Maximum Entropy Principle and Texture Formation

M. Arminjon & D. Imbault

Laboratoire "Sols, Solides, Structures",
Institut de Mécanique de Grenoble, Grenoble, France

1. The principle of maximum statistical entropy (MAXENT)

2. Implementation of MAXENT in physics of heterogeneous media

3. MAXENT with obvious constraints = Voigt-Taylor model

4. A less trivial MAXENT model (one more constraint)

1. The principle of maximum statistical entropy (MAXENT)

MAXENT (Jaynes 1957) = link between information theory and statistical mechanics.

Information theory: the "amount of uncertainty" represented by a probability distribution $(p_i)_{i = 1, ..., M}$ on a finite set $E = \{x_1, ..., x_M\}$ is the *statistical entropy* given by

$$S = - \sum_{i=1}^{M} p_i \, \mathrm{Log}\, p_i .$$

Case where only some expectation values

$$\langle \phi_q \rangle \equiv \sum_{i=1}^{M} p_i \, \phi_q(x_i) = a_q \quad (q = 1, ..., Q), \qquad (1)$$

are known (with ϕ_q *known* functions and $Q << M$) : the distribution $(p_i)_{i = 1, ..., M}$ is, of course, not determined by data a_q $(q = 1, ..., Q)$.

MAXENT: *the relevant distribution* (p_i) *makes S a maximum with the Q constraints (1).*

Amounts to selecting the *broadest* probability distribution compatible with the available information.

Statistical Mechanics:

Consider a system of a huge number N of "elementary constituents" (*e.g.* molecules in the kinetic theory of gases).

The *micro-state* (velocity + position) of each molecule is in one among M possible boxes, with $1 << M << N$.

Let l_i ($i = 1, ..., M$) = number of molecules in box (i).
Corresponding fraction: $p_i = l_i/N$ ($p_1 + ... + p_M = 1$).

Macro-state: pressure P, density ρ, temperature T, ...
Each should be computable from the probability distribution (p_i), as the average (expectation) $\langle \phi \rangle$ of some known function ϕ.

A given probability distrib^n (p_i) obtainable by a large number of distinct *configurations* [= the mapping: molecule $\rightarrow$ box (i)].

Stat. Mech. Hypothese: "real" distribution = the one that may be obtained by the largest number of distinct configurations.

But: possible configs. must be compatible with given macro-state.
$\Rightarrow$ The most general version of this hypothese is MAXENT.

2. Implementation of MAXENT in physics of heterogeneous media

Consider a heterogeneous medium, *e.g.* a *polycrystal*, with 2 micro fields: stimulus and response, *e.g. strain-rate* $\mathbf{d}$ and *stress* $\boldsymbol{\sigma}$.
(or: porous medium with pressure gradient and filtration velocity, ...)

Ideal (optimistic) aim of the *macro-to-micro* transition: determine the microfields $\mathbf{d}(\mathbf{x})$, $\boldsymbol{\sigma}(\mathbf{x})$ from data of macro-stimulus, say $\mathbf{D}$.

For a polycrystal, heterogeneity due to anisotropy of crystal, *i.e.* micro constitutive law $\mathbf{d}(\mathbf{x})$-$\boldsymbol{\sigma}(\mathbf{x})$ known from local orientation $\mathbf{R}(\mathbf{x})$

Discretize the existing (evolving) orientations :

at time t, $\mathbf{R} \in \{\mathbf{R}_1(t), ..., \mathbf{R}_n(t)\}$.

The volume fractions f_k ($k = 1, ..., n$) of the orientations are given (initial *texture* $\Leftrightarrow$ data of $(f_1, \mathbf{R}_1(t_0)), ..., (f_n, \mathbf{R}_n(t_0))$.

The fractions f_k are conserved (incompressibility condition).
Yet the texture evolves due to the evolution of the orientations.

