

https://dspace.org/

General multivariate approximation techniques applied to the finite element method

Hassoulas, Vasilios

2015-01-26
http://hdl.handle.net/10539/16739

Downloaded from a DSpace repository. https://dspace.org/

GENERAL MULTIVARIATE APPROXIMATION TECHNIQUES APPLIED TO THE FINITE ELEMENT METHOD

V. HASSOULAS

GENERAL MULTIVARIATE APPROXIMATION TECHNIQUES APPLIED TO THE FINITE ELEMENT METHOD

Vasilios Hassoulas

A Dissertation Submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg in fulfilment of the requirements for the degree of Master of Science.

Johannesburg 197.

GENERAL MULTIVARIATE APPROXIMATION TECHNIQUES APPLIED TO THE FINITE ELEMENT METHOD

V. Hassoulas

To my uncle, mother and brothers

To each of whom
I owe more
than I can possibly express.

DECLARATION

I, Vasilios Hassoulas, hereby declare that this dissertation is my own work and has not been submitted by me for a degree at any other University.

SUMMARY

It is by now well known that one of the most useful applications of the Approximation Theory is in approximating the solutions of boundary-value problems, where the recently developed Method of Finite Elements, in all of its different versions, has become one of the most popular and effective means for the discretization of the continuous problem. The purpose of this study is twofold. At first, an attempt has been made to present a unified treatment of the general multivariate approximation problem by bringing together the many published results concerning this problem. Then, the rate of convergence of the Finite Element Method is established which is found to depend upon the order to which the exact solution u can be approximated by the trial space S_h of piecewise polynomials.

For organizational purposes the content of the whole analysis has been divided into three Parts. The first Part is primarily concerned with an introduction to the Finite Element Method. The intention of this introduction, mainly one of expository nature, is to present the three major steps implied by the numerical solution of boundary-value problems through direct variational methods; viz: the variational formulation of the problem, the approximation of the solution of the

variational problem and finally, the numerical solution of the approximation problem.

Part II is exclusively devoted to the study of the multivariate approximation problem, where a distinction is made between three different approaches. In the first, Multivariate Pointwise Approximation of § 2.2, the multivariate Hermite approximation technique is analysed and the result of Theorem I provides us with a pointwise approximation estimate obtained through the standard Taylor series approach. The domain Ω of the problem, assumed to be a polyhedral type domain, is decomposed into a finite number of n-simplices and this permits us to consider the approximation problem over each n-simplex at a time. In the second approach, Multivariate Sobolev Approximation of § 2.3, which from the mathematical point of view is more general than the first, the approximation problem is posed over the Sobolev spaces $H_p^k(\Omega)$, $k \geq 1$, $1 \leq p < \infty$, and certain results from functional Analysis are employed in order to compute the mean-square approximation estimate of Theorem II. In particular, linear functionals which annihilate polynomials of a certain degree or less are of central importance in the study of this general approximation scheme. In the third approach to the approximation problem, Approximation by Convolution of § 2.4, the possibility of approximating a function u which does not meet the necessary continuity requirements for its interpolating polynomial to be defined is considered and the

outcome of this analysis is the result of Theorem III. Although Prof. G. Strang in a published paper on the Approximation of the Finite Element Method tackles the same problem by using Fourier transforms, I have used the general procedure of the previous paragraph 2.3 over Sobolev spaces in order to give, what I believe to be a more elegant mathematical treatment of this particular question of the approximation problem. Finally, for the sake of completeness only, an approximation technique over curved elements is briefly outlined in § 2.5, a situation which is encountered very frequently in the various practical applications of the method.

Then, the convergence of the Finite Element Method been governed by a single fundamental principle, viz: with respect to the energy inner-product $a(u,v)$ the approximate solution u_h is the projection of the exact solution u onto the subspace S_h, the main theme of the analysis of the third Part consists of an appropriate utilization of the approximation results of the second Part in order to compute the exact order of convergence of the finite element approximation u_h to the exact solution u. Finally, the main purpose of the two simple examples, which are given at the very end, is to illustrate to some extent the procedure which is usually followed in practice in order to determine several error bounds between the exact solution u and its approximation u_h.

