) and (24), we obtain

$$\bar{E} = \sum_{i=1}^{n} \frac{V_i - vN_i}{V_i} \frac{V_i}{V} = \sum_{i=1}^{n} E_i p_i \tag{25}$$

i.e., the average void fraction $\bar{E}$ equals the expectation value of the local void fraction with respect to the probabilities (24). (Note that if the void ratio $e = E/(1-E)$ is used instead of the void fraction E, the corresponding probabilities are given by N_i/N instead of (24) as noted by Shahinpoor [4].)

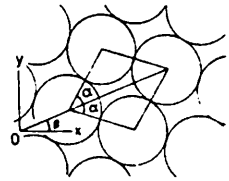

Fig. 2. A typical microstate of the two-dimensional model granular material.

Assuming that the packing is sufficiently dense, we consider the microstate to be the regular array of particles as is shown in Fig. 2. Consider a unit parallelogram made by the four lines passing the centres of the particles in contact. The microstate is uniquely specified by the two parameters α and β as indicated in Fig. 2. In the limit of $n \to \infty$, they are considered to be continuous variables such that

$$\pi/6 < \alpha < \pi/3, \qquad 0 \leqslant \beta < \pi/2 \tag{26}$$

If $\alpha = \pi/6$ or $\pi/3$, the microstate is the densest packing. The associated void fraction is given by

$$E(\alpha, \beta) = 1 - \frac{\pi}{4 \sin 2\alpha} \tag{27}$$

Now, we must define the state density function $\Omega(\alpha, \beta)$, which specifies the limit-taking process and describes the most random distribution. First, we assume that Ω does not depend on β, *i.e.*, it is uniform with respect to β. Since β specifies the angle of rigid rotation, we can safely assume that the sample contains microstates of all directions uniformly. Next, we assume that the angle α distributes uniformly in the interval $[\pi/6, \pi/3]$ in the most random sample. Thus, we put

$$\Omega(\alpha, \beta) = 6/\pi \tag{28}$$

We again emphasize the fact that we have now defined a sample that has the most random packing.

Changing the parameters from (α, β) to E, we obtain from eqn. (27) the state density for the void fraction E as follows:

$$\Omega(E) = \frac{6}{(1-E)\sqrt{16E^2 - 32E + 16 - \pi^2}} \tag{29}$$

(Note that $\Omega(E)$ is given by $\Omega(E)\, dE = \Omega(\alpha, \beta)\, d\alpha$ and hence $\Omega(E) = \Omega(\alpha, \beta)/(dE/d\alpha)$.) This distribution is plotted

in Fig. 3, where E_0 is the minimum value of E (attained when $\alpha = \pi/6$ or $\pi/3$) and E_1 the maximum value (attained when $\alpha = \pi/4$). They are respectively

$$E_0 = 1 - \sqrt{3}\pi 6 = 0.0931\ldots$$
$$E_1 = 1 - \pi/4 = 0.2146\ldots \quad (30)$$

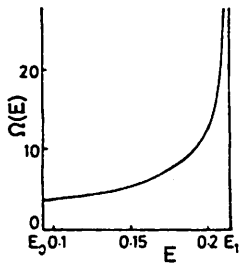

Fig. 3. The state density function, which defines the 'most random' distribution.

Let us consider the entropy maximization subject to the constraint that the average void fraction is $\bar{E}$. This is achieved by working on parameter α instead of E, because our formalism is *invariant to parameter transformations*. The canonical distribution for α is then given by

$$p(\alpha) = (6/\pi)e^{\theta E(\alpha)}/Z(\theta) \quad (31)$$

$$Z(\theta) = (6/\pi)\int_{\pi/6}^{\pi/3} e^{\theta E(\alpha)}\,d\alpha \quad (32)$$

where the conjugate average void fraction θ is determined from $\bar{E}$ by solving

$$\bar{E} = (d/d\theta)\log Z(\theta) \quad (33)$$

Equation (31) is plotted in Fig. 4, and the parameter transformation yields the canonical distribution for E, which is shown in Fig. 5. The conjugate average void fraction θ is a function of $\bar{E}$ and is plotted in Fig. 6. Figures

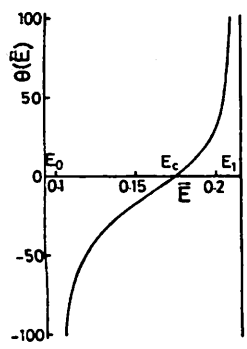

Fig. 4. Conjugate average void fraction $\theta(\bar{E})$.