A more modest, but attainable aim for the macro-to-micro transition (*e.g.* in a polycrystal):

Try to calculate the list $(\mathbf{D}_k)_{k = 1, \ldots, n}$
with $\mathbf{D}_k = \mathbf{D}_k(t)$ the *average strain-rate* in the orientation $\mathbf{R}_k(t)$

Then using the constitutive law for this orientation, the (average) stress $\boldsymbol{\sigma}_k(t)$ is obtained.
The (average) rotation rates $\boldsymbol{\Omega}_k(t) = \dot{\mathbf{R}}_k \mathbf{R}_k(t)$ are also obtained, hence an evolution (averaged over each orientation) is got.

Now consider a fixed time t.
How to use MAXENT so as to calculate the distribution $(\mathbf{D}_k)$?

Micro-state: (crystal orientation **R** + strain-rate **d**), function of **x**.
Orientation already discretized: discretize also strain-rate.

Thus the strain-rate field has been discretized:

$$\mathbf{d} \in \{\mathbf{D}^1, ..., \mathbf{D}^m\},$$

as well as the orientations: $\mathbf{R} \in \{\mathbf{R}_1, ..., \mathbf{R}_n\}$.

Hence μ-state $(\mathbf{d}, \mathbf{R}) \in \{\mathbf{D}^1, ..., \mathbf{D}^m\} \times \{\mathbf{R}_1, ..., \mathbf{R}_n\}$ so $M = m \times n$.

Let l^j_k = number of « elementary crystals » with μ-state $(\mathbf{D}^j, \mathbf{R}_k)$:

$\sum_{j,k} l^j_k = N$. Relevant probability distribution: $p^j_k = l^j_k / N$.

Thus p^j_k = volume fraction of polycrystal where

orientation $\mathbf{R}(\mathbf{x}) = \mathbf{R}_k$ and strain-rate such that $\|\mathbf{d}(\mathbf{x}) - \mathbf{D}^j\|_\infty \leq \varepsilon/2$

(ε is the size of the « **d**-grid »)

Volume average of strain-rate in orientation $\mathbf{R}_k$:

$$\mathbf{D}_k = (p^1_k \mathbf{D}^1 + ... + p^m_k \mathbf{D}^m)/f_k.$$

Obvious constraints [imposed to $(p^j{}_k)$] for MAXENT:

1) The volume fraction of polycrystal in the orientation $\mathbf{R}_k$ is the data f_k. Since the « **d**-grid » covers all possible values for **d**(**x**), each « elementary crystal » in the orientation k must find a $\mathbf{D}^j$, *i.e.*

$$\sum_{j=1}^{m} p^j{}_k = f_k \qquad (k = 1, ..., n). \tag{1}$$

2) The average strain-rate is the applied macro strain-rate **D** :

$$\sum_{j,k} p^j{}_k \, \mathbf{D}^j = \mathbf{D}. \tag{2}$$

Hence we may define a model based on MAXENT as follows:

Maximize $S = - \sum_{k=1}^{n} \sum_{j=1}^{m} p^j{}_k \, \mathrm{Log}\, p^j{}_k$ under constraints (1) and (2).

Since the statistical entropy S is a measure of disorder, the latter « MAXENT model with obvious constraints » may be called the « Volume-fraction model with maximum disorder ».

Now it is often said that « the self-consistent models describe a situation with perfect disorder » (in the sense that spatial correlations of a finite range do not exist).

It is hence natural to wonder whether the « obvious MAXENT model » is something like a self-consistent model.
However, we can prove that:

3. « Obvious MAXENT model » = Voigt-Taylor model !!

