Pergamon

Int. J. Solids Structures Vol. 33, No. 24, pp. 3549–3560, 1996

0020–7683/96 $15.00 + .00

0020–7683(95)00193–X

ENTROPY CHARACTERIZATION OF FINITE ELEMENTS

ABRAHAM I. BELTZER
Center for Technological Education Holon, affiliated to the Tel-Aviv University, Tel-Aviv, Israel

(*Received* 15 *February* 1995; *in revised form* 3 *August* 1995)

Abstract—The paper introduces the concept of entropy of a finite element. The general approach is based on the maximum-entropy (minimum-bias) principle. The results obtained for basic elements show that this entity may provide a quantitative evaluation of the complexity of a finite element in the context of the adopted physical model and serve as a parameter in its overall characterization.

1. INTRODUCTION

Since Shannon introduced the entropy as the central concept of information theory in his celebrated paper (Shannon, 1948), it has been extensively used in various fields. The entropy quantifies the degree of uncertainty (and thereby the value of information to be obtained) which is relevant for numerous situations requiring the decision making processes. As finite element analysis involves a certain selection procedure of a suitable element among the available ones, this concept may be useful in the context of this analysis.

The variety of finite elements available [see, for example, Zienkiewicz (1977), Kardestuncer (1987), Beltzer (1990), MacNeal (1994)], which is perhaps the most attractive feature of the finite element method, poses a problem of their comparative quality or performance. This problem may be approached from various points of view, such as simplicity of the finite element, its completeness, conformability, etc. There are aspects of the finite element performance which still lack a quantative evaluation and introducing the entropy concept may help fill the gap.

The present work deals with the definition and evaluation of the information "stored" by a finite element. This quantity may adequately describe what is usually called the complexity of the element, in a sense that a more complex element stores more information. Accordingly, it may provide an additional parameter for evaluation of the quality of a finite element, and may be useful at the stage of a finite element design as well as for the classification of finite elements.

The treatment below is based on the maximum-entropy (or minimum-bias) principle, which states that the most likely probability distribution follows from maximization of the entropy subjected to the given constraints. Jaynes (1957, 1979) provided a particularly elegant formalism for implementing this principle to physical systems. Applications to evaluation of the accuracy of approximate methods (Ritz, Galerkin) were given by Beltzer (1973, 1974). The present work and short communication by Beltzer and Gotlib (1995) are sequels to the above works. For an extensive exposition of the entropy concept, the reader is referred to the monograph by Kapur (1989).

Section 2 below deals with the derivation of basic equations and relevant considerations. It is shown that the account for the prior constraints (such as estimates and bounds) on the value of the unknown field gives rise to the finite value of the entropy associated with the finite element. This is carried out with the help of the so-called hypervolume which arises in the multi-dimensional space of the nodal values when proper constraints are taken into account. Section 3 deals with calculations of the hypervolume for some typical finite elements and Section 4 with calculations of their entropy. Conclusions are given in Section 5.

2. BASIC EQUATIONS

The finite element is characterized by its geometry and shape functions. For the sake of certainty, consider the assumed dimensionless displacement field, U, and set the basic relation

$$U(x) = \sum_{i=1}^{n} u_i g_i(x) \tag{1}$$

with u_i being the unknown nodal values (degrees of freedom) of a finite element and $g_i(x)$ the given shape functions. As far as the nodal values u_i are unknown, one may ascribe uncertainty (entropy) to eqn (1). The greater the uncertainty, the greater value of information emerges when the nodal values u_i are specified. Consequently, from this viewpoint, a finite element should be so designed as to ensure a maximal prior uncertainty (or minimal bias) under the given constraints. This paves the way for application of Jaynes' principle.

To this end, introduce the n-dimensional space of the nodal values $\{\mathbf{u}|\mathbf{u} \in \mathrm{R}^n\}$ and the k-dimensional coordinate space $\{\mathbf{x}|\mathbf{x} \in \mathrm{R}^k\}$. The way of arriving at the finite value of the uncertainty (entropy) associated with the representation of eqn (1) is the account for the prior information (constraints) available on $U(x)$. Depending on a particular problem, the prior constraints may include bounds or estimates for, say, the average field of U, its maximal value, etc. We consider below the case when the prior constraint on $U(x)$ is as follows:

$$\int_S U^2(x)\,\mathrm{d}\mathbf{x} \leqslant Q, \tag{2}$$

where S is the volume of the finite element in R^k.

The value of Q depends on a particular model under consideration. For example, if the problem deals with a plate bending, then $U(x)$ is the deflection function in the frameworks of the linear theory, and Q may be estimated as

$$Q = \mathrm{O}(\sigma^2\Omega), \tag{3}$$

where $\sigma \ll 1$ is a small parameter, and Ω is a typical plate area.

