Historical Comments on Finite Elements

J. Tinsley Oden

1. Introduction

Finite elements: Perhaps no other family of approximation methods has had a greater impact on the theory and practice of numerical methods during the twentieth century. Finite element methods have now been used in virtually every area of engineering that can make use of models of nature characterized by partial differential equations.

Why have finite element methods been so popular in both the engineering and mathematical community? I believe that a principal reason for the success and popularity of these methods is that they are based on the weak, variational formulation of boundary and initial value problems. This is a critical property, not only because it provides a proper setting for the existence of very irregular solutions to differential equations (such as distributions), but also because the solution appears in the integral of a quantity over a domain. The simple fact that the integral of a measurable function over an arbitrary domain can be broken up into the sum of integrals over an arbitrary collection of almost disjoint subdomains whose union is the original domain is a vital property. Because of it, the analysis of a problem can literally be made locally, over a typical subdomain. Also, by making the subdomain sufficiently small, one can argue that polynomial functions of various degrees are adequate for representing the local behavior of the solution. This summability of integrals is exploited in every finite element program. It allows the analysts to focus their attention on a typical finite element domain and to develop an approximation independent of the ultimate location of that element in the final mesh.

The simple integral property also has important implications in physics and in most problems in continuum mechanics. Indeed, the classical balance laws of mechanics are global in the sense that they are integral laws applying to a given mass of material, a fluid, or solid. The only mathematical requirement is that the primitive variables be sufficiently regular for these integrals to be well-defined. Moreover, since these laws are supposed to be fundamental axioms of physics, they must hold over every finite portion of the material: every finite element of the continuum. Thus, once again, one is encouraged to think of approximate methods defined by integral formulations over typical pieces of the continuum under consideration.
2. The Origin of Finite Elements

When did finite elements begin? It is difficult to trace the origins of finite element methods because of a basic problem in defining precisely what constitutes a “finite element method.” To most mathematicians, it is a method of piecewise polynomial approximation, and its origins, therefore, are frequently traced to the appendix of a paper by Courant [24] that discusses piecewise linear approximations of the Dirichlet problem over a network of triangles. Also, the “interpretation of finite differences” by Polya [58] is regarded as embodying piecewise polynomial approximation aspects of finite elements.

On the other hand, the approximation of variational problems on a mesh of triangles goes back as many as 92 years. In 1851, Schellbach [64] proposed a finite-element–like solution to Plateau’s problem of determining the surface $S$ of minimum area enclosed by a given closed curve. Schellbach used an approximation $S_h$ of $S$ by a mesh of triangles over which the surface was represented by piecewise linear functions, and he then obtained an approximation of the solution to Plateau’s problem by minimizing $S_h$ with respect to the coordinates of hexagons formed by six elements [74]. This is not quite the conventional finite-element approach, but certainly as much a finite-element technique as that of Courant.

Some say that there is even an earlier work that uses some of the ideas underlying finite-element methods: Leibniz himself employed a piecewise linear approximation of the Brachistochrone problem proposed by Bernoulli in 1696 [39]. With the help of his newly developed calculus tools, Leibniz derived the governing differential equation for the problem, the solution of which is a cycloid. However, most would agree that to credit this work as a finite element approximation is stretching the point. Leibniz had no intention of approximating a differential equation; rather, his purpose was to derive one. Two and a half centuries later it was realized that useful approximations of differential equations could be determined by keeping the elements finite in size, not necessarily by taking infinitesimal elements, as in the calculus. This idea is, in fact, the basis of the term “finite element.”

There is also some difference in the process of laying a mesh of triangles over a domain, on the one hand, and generating the domain of approximation by piecing together triangles, on the other. Although these processes may look the same in some cases, they may differ dramatically in how the boundary conditions are imposed. Thus, neither Schellbach nor Courant—nor for that matter Synge, who used triangular meshes many years later—was particularly careful as to how boundary conditions were to be imposed or as to how the boundary of the domain was to be modeled by elements, issues that are now recognized as important features of finite-element methodologies. If
Methods

A finite-element method is one in which a global approximation of a partial differential equation is built up from a sequence of local approximations over subdomains, then credit must go back to the early papers of Hrennikoff [32], and perhaps beyond. Hrennikoff chose to solve plane elasticity problems by breaking up the domain of the displacements into little finite pieces, over which the stiffnesses were approximated using bars, beams, and spring elements. McHenry used a similar "lattice analogy" [42]. While these works are draped in the most primitive physical terms, it is nevertheless clear that the methods involve some sort of crude, piecewise linear or piecewise cubic approximation over rectangular cells. Miraculously, the methods also seem to be convergent.