ACKNOWLEDGEMENTS

I should like to take this opportunity to express my sincere appreciation to Mr. C. Wright of the Department of Applied Mathematics for the considerable assistance, guidance and encouragement which I have received throughout his supervision of this investigation. I should also like to thank Dr. R. Jones for reading the final notes and making a number of pertinent criticisms and suggestions.

TABLE OF CONTENTS

	Page
DECLARATION	i
SUMMARY	ii
ACKNOWLEDGEMENTS	v
TABLE OF CONTENTS	vi
PART I - FINITE ELEMENT SOLUTIONS OF BOUNDARY VALUE PROBLEMS	
1.1 INTRODUCTION	1
1.2 VARIATIONAL FORMULATION OF THE PROBLEM	5
1.2.1 The Space $L_2(\Omega)$	6
1.2.2 The Space $H^k(\Omega)$	8
1.2.3 Symmetric, Positive and Positive Definite Differential Operators	9
1.2.4 The Equivalent Variational Problem	11
1.2.5 An Example	14
1.2.6 The Finite Element Approximation $u_h(x)$	16
1.3 CONVERGENCE OF THE METHOD	22
1.3.1 The Minimum Principle	22
1.3.2 Consistency Plus Stability Implies Convergence	25
PART II - THE GENERAL MULTIVARIATE APPROXIMATION PROBLEM	
2.1 FINITE ELEMENTS AND APPROXIMATING SUBSPACES	29
2.1.1 The Simplicial Finite Element	32
2.1.2 The Hypercubic Finite Element	43
2.1.3 Finite Elements in the Plane	44
2.1.4 The Quadrilateral Element and the Isoparametric technique	50

2.2 MULTIVARIATE POINTWISE APPROXIMATION 52

2.2.1 Fréchet Differentiation 55

2.2.2 Multivariate Hermite Interpolation 62

2.2.3 Multivariate Lagrange Interpolation 83

2.3 MULTIVARIATE SOBOLEV APPROXIMATION 88

2.3.1 Sobolev Spaces 88

2.3.2 Sobolev Lemma 89

2.3.3 The Quotient Space 91

2.3.4 The Quotient Norm 93

2.3.5 Equivalence of the Norm 96

2.3.6 Bramble and Hilbert Lemma 99

2.3.7 Multivariate Sobolev Interpolation 101

2.4 APPROXIMATION BY CONVOLUTION 114

2.4.1 The Space $W_p^k(R^n)$ 115

2.4.2 The Convolution Operator 116

2.4.3 Application to the Approximation Problem 117

2.4.4 Some Useful Remarks 129

2.5 MULTIVARIATE APPROXIMATION OVER CURVED ELEMENTS 131

PART III - ERROR ESTIMATES FOR THE FINITE ELEMENT APPROXIMATION

3.1 ERROR ANALYSIS FOR CONFORMING FINITE ELEMENTS 142

3.1.1 A Model Problem 142

3.1.2 Admissible Triangulation of Ω 144

3.1.3 The Error in the Energy Norm 146

3.1.4 The Constant Strain Condition 149

3.1.5 The Error in the H^m-norm 150

3.1.6 The Nitsche Trick 150

3.1.7 The Error in Displacement 156

3.1.8 The Phenomenon of Superconvergence 158

3.2 ERROR ANALYSIS FOR NON-CONFORMING FINITE ELEMENTS 160

3.3 TWO ILLUSTRATIVE EXAMPLES 168

CONCLUSIONS 182

REFERENCES 184

PART I

FINITE ELEMENT SOLUTIONS OF LINEAR BOUNDARY VALUE PROBLEMS

1.1 INTRODUCTION

A widely used technique in approximating the solutions of boundary-value problems arising in the theory of partial differential equations is provided by the Finite Element Method which, from the mathematical point of view, falls into the framework of the classical Ritz-Galerkin technique. It operates on problems posed in the variational form rather than directly on the differential equation itself, which is the case with the well-known method of finite differences. In order to implement the Finite Element Method, the domain Ω is replaced by a finite number of subdomains $e_i \subset \Omega$ and certain families of functions are considered which have different analytical expressions within each subdomain e_i. Then, by the term *finite element* we understand a closed subdomain $\bar{e}_i$ and a family of functions are allowed to occur within it. This family is a linear combination, with coefficients $q_j^{e_i}$, of a finite number of basic functions so that each function of the family corresponds to ascribing particular values to the parameters $q_j^{e_i}$. The values of the components of the function and, possibly those of some of its partial derivatives at a certain number of points