Applying the above maximum entropy principle to eqns (1) and (2), one arrives at the following formulation: find the joint probability density function $p(\mathbf{u}) = p(u_1, \ldots u_n)$ of the nodal values which delivers the maximum to the entropy functional

$$H = -\int_{\mathrm{R}^n} p(\mathbf{u}) \ln p(\mathbf{u})\,\mathrm{d}\mathbf{u} \tag{4}$$

under the constraint (2)

$$\int_S U^2(x)\,\mathrm{d}\mathbf{x} \leqslant Q \tag{5}$$

and the normalization condition for $p(\mathbf{u})$

$$\int_{\mathrm{R}^n} p(\mathbf{u})\,\mathrm{d}\mathbf{u} = 1. \tag{6}$$

Note that eqn (4) may contain an additive arbitrary constant.

On the substitution of the expansion (1) in eqn (5) one gets the constraint

$$\sum_{i=1}^{n}\sum_{j=1}^{n} B_{ij}u_iu_j \leqslant Q \tag{7}$$

with the coefficients

$$B_{ij} = \int_S g_i(x)g_j(x)\,\mathbf{d}\mathbf{x}. \tag{8}$$

The expression

$$\sum_{i=1}^{n}\sum_{j=1}^{n} B_{ij}u_iu_j = Q \tag{9}$$

is a symmetric quadratic form in the space of nodal values **u**. The geometrical type associated with eqn (7) depends on the eigenvalues of its symmetrical matrix **B**. If all the eigenvalues are positive (and naturally real ones), then this quadratic form specifies an ellipsoid.

Given the conditions (5) and (6) only, the extremizer of eqn (4) is the uniform distribution over the volume V in the n-dimensional space as specified by the form of eqn (7). Therefore $p(\mathbf{u})$ = constant = V^{-1}. Substituting this probability density $p(\mathbf{u})$ in eqn (4), one gets the expression for the prior entropy given by

$$H = -\int_V \frac{1}{V}\ln\frac{1}{V}\,\mathbf{du} = \ln V. \tag{10}$$

In case of the n-dimensional ellipsoid, which is the one treated herein, its volume measure V_n is given by (Fikhtengolts, 1965)

$$V_n = \frac{\pi^{n/2}}{(n/2)\Gamma(n/2)}\prod_{i=1}^{n} a_i, \tag{11}$$

where $a_i (i = 1, \ldots, n)$ are the semi-axes of the ellipsoid. This may be expressed in terms of the above **B** matrix in a more direct way. Let S be the element volume and set

$$\mathbf{C} = \mathbf{B}/S. \tag{12}$$

Now eqn (9) takes the form

$$\sum_{i=1}^{n}\sum_{j=1}^{n} C_{ij}u_iu_j = q \tag{13}$$

with $q = Q/S$, or in the main axes

$$\sum_{i=1}^{n} \lambda_i\tilde{u}_i^2 = q \tag{14}$$

where λ_i are the eigenvalues of **C**. Thus

$$\prod_{i=1}^{n} a_i = q^{n/2}\prod_{i=1}^{n}\left(\frac{1}{\lambda_i}\right)^{1/2} \tag{15}$$

and eqn (11) may be conveniently set as

$$V_n = \frac{\pi^{n/2} q^{n/2}}{(n/2)\Gamma(n/2)} (\det \mathbf{C})^{-1/2}. \tag{16}$$

For the case of two nodes, this quantity is merely the area and for that of three nodes the volume. Nevertheless, since the nodal space R^n is, in general, n-dimensional, the above quantity will be referred to below as the hypervolume.

Thus, the prior entropy (10) is given by

$$H_n = \ln V_n \tag{17}$$

where the subscript n indicates the number of nodal values (degrees of freedom) and may also be considered as a constraint. If the nodal values u_i are then specified with absolute precision, there would be no residual entropy H^{res} and eqn (17) also yields the value of information obtained. However, this is not the case as there are inevitable (inherent) inaccuracies, due to round-off errors, approximate integration, etc. Consequently, the final entropy of the element is given by

$$H_n^{fin} = H_n - H_n^{\text{res}} \tag{18}$$

which provides the value of the information obtained when the nodal values u_i are specified. In addition, it seems natural to define the entropy (information) per nodal value by

$$h^{fin} = \frac{H_n^{fin}}{n}. \tag{19}$$

The next section deals with computation of the hypervolume as given by eqns (11) and (16) for basic finite elements.

3. HYPERVOLUME

Figure 1 shows the types of finite elements to be considered herein in a subsequent order.