To the average practitioner, finite elements are much more than a method of piecewise polynomial approximation. Partitioning a domain, assembling elements, and applying loads and boundary conditions, and, of course, local polynomial approximation, are all components of the finite-element method.

If this is so, then one must acknowledge the early papers of Gabriel Kron. Kron developed his "tensor analysis of networks" in 1939 and applied his "method of tearing" and "network analysis" to the generation of global systems from large numbers of individual components in the 1940s and 1950s [37,38]. Of course, Kron never necessarily regarded his method as one of approximating partial differential equations; rather, the properties of each component were regarded as exactly specified, and the issue was an algebraic one of connecting them all appropriately together.

In the early 1950s, Argyris [1,2] began to put these ideas together into what some call a primitive finite-element method: He extended and generalized the combinatoric method of Kron and other ideas that were being developed in the literature on system theory at the time and added variational methods of approximation to them. This was a fundamental step toward true finite-element methodology.

Around the same time, Synge [69] described his "method of the hypercircle," in which he also spoke of piecewise linear approximations on triangular meshes, but not in a rich variational setting and not in a way in which approximations were built by either partitioning a domain into triangles or assembling triangles to approximate a domain. (Indeed, Synge's treatment of boundary conditions was clearly not in the spirit of finite elements, even though he was keenly aware of the importance of convergence criteria and of the "angle condition" for triangles, later studied in some depth by others.)

It must be noted that during the mid-1950s, there were a number of independent studies underway that made use of "matrix methods" for the analysis of aircraft structures. A principal contributor to this methodology was Levy [40], who introduced the "direct stiffness
method” wherein he approximated the structural behavior of aircraft wings using assemblies of box beams, torsion boxes, rods, and shear panels. These assuredly represent some sort of crude local polynomial approximation in the same spirit as the Hrennikoff and McHenry approaches. The direct stiffness method of Levy had a great impact on the structural analysis of aircraft, and aircraft companies throughout the United States began to adopt and apply some variant of this method or of the methods of Argyris to complex aircraft structural analyses. During this same period, similar structural analysis methods were being developed and used in Europe, particularly in England. One must mention in this regard the work of Taig [70], in which shear lag in aircraft wing panels was approximated using basically a bilinear finite-element method of approximation. Similar element-like approximations were used in many aircraft industries as components in various matrix methods of structural analyses. Thus, the precedent for piecewise approximations of some kind was established by the mid-1950s.

To a large segment of the engineering community, the work representing the beginning of finite elements was the paper by Turner and colleagues [71], in which a genuine attempt was made at both a local approximation (of the partial differential equations of linear elasticity) and the use of assembly strategies essential to finite-element methodology. It is interesting that in this paper, local element properties were derived without the use of variational principles. It was not until 1960 that Clough [23] actually dubbed these techniques as “finite-element methods” in a landmark paper on the analysis of linear plane elasticity problems.

The 1960s were the formative years of finite-element methods. Once it was perceived by the engineering community that useful finite-element methods could be derived from variational principles, variationally based methods significantly dominated all the literature for almost a decade. If an operator was unsymmetric, it was thought that the solution of the associated problem was beyond the scope of finite elements because it did not lend itself to a traditional extremum variational approximation in the spirit of Rayleigh and Ritz.

From 1960 to 1965, a variety of finite-element methods were proposed. Many were primitive and unorthodox; some were innovative and successful. During this time, numerous attempts at solving the biharmonic equation for plate-bending problems were proposed that employed piecewise polynomial approximations but did not provide the essentials for convergence. This led to the concern of some as to whether the method was indeed applicable to such problems. On the other hand, it was clear that classical Fourier series solutions of plate problems were, under appropriate conditions, convergent and could be fit together in an assemblage of rectangular components [46]. Thus,
Methods

a form of “spectral finite element methods” was introduced early in the study of such problems. However, such high-order schemes never received serious attention in this period, as it was thought that piecewise polynomial approximations could be developed that did give satisfactory results. It was not until the mid-to late 1960s that papers on bicubic spline approximations [13,14] provided successful polynomial finite-element approximations for these classes of problems.

Many workers in the field believe that the famous Dayton conferences on finite elements (at the Air Force Flight Dynamics Laboratory in Dayton, Ohio) represented landmarks in the development of the field [59]. Held in 1965, 1968, and 1970, these meetings brought specialists from all over the world to discuss their latest triumphs and failures. The pages of the proceedings, particularly the earlier volumes, were filled with remarkable and innovative accomplishments from a technical community just beginning to learn the richness and power of this new collection of ideas. In these volumes, one can find many of the premier papers of now well-known methods. In the first volume alone one can find mixed finite element methods [31], Hermite approximations [56], C1-bicubic approximations [14], hybrid methods [57], methods for highly nonlinear elliptic systems [3], and other contributions. In later volumes, further assaults on nonlinear problems and special element formulations can be found.

