From Formal Solutions to Methods Avoiding Passages to the Limit

E. R. de Arantes e Oliveira[1]

1 From formal solutions to differential methods

Before the advent of digital computers, *formal solutions* were the only available solutions to differential equations.

Formal solutions can be *closed solutions* or *series solutions*. Closed solutions make it possible to express the unknown functions by formulae involving a finite number of algebraic operations. Numerical analysis is reduced to calculating the values of the unknowns appearing in such formulae.

Series solutions require an infinite number of algebraic operations. The truncation of the series becomes then necessary, and the concepts of *truncation error* and *convergence* are involved.

In any case, numerical analysis starts only after integration is performed.

A more basic classification of computational methods consists in distinguishing between those in which numerical analysis starts before or after integration. In case it starts before, computational methods may be called *differential computational methods*. In case it starts after, computational methods may be called *integral computational methods*.

Finite difference and *finite element methods* are examples of differential computational methods. They generate solutions that only approximately satisfy both the equations and their boundary conditions. Their application lead to the solution of systems of algebraic equations whose unknowns are associated to points, called *nodes*, some of which are located within the domain corresponding to the body, and some on the boundary of such domain.

Integral methods lead to solutions that exactly satisfy the field equations, but only approximately satisfy the boundary conditions. In principle, all formal methods should be integral methods. However, once the truncation of a series solution is performed, the solution generated by a series technique ceases to satisfy the differential equation. In other words, formal solutions involving an infinite number of algebraic operations cannot be used as integral methods because numerical analysts cannot perform such an

[1] Instituto Superior Técnico, Universidade Técnica de Lisboa, Portugal (arantes@civil.ist.utl.pt).

infinite number of operations. Among formal solutions, only the rare close solutions available for solving differential equations may lead to integral methods.

2 From finite difference methods to meshless methods

The classical finite difference method for solving differential equations on a given domain is characterised by covering the domain by a *mesh*, and replacing the derivatives by their difference approximations with respect to such mesh. The values of the function at the nodes of the mesh become the finite number of unknowns to be determined, and the approximate solution is expressed in terms of such unknowns.

It is important to point out that the effectiveness of a finite difference analysis is usually not assessed in terms of such error, but in terms of the error committed in replacing derivatives by differences. In other words, what is usually assessed is the precision with which the differential operator is approximated, and not, as it should be, and really is in the finite element method, the precision with which the inverse of such operator is approximated.

The author published an article as early as 1985 (Arantes e Oliveira, 1985) on the theory of finite difference methods, aiming to evaluate the approximation error as it is usually done in the case of the finite element method. In this paper, the author characterised the finite difference method by two essential features: a *local interpolation* around each node, and the use of the *collocation technique* for determining the approximate solution. No mention was needed to any kind of mesh. The classical finite difference technique being a particular case of such technique, this latter was called by the author the *generalised finite difference method*.

The generalised finite element method is truly a *meshless method*, as Belytschko (Belytschko and all., 1996) indicates in an excellent state of art on meshless methods. The author's paper is not quoted, perhaps because it does not involve the key-word *meshless*, which only much later started to be used.

If elements of this kind are used, the finite difference and the finite element methods simply differ in that the approximate solution is generated by *collocation* in the first and by *variational techniques* in the latter.

However, the author could show, in papers (Arantes e Oliveira, 1985 and 1981) aimed to evaluate the approximation/discretisation error associated to the finite difference method, that the approximate solution generated by collocation can be obtained also by resorting to variational techniques. Parallel descriptions of both methods become thus possible which contribute to clarify their similarities and distinctions.

Other authors, such as J. Cea (1964) and Raviart (1967), have attempted to present the finite difference method as a variational technique. Cea wrote on variational approximations to elliptical problems and considered the finite difference method as capable of providing one of such approximations. Raviart did the same for certain evolution equations. No analogy with the finite element method was established in such papers. However, the fact that the approximate solution generated by collocation can be obtained also by a variational technique make parallel descriptions of the finite difference and finite element methods possible, and allows a comparison between both which clarify their relative advantages and disadvantages.