(1) It would be instructive to begin with a simple linear element. In the case of two nodes ($k = 1$, $n = 2$) the shape functions are

$$g_1(x) = \frac{x}{S} \text{ and } g_2(x) = 1 - \frac{x}{S} \tag{20}$$

with S being the element length. The quadratic form of eqn (13) is reduced to

$$\tfrac{1}{3}u_1^2 + \tfrac{1}{3}u_1 u_2 + \tfrac{1}{3}u_2^2 = q, \tag{21}$$

where $q = Q/S$ and all the coefficients follow from eqns (8) and (12). The expression (21) describes an ellipse. Its area (hypervolume) V is

$$V_2 = 2\pi\sqrt{3}q \approx 10.883q. \tag{22}$$

In the case of a three-node linear element ($k = 1$, $n = 3$), the question arises about an optimal

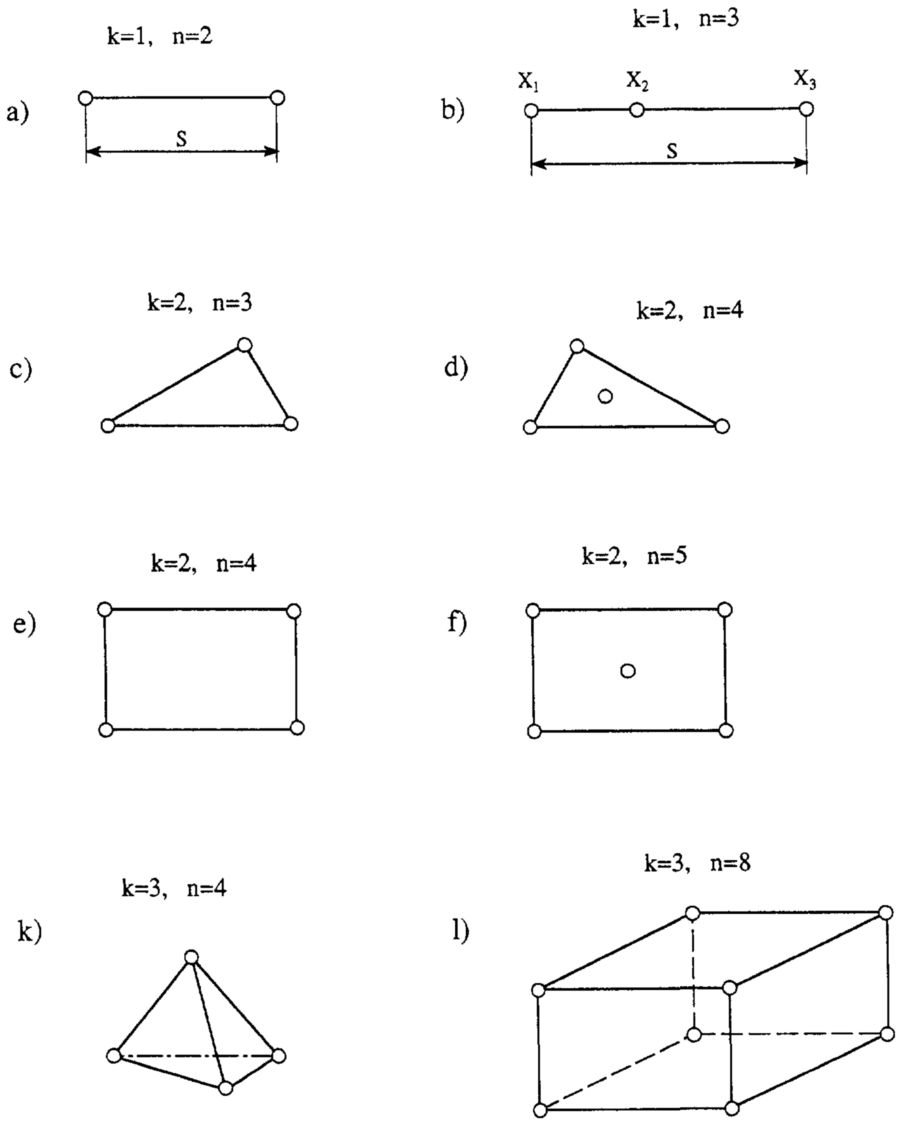

Fig. 1. Types of finite elements.

location of the internal "node". Although it is well-known that the location $x = S/2$ is in general the best one, this author is not aware of any pertinent quantitative analysis.