In the late 1960s and early 1970s, there finally emerged the realization that the method could be applied to unsymmetric operators without difficulty, and thus, problems in fluid mechanics were brought within the realm of application of finite-element methods. In particular, finite element models of the full Navier-Stokes equations were first presented during this period [47,48,52].

The early textbook by Zienkiewicz and Cheung [75] did much to popularize the method with the practicing engineering community. However, the most important factor leading to the rise in popularity during the late 1960s and early 1970s was not purely the publication of special formulations and algorithms, but the fact that the method was being successfully used to solve difficult engineering problems. Much of the technology used during this period was the work of Bruce Irons, who with his colleagues and students developed a multitude of techniques for the successful implementation of finite elements. These included the frontal solution technique [34], the patch test [35], isoparametric elements [28], and numerical integration schemes [33]. The scope of finite element applications in the 1970s would have been significantly diminished without these contributions.

3. The Mathematical Theory

The mathematical theory of finite elements was slow to emerge from this caldron of activity. The beginning works on this theory were
understandably concerned with one-dimensional elliptic problems and used many of the tools and jargon of Ritz methods, interpolation, and variational differences. Early works in this line include Varga's study of “Hermite interpolation-type Ritz methods” for two-point boundary value problems [72] and the contribution of Birkhoff, de Boor, Schwartz, and Wendroff on “Rayleigh-Ritz Approximation by Piecewise Cubic Polynomials” [12]. The latter is certainly one of the first papers to deal with the issue of convergence of finite-element methods, although some papers on variational differences yielded similar results though they did not focus on the piecewise polynomial features of finite elements. The work of Feng Kang [30], published in Chinese (a copy of which I have not been able to acquire for review), may fall into this category and is sometimes noted as relevant to the convergence of finite-element methods.

The mathematical theory of finite elements for two-dimensional and higher-dimensional problems began in 1968, and several papers were published that year on the subject. One of the first papers in this period to address the problem of convergence of a finite method in a rigorous way and in which a priori error estimates for bilinear approximations of a problem in a plane elasticity are obtained is the often-overlooked paper of Johnson and McClay [36]. This paper correctly developed error estimates in energy norms, and it even attempted to characterize the deterioration of convergence rates owing to corner singularities. In the same year there appeared the first of two papers by Ogenesjan and Ruchovec [53,54], in which “variational difference schemes” were proposed for linear second-order elliptic problems in two-dimensional domains. These works dealt with the estimates of the rate of convergence of variational difference schemes.

Also in 1968, there appeared the paper of Zlamal [76] discussing in detail the interpolation properties of a class of triangular elements and their application to second-order and fourth-order linear elliptic boundary-value problems. This paper attracted the interest of a large segment of the numerical analysis community, and several very good mathematicians began to work on finite-element methodologies. Zlamal's paper represented a departure from studies of tensor products of polynomials on rectangular domains and provided an approach toward approximation in general polygonal domains. In the same year, Ciarlet [18] published a rigorous proof of convergence of piecewise linear finite-element approximation of a class of linear two-point boundary-value problems and proved $L^2$ estimates using a discrete maximum principle. Oliveira's work [55] on convergence finite-element methods established correct rates-of-convergence of certain problems in appropriate energy norms. In subsequent years, Schultz presented error estimates for “Rayleigh-Ritz-Galerkin methods” for multidimensional problems [65] and $L^2$-error bounds for these methods [66].
Methods

By 1972, finite-element methods had emerged as an important new area of numerical analysis in applied mathematics. Mathematical conferences were held on the subject on a regular basis, and there began to emerge a rich volume of literature on mathematical aspects of the method applied to elliptic problems, eigenvalue problems, and parabolic problems. A conference of special significance in this period was held at the University of Maryland in 1972. It featured a penetrating series of lectures by Ivo Babuška [8] and several important mathematical papers by leading specialists in the mathematics of finite elements, all collected in the volume edited by Aziz [5].

One unfamiliar with aspects of the history of finite elements may be led to the erroneous conclusion that the method of finite elements emerged from the growing wealth of information on partial differential equations, weak solutions of boundary-value problems, Sobolev spaces, and the associated approximation theory for elliptic variational boundary-value problems. This is a natural mistake, because the seeds for the modern theory of partial differential equations were sown about the same time as those for the development of modern finite-element methods, but in an entirely different garden.