placed on the boundary and, perhaps, in the interior of the finite elements, usually called *nodes* or *nodal points*, are chosen as parameters. Therefore, the *approximate* solution, which is computed separately within each element e_i, is assumed to be a linear combination:

$$\sum_j q_j^{e_i} \phi_j^{e_i}(x) \qquad \ldots \qquad (1\text{-}1)$$

of given basis functions $\phi_j^{e_i}(x)$ and the parameters $q_j^{e_i}$ in (1-1) are computed from the underlying variational principle. For example, it may be required that the exact solution $u(x)$ minimizes a given expression of potential energy corresponding to some physical system.

The crucial difference between the Finite Element Method and the classical Ritz-Galerkin technique lies in the construction of the basis functions. In the former they are piecewise polynomials and their main feature is that they vanish over all but a fixed number of the elements into which the given domain Ω has been divided. The following two disadvantages of the classical Ritz approach contribute to the success of Finite Element Method:

(a) In practice, the construction of the basis functions $\phi_j^{e_i}(x)$ is only possible for some special domains Ω, and

(b) even for these simple domains the method can be highly unstable.

Furthermore, the capability of dealing with arbitrary complex

geometrical r[illegible]ions as well as the banded nature of the resulting sy[illegible] simultaneous linear equations of the assembled p[illegible], are some more of the basic reasons which have turned the balance of choice in favour of the newly developed method of finite elements.

The true originators of the idea were the engineers, who regarded the method as a means of generating a discrete model of a physical system, and, for a number of complicated problems in elasticity and structural analysis the Finite Element Method superceded the then well-established method of finite differences. Only recently has the method become attractive to mathematicians and this is due to the latest developments in approximating theory and variational principles. Historically, however, the first idea goes back to Courant [9]. He suggested - in a lecture delivered some 35 years ago - a triangulation of the given domain and th[illegible]se of linear trial functions over each triangle in order to solve second order boundary-value problems. Prior to Courant, sine and cosine functions as well as Legendre polynomials were commonly used in approximation problems. On regular domains these functions are still entirely adequate, but for irregular regions the situation is completely different as these functions become almost useless. Although Courant's idea was forgotten, when finally it was recalled again - about 15 years after his remarkable lecture - and was combined with the newly developed theory on approximation of functions by piecewise polynomials (so far this new theory on approximation having been constructed

completely independent from what Courant had said in his lecture), was to be the most powerful technique for numerical solution of partial differential equations: The Finite Element Method.

In addition to the Ritz version of the Finite Element Method which requires the minimization of a certain functional, there are also some other forms of the method for problems where convenient variational principles are not available. The most popular of these involve classical methods such as Galerkin, Least squares, Collation, etc. However, regardless of the principle applied, we shall always use the term *Finite Element Method* if the basic functions constructed have the above mentioned characteristic property to being piecewise polynomials.

A fundamental mathematical problem is to determine how efficiently piecewise polynomials can approximate an unknown solution u. In other words, to obtain an estimate for the error as closely as possible and determine how rapidly the error decreases as the number of the finite *elements e_i is increased. Then, the accuracy of the* approximation can be increased by simply refining the subdivision of the domain. Therefore, the following two important questions need an answer as far as the approximation problem is concerned:

1. What is the degree of the approximation which can be achieved by the Finite Element Method?

2. What is the error estimate for the difference $u - u_h$, where u denotes the exact solution of the problem and u_h its finite element approximation?

The parameter h, of course, in some sense measures the size of the finite elements e_i so that we are working with a sequence of approximations with $h \to 0$. In recent years, much of the mathematical literature of the method has been concerned with forming a wider basis for the finite element approximation from the point of view of Functional Analysis and thus, without any doubt, the problem requires some sort of rigorous mathematical treatment in order to determine its order of convergence.

1.2 VARIATIONAL FORMULATION OF THE PROBLEM

In numerically solving a given partial differential equation, we first express approximately the solution in terms of a finite number of parameters and, since in general the solution is sought in some class of functions, it is essential that we are able to express any function of the class in terms of a finite number of parameters with a reasonable accuracy. Then, the given differential operator is transformed into expressions relating these parameters and, if the differential operator is linear, these relations are also linear and we are led to a linear system of algebraic equations. Therefore, for the numerical solution of any differential equation it becomes clear that the following two principles

are essential:

(i) The choice of the local parameters of the solution, and

(ii) The use of an appropriate variational principle for transforming the given differential equation into relations among the parameters of the solution.