To this end, set three shape functions as

$$g_1(x) = \frac{(x-x_2)(x-x_3)}{(x_1-x_2)(x_1-x_3)}$$

$$g_2(x) = \frac{(x-x_1)(x-x_3)}{(x_2-x_1)(x_2-x_3)}$$

$$g_3(x) = \frac{(x-x_1)(x-x_2)}{(x_3-x_1)(x_3-x_2)}. \tag{23}$$

Specifying $x_1 = 0$, $x_3 = S$ and $x_2 = \delta S$ with $0 < \delta < 1$, get from eqn (8) the matrix $\mathbf{B} = [B_{ij}]$ as follows:

$$S\begin{bmatrix} \frac{1}{30}\frac{-5\delta+1+10\delta^2}{\delta^2} & -\frac{1}{60}\frac{-2+5\delta}{\delta^2(\delta-1)} & \frac{1}{60}\frac{3+10\delta^2-10\delta}{\delta(\delta-1)} \\ -\frac{1}{60}\frac{-2+5\delta}{\delta^2(\delta-1)} & \frac{1}{30}\frac{1}{\delta^2(\delta-1)^2} & -\frac{1}{60}\frac{-3+5\delta}{\delta(\delta-1)^2} \\ \frac{1}{60}\frac{3+10\delta^2-10\delta}{\delta(\delta-1)} & -\frac{1}{60}\frac{-3+5\delta}{\delta(\delta-1)^2} & \frac{1}{30}\frac{6+10\delta^2-15\delta}{(\delta-1)^2} \end{bmatrix}. \tag{24}$$

This provides the determinant of **C** [see eqn (12)] as

$$\det \mathbf{C} = \frac{1}{2160\delta^2(\delta-1)^2}. \tag{25}$$

The substitution of this simple expression in eqn (16) shows that $\delta = 1/2$, $(0 < \delta < 1)$, provides the strong maximum to V_3 and thereby specifies the optimal location of the internal node, in agreement with our intuition.

Completing the case of the three-node linear element for $\delta = 1/2$, get the eigenvalues λ_i, $i = 1, 2, 3$ of the matrix $15\mathbf{C}$ as

$$\lambda_1 = 1.2056, \quad \lambda_2 = 2.5000, \quad \lambda_3 = 8.2944 \tag{26}$$

and the semi-axes as

$$a_i = \left(\frac{15q}{\lambda_i}\right)^{1/2}, \quad i = 1, 2, 3. \tag{27}$$

Now eqn (11) or (16) yields the relevant volume as

$$V_3 = 48.669q^{3/2} \tag{28}$$

instead of the result for the two-node element given by eqn (22).

(2) In the case of the three-node triangular element ($k = 2$, $n = 3$) the computations are particularly simple in the area coordinates α_i, $i = 1, 2, 3$

$$U = \sum_{i=1}^{3} u_i \alpha_i. \tag{29}$$

The **C** matrix gets the form

$$\mathbf{C} = \frac{\mathbf{B}}{S} = \frac{1}{12}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \tag{30}$$

and the semi-axes are given by

$$a_i = \left(\frac{12q}{\lambda_i}\right)^{1/2}, \quad i = 1, 2, 3 \tag{31}$$

with

$$\lambda_1 = 4, \quad \lambda_2 = 1, \quad \lambda_3 = 1. \tag{32}$$

This yields for the three-node triangle

$$V_3 = 87.063\, q^{3/2}. \tag{33}$$

Considering the four-node triangle, place a "node" at its centroid with the coordinates $\alpha_i = 1/3$, $i = 1, 2, 3$. Now the nodal representation for the field U is

$$U = \sum_{i=1}^{3} u_i \alpha_i + u_4 27 \alpha_1 \alpha_2 \alpha_3 \tag{34}$$

where the last term reflects the contribution of the "bubble" function.

The $\mathbf{C}$ matrix gets the form

$$\mathbf{C} = \begin{bmatrix} 1/6 & 1/12 & 1/12 & 3/20 \\ 1/12 & 1/6 & 1/12 & 3/20 \\ 1/12 & 1/12 & 1/6 & 3/20 \\ 3/20 & 3/20 & 3/20 & 81/280 \end{bmatrix} \tag{35}$$

with the eigenvalues

$$\lambda_1 = 0.57205, \quad \lambda_2 = \lambda_3 = 0.08333, \quad \lambda_4 = 0.05057. \tag{36}$$

The hypervolume is

$$V_4 = 348.17\, q^2 \tag{37}$$

as compared to that of the three-node triangle given by eqn (33).