The relative advantages of the finite element method are associated with the fact that the function derivatives involved in the functional are, in the finite difference case, of the same order as those involved in the differential equation, while they are of half such order in the finite element method. Owing to this circumstance, the co-ordinate functions are required to satisfy more severe conditions in the finite difference method than in the finite element method. The completeness conditions are also more severe in the finite difference method than in the finite element method, as a complete polynomial of the $(2p-1)^{th}$ degree is required to be contained in the expression of the function in the first, and a complete polynomial of the $(p-1)^{th}$ degree in the latter.

The relative advantages of the finite difference method stem from the fact that the associated approximate solution is obtainable using the collocation technique. The main advantage of the method is that, using the finite difference technique, the degree of accuracy depends exclusively on the degree of completeness, and, consequently, that no conditions other than the *completeness conditions*, are necessary for convergence. On the contrary, the finite element method requires the satisfaction of the completeness conditions and of a *supplementary condition*, like the satisfaction of the *patch test*.

3 From hybrid elements to block elements

In the beginnings of the finite element method, researchers looked for *conforming* finite elements, i.e., for finite elements systems providing approximate solutions in which compatibility was satisfied both within the elements and on the interfaces between elements. It was soon realised that the only advantage that conformity can bring was making the finite element method a particular case of the *Ritz method*, characterised by co-ordinate functions piecewise defined over the system of subdomains corresponding to the elements. Completeness was therefore a sufficient condition for convergence.

This is a good example of how a good computational technique can be rejected due to the insufficiency of the theoretical tools available. In other words, *non-conforming* elements started being rejected because their convergence theory had not yet been developed. However, the use of non-conforming finite elements represented a very important step in the history of computational analysis. A second important step was the implementation (Pian, 1964) of *hybrid elements*, i. e., of elements that were both non-conforming and non-compatible.

Some years ago, other types of elements, called *block elements*, were invented (Cundall, 1971; Kawai, 1971) that were themselves rigid, the deformation of the system being concentrated in the internal boundary elements, i. e., in the interfaces between elements. The advantage of such elements consisted in avoiding integrations over the element domains, and replacing them by integrations over the interfaces between elements.

Hybrid elements are such that:

i) the stress fields

$$\boldsymbol{s} = \boldsymbol{\chi}^e \, \boldsymbol{s}^e \tag{1}$$

allowed within the elements are defined through polynomial stress fields assumed to equilibrate vanishing body forces, and depending on a set of generalized stresses $\boldsymbol{s}^e$;

ii) the stresses $\boldsymbol{s}$ allowed within the elements are equilibrated but not necessarily compatible;
iii) the nodes are located on the element boundaries, so that each node belongs to several elements; the topology of a system of hybrid element can thus be defined by providing for each element the set of its nodes;
iv) the interaction between contiguous elements is defined by compatibility conditions expressed in terms of the nodal displacements $\boldsymbol{q}e$, such that

$$\boldsymbol{u} = \boldsymbol{\Phi} e\ \boldsymbol{q} e \quad . \tag{2}$$

Let us select a set of nodes on the hybrid element boundaries, and suppose that a set of nodal displacements is assigned to each node, such that the displacements along the segment between any two contiguous nodes on the element boundary can be defined in terms of the nodal displacements at the ends of such segment. Then:
i) displacements along the boundary can be expressed in terms of the element nodal displacements;
ii) the approximate solution associated to the decomposition of the body into hybrid elements is obtained by equalising the generalized displacements at the coinciding nodes of adjacent elements;
iii) hybrid elements are deformable, although rigidly connected along their boundaries;
iv) deformability in a system of hybrid elements is essentially associated to the elements themselves, not to the element interfaces.

In hybrid elements, *the field that minimizes the element total complementary energy on a linear space of allowed fields* (1) *is the one belonging to that space that is nearest to the exact solution - equilibrated and compatible - associated to given boundary displacements*. In other words, introducing (1) in the complementary energy of the element, differentiating and equating to zero, we obtain the *generalized stress-displacement equations*:

$$\boldsymbol{se} = \boldsymbol{He}\ \boldsymbol{De}\ \boldsymbol{ue}\,, \tag{3}$$

where

$$\boldsymbol{He} = [\ \int_{\boldsymbol{\Omega e}} \boldsymbol{\chi} e^{\mathrm{T}} \boldsymbol{H}^{1} \boldsymbol{\chi} e\ d\boldsymbol{\Omega}\ ]^{-1} \quad , \qquad \boldsymbol{De} = \int_{\boldsymbol{\partial\Omega e}} \boldsymbol{\chi} e^{\mathrm{T}}\ \boldsymbol{N}^{\mathrm{T}}\ \boldsymbol{\Phi} e\ d\boldsymbol{\partial\Omega e} \quad . \tag{4, 5}$$