I.e., this model leads to the following prediction for the volume average $\mathbf{D}_k$ of the strain-rate in orientation $\mathbf{R}_k$:

$$\mathbf{D}_k = \mathbf{D} \quad \text{for all } k = 1, ..., n.$$

Sketch of the proof

Use method of Lagrange multipliers: search stationary point of

$$\Phi \equiv -\sum_{j,k} p^j{}_k \operatorname{Log} p^j{}_k - \sum_{k=1}^{n} \lambda_k \left(f_k - \sum_{j=1}^{m} p^j{}_k \right) - \sum_{l=1}^{6} \mu_l \left(D_l - \sum_{j,k} p^j{}_k D^j{}_l \right)$$

S constraint (1) constraint (2)

($D_l = l^{\text{th}}$ component of $\mathbf{D}$, $D^j{}_l = l^{\text{th}}$ component of $\mathbf{D}^j$)

Writing that $\dfrac{\partial \Phi}{\partial p^j{}_k} = 0$ leads to

$$p^j{}_k = e^{\lambda_k - 1} \exp\left(\sum_l \mu_l D^j{}_l \right). \qquad (3)$$

With (3), constraint (1) allows us to eliminate λ_k by

$$e^{\lambda_k - 1} = f_k \,/\, \sum_{j=1}^{m} \exp\left(\sum_{l=1}^{6} \mu_l D^j{}_l \right). \qquad (4)$$

Calculating $D_{k\,l} \equiv (\mathbf{D}_k)_l = \sum_j p^j{}_k D^j{}_l \,/\, f_k$ using (3) and (4) gives

$$D_{k\,l} = \sum_{j=1}^{m} \frac{D^j{}_l \exp\left(\sum_{l=1}^{6} \mu_l D^j{}_l\right)}{\sum_{j'} \exp\left(\sum_{l=1}^{6} \mu_l D^{j'}{}_l\right)}.$$

Since this does not depend on k and since $\sum_k f_k \mathbf{D}_k = \mathbf{D}$,

it follows that $\mathbf{D}_k = \mathbf{D}$ for all k, as announced.

Thus, the polycrystal model associated with the « maximum disorder » is the Voigt-Taylor model, a rather crude one.

Can we build a better model also based on MAXENT ?

4. A less trivial MAXENT model (one more constraint)

To obtain a better model: add information, *i.e.* add constraint(s).

Possible new constraint: average potential known,

$$\langle u \rangle \equiv \sum_k f_k \, u_k(\mathbf{D}_k) \equiv \sum_k f_k \; u_k\left(\sum_j p^j{}_k \, \mathbf{D}^j \; / \; f_k \right) \equiv U(\mathbf{D}) \text{ known} \quad (5)$$

(= minimum information to add in order to determine the macroscopic behavior).

Means the *micro-to-macro* transition is solved.

« Inhomogeneous variational model » well-suited in this context:

Data $U(\mathbf{D})$ replaced by data of average heterogeneity h, with

$$h^p \equiv \sum_{k=1}^{n} f_k \left\| \mathbf{D}_k - \mathbf{D} \right\|^p$$

It has been shown that the *macro-to-micro* transition, *i.e.*, determining the distribution $(\mathbf{D}_k)_{k = 1, ..., n}$ from data $\mathbf{D}$ (+ ...), is very close in that model and in the above MAXENT model.

Yet it's simpler to impose directly the average heterogeneity! For

$$h^p = \sum_{k=1}^{n} \left\| \sum_{j=1}^{m} p^j{}_k \, \mathbf{D}^j - f_k \mathbf{D} \right\|^p \; / \; f_k{}^{p-1}.$$

Conclusions

1) A general formulation of the Maximum Entropy Principle (MAXENT) has been given for the macro-to-micro transition in a heterogeneous medium.
This formulation was illustrated for a textured polycrystal.

2) MAXENT demands constraints. The most obvious ones (volume fractions imposed + consistency condition ⟨micro-stimulus⟩ = macro-stimulus) lead just to Voigt-Taylor model.

3) Imposing ⟨micro-potential⟩ gives a model close to the inhomogeneous variational model. But it seems wiser to impose directly the average heterogeneity.

4) MAXENT provides a general method to build more and more accurate models by adding information (= constraints).