(3) For the four-node rectangular element $k = 2$ and $n = 4$. The shape functions are

$$g_1(x_1, x_2) = \frac{(x_1 - b)x_2 - ax_1 + ab}{4ab}, \quad g_2(x_1, x_2) = \frac{(x_1 + b)x_2 - ax_1 - ab}{4ab},$$
$$g_3(x_1, x_2) = \frac{(x_1 + b)x_2 + ax_1 + ab}{4ab}, \quad g_2(x_1, x_2) = \frac{(x_1 - b)x_2 + ax_1 - ab}{4ab}, \tag{38}$$

where $2a$ and $2b$ specify the size of the rectangle in the symmetry reference frame. Equation (13) is given by

$$2(u_1^2 + u_2^2 + u_3^2 + u_4^2) + 2u_1u_2 + u_1u_3 + 2u_1u_4 + 2u_2u_3 + u_2u_4 + 2u_3u_4 = 18q. \tag{39}$$

The eigenvalues are positive with their product equal to 81. The hypervolume is

$$V_4 = 72\pi^2 q^2 \approx 710.61 q^2. \tag{40}$$

Turning to the five-node rectangular element ($k = 2$, $n = 5$) introduce a "bubble" shape function

$$g_5(x_1, x_2) = \frac{(x_1^2 - b^2)(x_2^2 - a^2)}{(ab)^2}. \tag{41}$$

In order to evaluate the additional entropy due to this modification, compute the **C** matrix, which is

$$\mathbf{C} = \begin{bmatrix} 1/9 & 1/18 & 1/36 & 1/18 & 1/9 \\ 1/18 & 1/9 & 1/18 & 1/36 & 1/9 \\ 1/36 & 1/18 & 1/9 & 1/18 & 1/9 \\ 1/18 & 1/36 & 1/18 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 & 1/9 & 64/225 \end{bmatrix}. \tag{42}$$

Now, eqn (11) or (16) provides the hypervolume as

$$V_5 = 2571.1q^{5/2} \tag{43}$$

instead of the result given by eqn (40) for the four-node rectangle.

(4) Turning to the three-dimensional case, consider the four-node tetrahedron element ($k = 3$, $n = 4$). This time

$$U = \sum_{i=1}^{4} u_i \alpha_i \tag{44}$$

where α_i, $i = 1, 2, 3, 4$ are the volume coordinates. The **C** matrix is given by

$$\mathbf{C} = \frac{1}{20} \begin{bmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{bmatrix} \tag{45}$$

with the eigenvalues $\lambda = 1/4$ and $\lambda_2 = \lambda_3 = \lambda_4 = 1/20$. According to eqn (16) the hypervolume is

$$V_4 = 882.76q^2. \tag{46}$$

Finally, consider the eight-node brick element. In the natural coordinates ε, η and ζ the shape functions are given by

$$g_i = (1 + \varepsilon\varepsilon_i)(1 + \eta\eta_i)(1 + \zeta\zeta_i)/8 \quad i = 1, 2, \ldots 8 \tag{47}$$

where ε_i, η_i and ζ_i are the nodal coordinates. The eigenvalues of the 8×8 **C** matrix are

$$\lambda_1 = 0.12500, \quad \lambda_2 = \lambda_3 = \lambda_4 = 0.04167, \quad \lambda_5 = \lambda_6 = \lambda_7 = 0.01389, \quad \lambda_8 = 0.00463 \tag{48}$$

and

$$\det \mathbf{C} = 1.12156 \times 10^{-13}. \tag{49}$$

Now eqn (16) yields

$$V_8 = 1.2119\times 10^7 q^4. \tag{50}$$

Now the entropy of these elements easily follows from eqns (17)–(19).

4. ENTROPY

In order to avoid the scale effect, the above field U has been taken dimensionless

$$U = W/L, \tag{51}$$

where W is a physical displacement and L is a typical size of the finite element. In general, it may be put $L^k = S$ where $k = 1, 2, 3$, depending on the dimensions of the finite element. As to the value of Q appearing in constraint (2), it depends on a particular model under consideration and may be specified in various ways. Assuming that the finite element is not singular and the problem is that of linear elasticity, a basic constraint may be set as

$$|W| < \varepsilon L \tag{52}$$

where $\varepsilon \ll 1$ is a small parameter typical of the problem at hand. The incorporation of this constraint, as it is, does not seem to be easy. It is convenient to reformulate eqn (52) in a weaker form given by eqn (2), which yields $Q = \varepsilon^2 L^k$, $k = 1, 2, 3$ depending on the dimension of the finite element. This provides for all the elements

$$q = \frac{Q}{S} = \varepsilon^2 \tag{53}$$

and enables one to adapt the above results for the hypervolume to this particular model by substituting ε^2 for q. Then the corresponding prior entropy H_n follows from eqn (17).