The decomposition of the body into hybrid elements leads to a system of elements satisfying (1), which behaves more and more like a system of compatible finite elements, as the degree of the polynomial shape functions in (1) becomes larger and larger. Compatibility between elements is enforced by (2).

Compatibility between elements is enforced by introducing the system displacement vector, $\boldsymbol{U}$, and expressing vectors $\boldsymbol{u}$e as a linear function of $\boldsymbol{U}$, i.e., by writing

$$\boldsymbol{u}\mathrm{e} = \boldsymbol{A} e\ \boldsymbol{U}. \tag{6}$$

Therefore,

$$\boldsymbol{KU} = \boldsymbol{F}, \tag{7}$$

where

$$\boldsymbol{K} = \sum_e \boldsymbol{Ae}\, \boldsymbol{He}\, \boldsymbol{De} \tag{8}$$

or, as the allowed stresses are supposed to equilibrate vanishing body forces,

$$\boldsymbol{KU} = \boldsymbol{0}. \tag{9}$$

Block elements are such that:

i) the displacement fields

$$\boldsymbol{u} = \boldsymbol{\Psi e}\, \boldsymbol{ue} \tag{10}$$

allowed within the elements are defined through polynomial displacement fields continuous within each element, and depending on a set of generalized displacements $\boldsymbol{ue}$ assigned to a nodal point located inside the element;

ii) the fields allowed within the elements are compatible but not necessarily equilibrated;

iii) the topology of a system of block elements can thus be defined by providing, for each element, the set of adjacent elements;

iv) the interaction between adjacent elements is defined by equilibrium conditions expressed in terms of the generalized contact forces $\boldsymbol{re}$, such that

$$\boldsymbol{p} = \boldsymbol{\Gamma e}\, \boldsymbol{re} \quad . \tag{11}$$

Let us consider the block element boundaries decomposed into segments corresponding to the interfaces with the adjacent elements, and suppose that the tractions are distributed on each segment in such a way that they can be expressed in terms of generalized contact forces $\boldsymbol{re}$ associated to such segment. Then:

i) tractions acting on each segment can be expressed in terms of the generalized tractions;

ii) the approximate solution associated to the decomposition of the body into block elements is obtained by equilibrating the generalized tractions associated to coinciding segments of the adjacent element boundaries;

iii) block elements can themselves be rigid, although not rigidly connected to the adjacent elements along their common boundaries;

iv) deformability in a system of block elements is, therefore, essentially associated to the element interfaces, not necessarily to the elements.

In block elements, *the field that minimizes the element total potential energy on a linear space of allowed fields* (10) *is the one belonging to that space that is nearest to the exact solution - compatible and equilibrated - associated to given boundary tractions*. In other words, introducing (10) in the total potential energy of the element, differentiating, and equating to zero, we obtain the *generalized displacement-contact force equations*:

$$\boldsymbol{ue} = \boldsymbol{He}^{-1}\, \boldsymbol{Me}\, \boldsymbol{re} \quad , \tag{12}$$

where

$$\boldsymbol{He} = \int_{\Omega \boldsymbol{e}} (\boldsymbol{D}\,\boldsymbol{\Psi e})^{\mathrm{T}} \boldsymbol{H} (\boldsymbol{D}\,\boldsymbol{\Psi e})\, d\boldsymbol{\Omega} \qquad \boldsymbol{Me} = \int_{\partial \Omega \boldsymbol{e}} \boldsymbol{\Psi e}^{\mathrm{T}} \boldsymbol{\Gamma e}\, d\boldsymbol{\partial \Omega e} \quad . \tag{13,14}$$

The decomposition of the body into block elements leads to a system of elements satisfying (10), which behaves more and more like a system of equilibrated finite elements, as the degree of the polynomial shape functions in (10) becomes larger and larger. Equilibrium between elements is ensured by conditions (11).