But, in order to answer these two questions, we first need some sort of notation:

1.2.1 The space $L_2(\Omega)$:

Let Ω be a bounded and open domain in R^n with its boundary denoted by $\partial\Omega$, and let

$$C(\bar{\Omega}),$$

where $\bar{\Omega}$ is the closed domain resulting from the combination of Ω and $\partial\Omega$, be the space of all real-valued functions defined on Ω and which are continuous over $\bar{\Omega}$. Then we denote by:

$$L_2(\Omega) \quad \text{or by} \quad H_0(\Omega)$$

the space of all measurable functions $u(x)$ defined on Ω for which:

$$\int_{\Omega} [u(x)]^2 dx < \infty$$

with the integration being a Lebesgue integration. That is,

$$L_2(\Omega) = \{u(x) : \int_{\Omega} [u(x)]^2 dx < \infty\} \qquad \ldots (1.2)$$

and the real-valued function $\|.\|_{L_2(\Omega)}$ defined by:

$$\|u\|_{L_2(\Omega)} = \left(\int_\Omega [u(x)]^2 dx\right)^{\frac{1}{2}} \qquad \ldots \quad (1\text{-}3)$$

is known as the L_2-norm of the function $u(x) \in L_2(\Omega)$. One basic reason for the introduction of such a space of functions here is that the space $L_2(\Omega)$ is complete in the norm (1-3), whereas, the continuous space $C(\bar{\Omega})$ is not. This means that, if $\{u_n(x)\}$ is a sequence of functions in $L_2(\Omega)$ for which:

$$\|u_n(x) - u_m(x)\|_{L_2(\Omega)} \to 0, \text{ as } n,m \to \infty$$

there exists a function $u(x)$ in $L_2(\Omega)$ such that:

$$\|u_n(x) - u(x)\|_{L_2(\Omega)} \to 0, \text{ as } n \to \infty.$$

Note that $C(\bar{\Omega})$ is a linear subspace of the space $L_2(\Omega)$ and another equivalent *definition* for the space $L_2(\Omega)$ is that it is exactly the completion of the space $C(\bar{\Omega})$ with respect to the norm (1-3). This property of completeness becomes quite important when we reach the stage of constructing approximate solutions of differential equations by employing certain variational principles, which is the case with the Finite Element Method. The space $L_2(\Omega)$ is a Hilbert space with an inner-product defined by:

$$(u,v)_{L_2(\Omega)} = \int_\Omega u(x).v(x)\,dx; \qquad u,v \in L_2(\Omega)$$

1.2.2 The space $H^k(\Omega)$

For any nonnegative integer k, let $C^k(\bar{\Omega})$ be the space of all the real-valued functions defined on Ω and which have all their derivatives of up to the order k-th inclusive continuous over $\bar{\Omega}$. Then, the space of functions defined by:

$$H^k(\Omega) = \{u(x) : u(x) \in L_2(\Omega);\ D^\alpha u(x) \in L_2(\Omega),\ \text{for all } |\alpha| \le k\} \quad \ldots (1\text{-}4)$$

is the completion of the space $C^k(\bar{\Omega})$ with respect to the following norm:

$$\|u\|^2_{H^k(\Omega)} = \sum_{|\alpha| \le k} \|D^\alpha u\|^2_{L_2(\Omega)} \quad \ldots (1\text{-}5)$$

ie, $H^k(\Omega)$ is the space of functions which together with all their generalized derivatives of order up to the k inclusive are square-integrable over the domain Ω, where:

$$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n),\ |\alpha| = \sum_{i=1}^{n} \alpha_i,\ D^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdot \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}},\ dx = dx_1 . dx_2 \ldots dx_n$$

we have used the usual multi-index notation. The space defined through (1-4) is a Hilbert space with an inner-product defined by:

$$(u,v)_{H^k(\Omega)} = \sum_{|\alpha| \le k} (D^\alpha u, D^\alpha v)_{L_2(\Omega)} = \sum_{|\alpha| \le k} \int_\Omega D^\alpha u D^\alpha v dx;\ u, v \in H^k(\Omega)$$

Corresponding to the norm defined by (1-5), we define a semi-norm by the formula:

$$|u|^2_{H^k(\Omega)} = \sum_{|\alpha|=k} \|D^\alpha u\|^2_{L_2(\Omega)} = \sum_{|\alpha|=k} \int_\Omega [D^\alpha u]^2 \, dx \qquad \ldots (1\text{-}6)$$

where, unlike the norm, the semi-norm is always zero for a function of degree less than k.