$$H_n = \ln V_n. \tag{54}$$

As to the residual entropy H^{res}, it characterizes the accuracy of a particular method of solution applied to specify the nodal values u_i, as noted in Section 2 and thus depends on the problem at hand. Its exact evaluation appears difficult, and comments on the order of this value are given immediately below. Since the order of magnitude of u_i is $O(\varepsilon)$, as eqns (51) and (52) show, the error ϕ in specifying the nodal values may be set as $O(\varepsilon^m)$ with $m > 2$, unless there is a more specific knowledge of this inherent inaccuracy. Taking into account the dimension n of the space of the nodal values, the residual entropy is given by

$$H_n^{\text{res}} = \ln(\phi^n) = n\ln(\varepsilon^m). \tag{55}$$

The final value of the entropy (or of the obtained information) now follows from eqn (18)

$$H_n^{\text{fin}} = \ln V_n - n\ln\phi \tag{56}$$

and that of the entropy per nodal value from eqn (19)

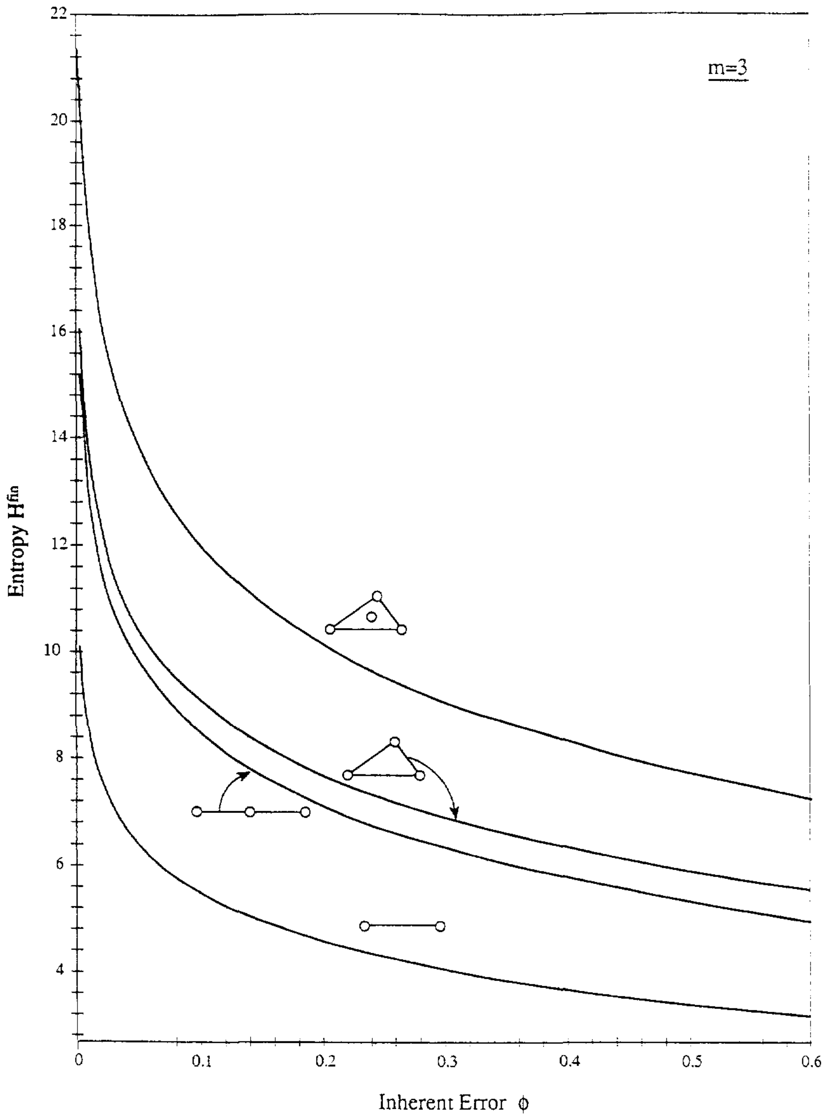

Fig. 2. Entropy versus inherent error.

$$h_n^{fin} = \frac{\ln V_n}{n} - \ln \phi. \tag{57}$$

Figures 2 and 3 show the results for H_n^{fin} and Fig. 4 for h_n^{fin} for $0 < \phi < 0.6$ and $m = 3$. There is a well-defined disparity in the values of the entropy for the different elements. A smaller inherent error ϕ leads to a greater value of the entropy as expected. On the other hand, there is decay in the entropy with increasing ϕ.