However, the remarkable difference between hybrid and block elements consists in that, in a system of block elements, deformability is essentially associated with the element interfaces, not necessarily with the elements. In case the elements are stiff, $\boldsymbol{H}^{-1}$ vanishes, as well as $\boldsymbol{He}^{-1}$.

On the other hand, the term $\sum T^{*}e$ of the total complementary energy cancels, and the total complementary energy is reduced to

$$T^{*} = \tfrac{1}{2} \cdot \sum_{ef} \boldsymbol{ref}^{\mathrm{T}} \boldsymbol{Hef}^{-1} \boldsymbol{ref} \quad , \tag{15}$$

where the summation extends over the whole set of interfaces between pairs e and f of adjacent elements, $\boldsymbol{ref}$ and $\boldsymbol{ref}$ being the generalized contact force vectors, and $\boldsymbol{Hef}$ a symmetric, positive definite matrix of magnitudes depending on the constants which characterize the mechanical properties of such interfaces.

For sake of equilibrium,

$$\boldsymbol{ref} = - \boldsymbol{rfe} \quad . \tag{16}$$

Let $\boldsymbol{uef}$ denote the corresponding generalized displacement vector at the interface of blocks e and f. Conformity not being supposed to be achieved,

$$\boldsymbol{uef} \neq \boldsymbol{ufe} \quad . \tag{17}$$

As the blocks are assumed to be rigid, the displacement field within a given block e can be completely defined by ascribing to an arbitrary nodal point Ne, selected within the block, the vector $\boldsymbol{Ue}$ of the generalized displacements (translations and rotations around Ne) of the block, so that the linear relation between $\boldsymbol{uef}$ and $\boldsymbol{Ue}$ may be expressed by

$$\boldsymbol{uef} = \boldsymbol{Aef}\, \boldsymbol{Ue} \quad . \tag{18}$$

Each block e being in equilibrium under forces $\boldsymbol{Fe}$ and $\boldsymbol{ref}$, the equilibrium equation for the block e assumes the linear form

$$\sum_{f} \boldsymbol{Aef}^{\mathrm{T}} \boldsymbol{ref} = \boldsymbol{Fe} \quad , \tag{19}$$

in which the summation in f extends over the whole set of blocks contacting with e, but the repeated indices summation convention is not used.

Combining (22), (24) and (25), the following equation results

$$\sum_{f} \boldsymbol{Aef}^{\mathrm{T}} \boldsymbol{Hef} (\boldsymbol{Aef}\, \boldsymbol{Ue} - \boldsymbol{Afe}\, \boldsymbol{Uf}) = \boldsymbol{Fe} \quad , \tag{20}$$

in which the repeated indices summation convention is not used either. The system of equations obtained by putting together the equations (26) for the whole set of blocks provides magnitudes $\boldsymbol{Ue}$. As it may easily be checked, the matrix is symmetric (Arantes e Oliveira and al., 1992).

4 From integral methods to methods avoiding passages to the limit

An integral method for the solution of plane elasticity problems was presented by the author as a thesis in 1965 (Arantes e Oliveira, 1965) and published internationally in 1968 (Arantes e Oliveira,1968), which could be used with any kind of boundary conditions, and for simply or multiply-connected domains. Before the author proposed his method, a major disadvantage of integral methods lied in their insufficient generality. Most, like Shermann-Lauricella's (Muskhelishvili, 1953) and Massonet's (Massonet, 1960)were based indeed on the distorsional field generated by a concentrated force applied at the boundary of a half-plane (*simple radial distribution*), which prevented their use in the case of multiply-connected domains. Difficulties also arose in connection with mixed boundary conditions.

It is well-known that, if simple radial distributions are used as a Green's function, the problem of the solution of a two-dimensional elasticity problem can be reduced to the solution of a system of Fredhom's integral equations of the second kind, whose theory is well-known. The author was bold enough to use as a Green's function another kind of singular field: the one generated by a concentrated force applied at a given point of an indefinite plan, which leads to a system of Fredholm's equation of the first kind. This has the advantage of making it possible the solution of problems associated with multiply-connected domains. It is important to stress, that Fredholm's integral equations of the first kind may also have solutions. The problem is that a mathematical theory was not available for them. For this reason, their use had been avoided. Avoiding the use of Fredholm's equations of the first kind is another example of how the lack of a sound theoretical base may affect the progress of computational practice.