Finally, for a negative integer k, we can define, by using the principle of duality, the following negative norm:

$$\|u\|_{H^{-k}(\Omega)} = \max_{v \in H^k(\Omega)} \frac{|(u,v)_{L_2(\Omega)}|}{\|v\|_{H^k(\Omega)}} = \max_{v \in H^k(\Omega)} \frac{\left|\int_\Omega u(x) \cdot v(x) \, dx\right|}{\|v\|_{H^k(\Omega)}}$$

where, of course, by the word *negative* we mean that the index k is negative.

1.2.3 Symmetric, Positive and Positive Definite Differential Operators

Consider the following general linear boundary-value problem of order 2m in the n-variables $x = (x_1, x_2, \ldots x_n)$:

$$A.u(x) = f(x), \quad x \in \Omega \qquad \ldots (1\text{-}7)$$

subject to:

several essential and/or natural boundary conditions on $\partial\Omega$ $\ldots$ (1-7)'

For the 2m-th order differential operator A, defined through the equation (1-7), we say that it is:

1. symmetric, if the following condition is satisfied:

$$(Au,v)_{L_2(\Omega)} = (u,Av)_{L_2(\Omega)}$$

for any two functions u and v from its field of definition. Furthermore, a symmetric operator A is said to be:

2. positive, if:

$$(Au,u)_{L_2(\Omega)} \geq 0$$

for any function u from its field of definition, and where equality occurs if and only if $u = 0$,

3. positive definite or elliptic, if:

$$(Au,u)_{L_2(\Omega)} \geq \gamma \|u\|^2_{H^m(\Omega)} \qquad \ldots \quad (1\text{-}8)$$

for some positive constant γ and any function u from its field of definition.

Then, one can easily prove that if the operator A is positive, the equation (1-7) cannot have more than one solution: Indeed, suppose that it has two solutions u_1 and u_2 such that:

$$Au_1 = f \quad \text{and} \quad Au_2 = f$$

or, by subtracting and using the fact that the operator A is linear:

$$A(u_1-u_2) = 0$$

Multiplying this equation scalarly by $u_1 - u_2$, we obtain:

$$(A(u_1-u_2),u_1-u_2)_{L_2(\Omega)} = 0$$

which means, since the operator A is assumed to be positive, that:

$$u_1 - u_2 = 0, \text{ or } u_1 = u_2$$

1.2.4 The Equivalent Variational Problem

We introduce now an equivalent variational approach to the problem (1-7) - (1-7)' which is a matter of great importance in our entire analysis from the point of view that it provides a means for the discretization of the continuous equation (1-7), ie,
The functional:

$$F(v) = (Av,v)_{L_2(\Omega)} - 2(f,v)_{L_2(\Omega)} \quad \ldots \quad (1\text{-}9)$$

is related to the differential equation (1-7) in the following way: if the equation (1-7) has a solution u, this solution minimizes the functional (1-9) and, conversely, if there exists an element u which minimizes the functional (1-9) this same element u satisfies the equation (1-7).

This method of solving boundary-value problems by replacing the differential equation by the problem of minimizing a certain functional is usually called, the *energy method* and the particular functional F(u) the *functional of the energy method*. An equivalent form for the functional (1-9) is given through the following:

$$F(v) = a(v,v) - 2(f,v)_{L_2(\Omega)} \quad \ldots \quad (1\text{-}10)$$

where the second order term $a(v,v)$ in (1-10) is obtained as a result of an integration by parts of the inner-product term $(Av,v)_{L_2(\Omega)}$ in (1-9) over the given domain Ω and by invoking the boundary conditions (1-7)' which are supposed to be satisfied by the functions v. We shall call the quantity $a(v,v)$ the *strain energy* of the function v and its square root the *energy norm* - or *natural norm* - of the function v. It is this norm with respect to which the Ritz method is minimizing the functional $F(v)$ and any error estimate will at the very first be given in terms of that kind of norm.