The results consistently show that for the same number of the nodal values and the same value of the inherent error, the plane element provides more information than the linear one and the solid element provides more than the plane one. A comparison of the four-node triangular and four-node rectangular elements shows that they provide information of the same order but the latter is more informative. In other words, the four-node rectangular element is "less biased" than the four-node triangular element. Among the elements considered, the brick element has by far the best entropy and the entropy per nodal value. Note that the above results are valid under constraint (2) and may change if other constraints come into play.

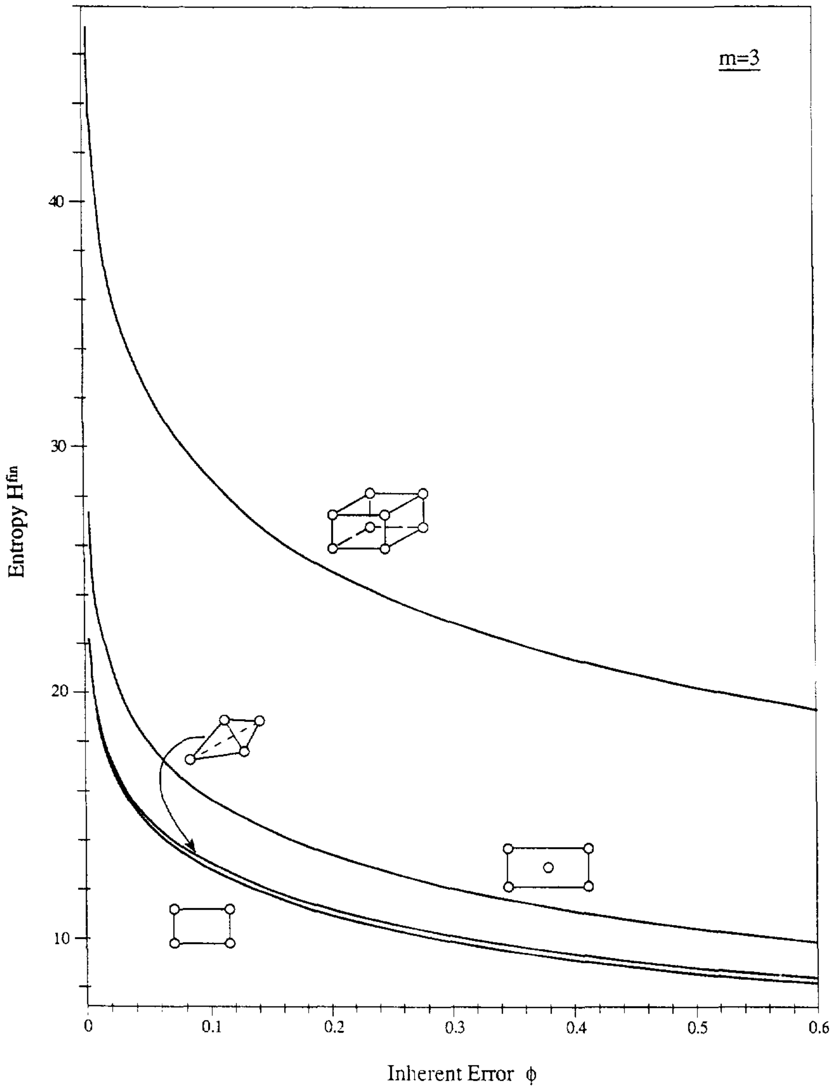

Fig. 3. Entropy versus inherent error.

5. CONCLUSIONS

The concept of entropy (information) of a finite element, introduced in the above considerations, depends, on the one hand, on its number of degrees of freedom, shape functions and geometry, in agreement with our intuition. On the other hand, it is influenced by a particular physical model under consideration, which comes heavily into play through prior constraints on the unknown field. It thus provides a possibility for a quantitative evaluation of the complexity (lack of bias) of the finite element in the frameworks of the adopted physical model and may serve as a parameter in its overall characterization. This in no way diminishes the role of other criteria such as completeness, conformability, etc., which should be considered from the very outset. As the example with the three-node linear element indicates, the entropy evaluation may be particularly useful for optimizing the node location. Though the computations deal with a particular type of constraint given by eqn (2), the presented approach remains valid for other types of restrictions too. In particular, the account for stiffer, more specific constraints should bring the entropy down. This case merits further investigations.