In the late 1940s, Laurent Schwartz was putting together his theory of distributions, around a decade after the notion of generalized functions and their use in partial differential equations appeared in the pioneering work of Sobolev. Many other names could be added to the list of contributors to the modern theory of partial differential equations, but that is not our purpose here. Rather, we must note only that the rich mathematical theory of partial differential equations that began in the 1940s and 1950s, blossomed in the 1960s, and is now an integral part of the foundations of not only partial differential equations but also approximation theory, grew independently and parallel to the development of finite-element methods for almost two decades. There was important work during this period on the foundations of variational methods of approximation, typified by the early work of Lions [41] and by the French school in the early 1960s; but, although this work did concern itself with the systematic development of mathematical results that would ultimately prove to be vital to the development of finite-element methods, it did not focus on the specific aspects of existing and already successful finite-element concepts. It was, perhaps, an unavoidable occurrence that in the late 1960s these two independent subjects, finite-element methodology and the theory of approximation of partial differential equations via functional analysis methods, united in an inseparable way. So firmly were they united that it is difficult to appreciate the fact that they were ever separate.

The 1970s must mark the decade of the mathematics of finite elements. During this period, great strides were made in determining a priori error estimates for a variety of finite element methods, for
linear elliptic boundary-value problems, for eigenvalue problems, and for certain classes of linear and nonlinear parabolic problems; also, some preliminary work on finite-element applications to hyperbolic equations was done. It is both inappropriate and perhaps impossible to provide an adequate survey of this large volume of literature, but it is possible to present a reference, albeit biased, to some of the major works along the way.

An important component in the theory of finite elements is an interpolation theory: How well can a given finite-element method approximate functions of a given class locally over a typical finite element? A great deal was known about this subject from the literature on approximation theory and spline analysis, but its particularization to finite elements involves technical difficulties. One can find results on finite-element interpolation in a number of early papers [6,7,8,11,15,65,76]. But the elegant work on Lagrange and Hermite interpolations of finite elements by Claper and Raviart [20] must stand as an important contribution to this aspect of finite-element theory. The 1972 memoir of Babuška and Aziz [8] on the mathematical foundations of finite-element methods interweaves the theory of Sobolev spaces and elliptic problems with general results on approximation theory that have a direct bearing on finite-element methods.

It was known that Cea’s lemma [17] established that the approximation error in a Galerkin approximation of a variational boundary-value problem is bounded by the so-called interpolation error; that is, by the distance in an appropriate energy norm from the solution of the problem to the subspace of approximations. Indeed, it was this fact that made the results on interpolation theory using piecewise polynomials particularly interesting to finite-element methods. The introduction of the “Inf-Sup” condition by Babuška [6] and Babuška and Aziz [8] dramatically enlarged this framework. The condition is encountered in the characterization of coerciveness of bilinear forms occurring in elliptic boundary-value problems. The characterization of this “Inf-Sup” condition for the discrete finite-element approximation embodies the essential elements for studying the stability in convergence of finite-element methods. Brezzi [16] developed an equivalent condition for studying constrained elliptic problems, and these conditions provide for a unified approach to the study of qualitative properties, including rates of convergence, of broad classes of finite-element methods.

The work of Nitsche [44] on $L^2$ estimates for general classes of linear elliptic problems must stand out as one of the most important contributions of the 1970s. Strang [68] pointed out “variational crimes” inherent in many finite-element methods, such as improper numerical quadrature, the use of nonconforming elements, and improper satisfaction of boundary conditions. All of these are common prac-
Method

tices in applications, but all frequently lead to acceptable numerical schemes. In the same year, Ciarlet and Raviart [21,22] also studied these issues. Many of the advances of the 1970s drew upon earlier results on variational methods of approximation based on the Ritz method and finite differences; for example, the Aubin-Nitsche method for lifting the order of convergence to lower Sobolev norms [4,43,54] used such results. In 1974, Brezzi [16] used such earlier results on saddle-point problems and laid the groundwork for a multitude of papers on problems with constraints and on the stability of various finite-element procedures. Although convergence of special types of finite-element strategies such as mixed methods and hybrid methods had been attempted in the early 1970s [49], the Brezzi results and the methods of Babuška for constrained problems provided a general framework for studying virtually all mixed and hybrid finite elements [9,60,61].

The first textbook on mathematical properties of finite-element methods was the popular book of Strang and Fix [68]. An introduction to the mathematical theory of finite elements was published soon after by Oden and Reddy [51], and the treatise on the finite-element method for elliptic problems by Ciarlet [19] appeared two years later.