Our method had another feature, however, which contributed to make it different from any of those which had been presented sofar, and from most of those which would later be presented: no attempt was made to formulate the problem in terms of integral equations. This represented a significant feature of an engineering approach to computational methods. For a mathematician, reducing the problem to the solution of a system of integral equations would be vital, since only then a reliable mathematical theory could be applicable.

In the author's method, the domain was supposed to be immersed in the indefinite plane, or more precisely, a simple- or multiply-connected contour forming the boundary of the domain was supposed to be drawn on the indefinite plane. Fictitious forces were supposed to be distributed, not on the actual contour, but on an auxiliary contour involving the actual domain from the outside, such that the two following conditions were met: (i) the homogeneous field equations were satisfied at all points of the domain; (ii) the prescribed boundary conditions were satisfied at all points on the boundary. Both the auxiliary contour location and the fictitious force distribution were defined by a finite number of parameters. This was done by assimilating the auxiliary contour to a N-vertices polygon, and by applying on such polygonal contour fictitious forces linearly distributed over each of its N sides, and defined by its distribution density at each vertex.

Integral equations could be obtained if: (i) the distances between vertices of the auxiliary boundary would become indefinitely small, and; (ii) the vertices of the

auxiliary contour would approach the actual boundary until coinciding with the latter. However, such passages to the limit were not even deemed necessary. In other words, approximate solutions of the equations could be obtained without the equations even being written.

References

1. Arantes e Oliveira, E. R., "A utilização de métodos numéricos integrais na resolução de problemas de elasticidade plana" (thesis), LNEC, Lisbon, 1965.
2. Arantes e Oliveira, E. R., Plane stress analysis by a general integral method*, J. Eng. Mech. Div., Proc. ASCE*, 1968, **94**, No. EM1, Proc. Paper 5791, 79-101.
3. Arantes e Oliveira, E. R., Finite elements and finite differences: distinctions and similarities, *Int. Conf. on Math. in Finite Elements and Applications*, Brunel, 1981.
4. Arantes e Oliveira, E. R., On the accuracy of finite difference solutions to elliptic partial differential equations, *Int. J. Num. Meth. Eng.*, **21**, 229-238, 1985.
5. Arantes e Oliveira, E. R., Pedro, J. O.. Pina, C., Lemos, J. V., Non-linear and time-dependent analysis in three-dimensional structures, *in Zhang Yajun (ed.), Proc. of the Int. Conf. on Education, Practice and Promotion of Comp. Methods in Eng. Using Small Computers - EPMESC IV*, Dalian University Technology Press, 1992.
6. Arantes e Oliveira, E. R., Sobre a teoria dos métodos sem malha aplicados à resolução de equações elípticas, *VI Cong, Nac. de Mec. Apl. e Comp*., Aveiro, 2000.
7. Belytschko, T., Krongauz, Y., Organ, D., Fleming, M. Krysl, P., Meshless methods: an overview and recent developments, *Comp. Methods Appl. Mech. Engrg,* **139**, 3-47, 1996.
8. Cea, J., Approximation variationelle des problèmes aux limites, *Ann. Inst. Fourier*, Grenoble, **14** (2), 345-444, 1964.
9. Cundall, P. A., A computer model for simulating progressive large scale movements in blocky rock systems, *Symposium on Rock Fracture*, Nancy, 1971.
10. Kawai, T, New element methods in discrete structural analysis, *Japan Soc. Naval Arch*., **141**, 174-180, 1971.
11. Pian, T. H., Derivation of element stiffness matrices by assumed stress distributions, *AIAA Journal*, **7**, 1964.
12. Raviart, P. A., Sur l'approximation de certaines équations d'évolution linéaires et non-linéaires, *J. Math. Pures et Appl*., **46**, 11-107, 1967.

Acknowledgement
The research reported in the present paper was carried out under the Line 2 of the Institute of Structural Engineering, Territory and Construction – IST/ICIST (Technical University of Lisbon).