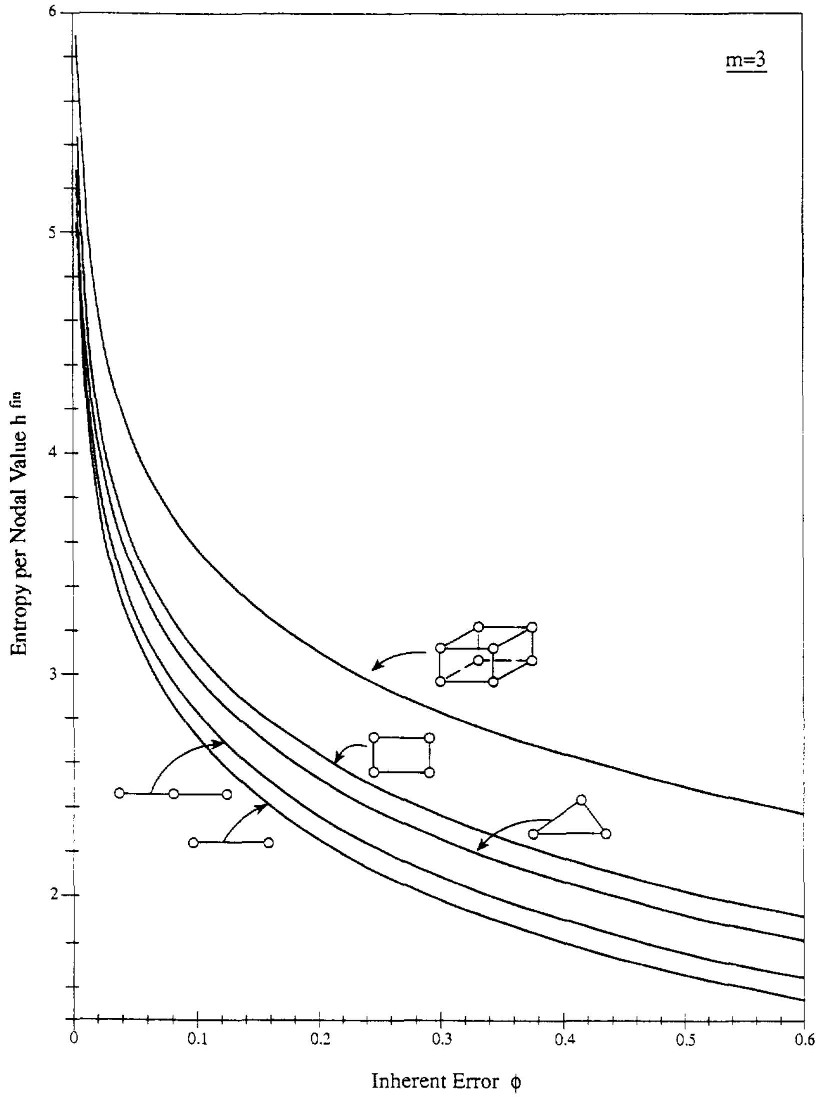

Fig. 4. Entropy per nodal value versus inherent error.

REFERENCES

Beltzer, A. I. (1973). Application of information theory to analysis of approximate solutions. In *Problems of Dynamics and Strength* (in Russian: *Voprosy Dynamiki e Prochnosti*), Vol. 27, pp. 119–125. Riga Polytechnic Inst.

Beltzer, A. I. (1974). Evaluation of informational properties of approximate solutions. In *Problems of Dynamics and Strength* (in Russian: *Voprosy Dynamiki e Prochnosti*), Vol. 28, pp. 108–114. Riga Polytechnic Inst.

Beltzer, A. I. (1990). *Variational and Finite Element Methods—A Symbolic Computation Approach*. Springer, Berlin.

Beltzer, A. I. and Gotlib, V. A. (1995). Assessment of informational properties of finite elements. In *Proc. Int. Conf. Computational Engineering Science, ICES-95*, 30 July–3 August 1995, Hawaii, U.S.A.

Fikhtengolts, G. M. (1965). *The Fundamentals of Mathematical Analysis*, Vol. II, p. 403. Pergamon Press, NY.

Jaynes, E. T. (1957). Information theory and statistical mechanics. *Phys. Revs.* **106,** 620–630; **108,** 171–197.

Jaynes, E. T. (1979). Where do we stand on maximum entropy? In *Maximum Entropy Formalism* (edited by R. D. Levine and M. Tribus), pp. 15–118. MIT Press, Cambridge, MA.

Kapur, J. N. (1989). *Maximum Entropy Models in Science and Engineering*. John Wiley, NY.

Kardestuncer, H. and Norrie, D. H. (Eds). (1987). *Finite Element Handbook*. McGraw-Hill, NY.

MacNeal, R. H. (1994). *Finite Elements: Their Design and Performance*. Marcel Dekker, NY.

Shannon, B. D. (1948). A mathematical theory of communication. *Bell System Tech. J.* **27,** 379–423, 623–659.

Zienkiewicz, O. C. (1977). *The Finite Element Method*. McGraw-Hill, London.