The work of Nitsche and Schatz [45] on interior estimates and of Schatz and Wahlbin [62,63] on $L^2$ estimates and singular problems represented notable contributions to the growing mathematical theory of finite elements. Progress was also made on finite-element methods for parabolic problems and hyperbolic problems [25,26,27] and on the use of elliptic projections for deriving error bounds for time-dependent problems [73].

The 1970s was also a decade in which the generality of finite-element methods began to be appreciated over a large portion of the mathematics and scientific community, and it was during this period that significant applications to highly nonlinear problems were made. The fact that very general nonlinear phenomena in continuum mechanics, including problems of finite deformation of solids and of flow of viscous fluids, could be modeled by finite elements and solved on existing computers was demonstrated in the early seventies [49]. By the end of that decade, several "general purpose" finite-element programs were being used by engineers to treat broad classes of nonlinear problems in solid mechanics and heat transfer. The mathematical theory for nonlinear problems also was advanced in this period, in particular by Falk's work [29] on finite-element approximations of variational inequalities.

It is not too inaccurate to say that by 1980, a solid foundation for the mathematical theory of finite elements for linear problems had been established and that significant advances in both theory and application into nonlinear problems existed. The open questions
that remain are difficult ones, and solving them will require a good understanding of the mathematical properties of the method.

4. Personal Reflections

The organizers of the meeting for which this discourse was prepared asked the invited authors to include "personal reflections" if possible. With such an invitation to relax customary standards of modesty and humility, I offer these closing comments on my own early introductions to the subject.

I remember very well my own introduction to finite elements. I had read thoroughly the work of Agyris and others on "matrix methods in structural mechanics" and had developed notes on the subject while teaching graduate courses in solid mechanics in the early 1960s, but none of the literature of the day had much impact on my university research at the time, or seemingly on the research of anyone in the university community. The aircraft industry was actively developing the subject during this period and was far ahead of universities in studying and implementing these methods.

Then, in 1963, I had the good fortune to enter the aerospace industry for a brief period and to meet and begin joint work with Gilbert Best, who had been charged with the responsibility of developing a large general-purpose finite-element code for use in aircraft structural analysis. Only the two of us worked on the project, but by fall 1963 we had produced some quite general results and one of the early working codes on finite elements. This code had features in it that were not fully duplicated for more than a decade. I still have copies of our elaborate report on that work [10].

It was Best who demonstrated to me the strength and versatility of the method. In our work we developed mixed methods, assumed-stress methods, and hybrid methods; explored algorithms for optimization problems, nonlinear problems, bifurcation, and vibration problems; and did detailed tests on stability and convergence of various methods by numerical experimentation. We developed finite elements for beams, plates, and shells; for composite materials; for three-dimensional problems in elasticity; for thermal analysis; and for linear dynamic analysis. Some of our methods were failures; most were effective and useful. Since convergence properties and criteria were not to come for another decade, our only way to test many of the more complex algorithms was to code them and compute solutions for test problems.

I returned to academia in 1964, and one of my first chores was to develop a graduate course on finite-element methods. At the same time, I taught mathematics and continuum mechanics, and it became clear to me that finite elements and digital computing offered hope
Methods

of transforming nonlinear continuum mechanics from a qualitative and academic subject into something useful in modern scientific computing and engineering. Toward this end, I began work with graduate students in 1965 that led to successful numerical analyses of problems in finite-strain elasticity (1965, 1966), elastoplasticity (1967), thermoelasticity (1967), thermostress (1969), and incompressible and compressible viscous fluid flow (1968, 1969). These works, many summarized in [49], include early (perhaps the first) uses of Discrete-Kirchhoff elements, incremental elastoplastic algorithms, conjugate-gradient methods for nonlinear finite-element systems, continuation methods, dynamic relaxation schemes, Taylor-Galerkin algorithms (then called “finite-element based Lax-Wendroff schemes”), primitive-variable formulations in incompressible flow, curvilinear elements, and penalty formulations; all these subjects have been resurrected in more recent times and have been studied in far more detail and better style and depth than was possible in the 1960s.

Although my later work, work in the 1970s and 1980s, was influenced by the competent mathematicians (and friends) who developed the subject during the period (Babuška, Ciarlet, Strang, Douglas, Nitsche, and many others), the work and guidance of G. Best was basic to my interest in this subject, and I dedicate this note to him.

5. Acknowledgments

Portions of this paper are excerpts from an article to appear in the Handbook of Numerical Analysis, edited by J. L. Lions and P. G. Ciarlet, North Holland Publishing Co., Amsterdam. I am grateful to North Holland for granting permission to use this material in the present volume.

References

Historical Comments on Finite Elements


Methods


Historical Comments on Finite Elements


Methods