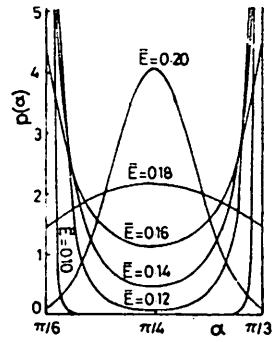

Fig. 5. Canonical distribution for the contact angle α.

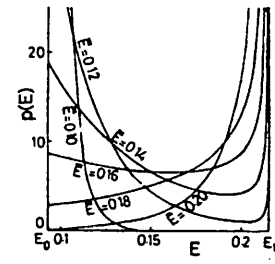

Fig. 6. Canonical distribution for the void fraction E.

4 and 5 show that as the average void fraction $\bar{E}$ decreases, the sample contains the larger proportion of dense microstates. The entropy for the canonical distribution is shown in Fig. 7. It is nonpositive, and it attains its maximum at

$$\bar{E} = E_c\ (= 0.1760\ldots) \quad (34)$$

at which $p(E)$ coincides with $\Omega(E)$ of eqn. (25). When $\bar{E} = E_0$, the sample consists entirely of the densest packing, and when $\bar{E} = E_1$, it consists entirely of the microstates of minimum density ($\alpha = \pi/4$). Thus, both of these extremes are *far from* random, and the entropy approaches $-\infty$. In other words, $-H$ measures the distance between $p(E)$ and $\Omega(E)$.

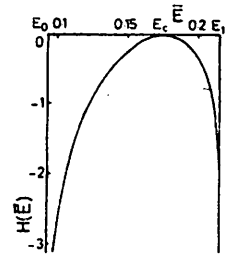

Fig. 7. Entropy of canonical distributions.

We have considered only the entropy of the canonical distribution as an example. However, the entropy is defined for any distribution. If the actual distribution $p(E)$ is given, we can measure the amount of difference between that distribution and $\Omega(E)$. Or we can take the canonical distribution (31) as the reference distribution and measure the difference from it. In this sense, the entropy plays an important role in measuring randomness of the sample.

4. CONCLUSION

We have pointed out that in order to define the entropy for continuous quantity we must introduce the state density function, which specifies the limit-taking process and determines the asymptotic form of the entropy, to assure the invariance to parameter transformations. Then we have reached a conclusion that the entropy maximization has its significance only when a certain standard distribution, which defines the most random distribution, is given as a reference. Thus the applicability of the entropy maximization is limited in this sense.

On the other hand, it has been shown that the entropy measures the distance between the distribution under consideration and the reference distribution. This observation widens the applicability of entropy, and it is in this sense that entropy plays an important role in the description of granular materials.

ACKNOWLEDGEMENT

The author wishes to thank Professor M. Shahinpoor of Clarkson College and the reviewers of the manuscript for helpful discussions.

REFERENCES

1 T. Mogami, *Soils Found.*, *5* (1965) 26.
2 P. W. Jowitt and J. Munro, *Proc. Application of Statistics and Probability in Soil and Structural Engineering, Aachen, 1975*, p. 365.
3 C. B. Brown, in S. C. Cowin and M. Satake (eds.), *Proc. U.S.–Japan Seminar, Sendai, 1978*, Gakujutsu Bunken Fukyu-kai, Tokyo, 1978, p. 98.
4 M. Shahinpoor, *Powder Technol.*, *25* (1980) 163.
5 K. Kanatani, *Int. J. Eng. Sci. (Lett. Appl. Eng. Sci.)*, *18* (1980) 989.
6 A. I. Khinchin, *Mathematical Foundations of Statistical Mechanics*, Dover, New York, 1949.
7 S. Kullback, *Information Theory and Statistics*, Wiley, New York, 1959.
8 K. Kanatani, *Powder Technol.*, *28* (1981) 167.
9 K. Kanatani, *Int. J. Eng. Sci.*, *17* (1979) 419.
10 E. T. Jaynes, *Phys. Rev.*, *106* (1957) 620; *108* (1957) 171.
11 E. T. Jaynes, in J. L. Bogdanoff and F. Kozin (eds.), *Proc. 1st Symp. on Engineering Applications of Random Function Theory and Probability*, Wiley, New York, 1963, p. 163.
12 E. T. Jaynes, in K. W. Ford (ed.), *Statistical Physics III*, Benjamin, New York, 1963, p. 182.
13 C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, Univ. of Illinois Press, Urbana, IL, 1964.
14 J. L. Finney, *Proc. R. Soc. London, A319* (1970) 479, 495.