As far as the differential equation (1-7) is concerned, it is easily seen that the operator A is acting on the space of functions which are 2m-times differentiable and satisfy all the boundary conditions imposed on $\partial\Omega$. That is, we say that the (unknown) exact solution u belongs to the space of functions $H_B^{2m}(\Omega)$, with the letter B referring to the boundary conditions. Then the natural question which arises is to find the admissible space of functions for which the functional (1-10), or equivalently (1-9), is well-defined so that the minimization process can proceed. It is an easy matter to see, however, that, since the minimization can proceed as long as the functional $F(v)$ remains finite and since the quadratic term $a(v,v)$ in (1-10) involves derivatives of up to the order m-th inclusive only, the functional (1-10) will be well-defined for all the functions v which have all their generalized derivatives of up to the order m-th inclusive square-integrable over Ω;

ie, for all the functions v which belong to the space $H^m(\Omega)$ and, further, they are required to satisfy only the essential boundary conditions of the problem. This last observation comes out as an immediate consequence of the following general rule (see G. Strang and G. Fix [24] p.8) which distinguishes between essential boundary conditions which remain and natural boundary conditions which go: 'Boundary conditions which involve only derivatives of order s will make sense in the H^s-norm; those involving derivatives of order s or higher will be unstable and will not apply to the functions belonging in the space H^s.' Therefore, the corresponding admissible space of functions for the equivalent variational principle, defined through the functional (1-10), will be the subspace:

$$V \subset H^m(\Omega)$$

which satisfies the essential (homogeneous or inhomogeneous) boundary conditions of the problem. Note that the only difference between the spaces V and $H^m(\Omega)$ is that the first contains all these functions:

$$v \in H^m(\Omega)$$

which satisfy the essential conditions of the problem. In the case where all the boundary conditions are natural, or of Neumann type, the two spaces V and $H^m(\Omega)$ coincide.

Then, the result of our whole analysis so far is that the same solution u can be approximated either from the differential equation (1-7) directly or from the corresponding variational principle (1-10). In the first, we replace

each derivative of the given differential equation by a suitable difference quotient and we are led to a system of difference equations from which an approximate solution is obtained. This is the finite difference scheme which operates directly on the differential equation. In the latter, however, we are looking for a function u which minimizes the quadratic functional (1-10) over the infinite-dimensional space V, which means that we now have the enormous advantage of trying functions which do not belong to the originally admissible space $H_B^{2m}(\Omega)$. Furthermore, the space V contains, by its construction, those functions v provided they can be obtained as the limit of a sequence v_n in $H_B^{2m}(\Omega)$, where by the word *limit* we mean that the second order term of the energy functional (1-10) converges; ie,

$$a(v-v_n, v-v_n) \to 0, \text{ as } n \to \infty.$$

However, it can easily be proved that the energy norm $a(u,u)$ is equivalent to the H^m-norm defined by (1-5), ie, there exist two positive constants C_1 and C_2 such that:

$$C_1 \|u\|^2_{H^m(\Omega)} \leq a(u,u) \leq C_2 \|u\|^2_{H^m(\Omega)} \quad \ldots \quad (1\text{-}11)$$

so that, the completion of the space $H_B^{2m}(\Omega)$ to that of V can be carried out in either norm. We note here, however, that such an enlargement of the space from $H_B^{2m}(\Omega)$ to V does not lower the minimum value of the functional (1-10), since every new value of F(u) is the limit of old values $F(u_n)$.

1.2.5 An Example

As example, consider the following second order partial

differential equation of elliptic type:

$$Au(x) = -\sum_{i,j=1}^{2} \frac{\partial}{\partial x_i}\left(q_{ij}(x)\frac{\partial u}{\partial x_j}\right) = f(x),\ x \in \Omega \quad \ldots \quad (1\text{-}12)$$

subject to the Dirichlet boundary conditions:

$$u(x) = 0,\quad x \in \partial\Omega \quad \ldots \quad (1\text{-}13)$$

where the functions $q_{ij}(x)$, $1 \le i,j \le 2$, are real-valued functions with:

$$q_{ij}(x) \in C^1(\bar{\Omega})$$

and are such that:

$$q_{ij}(x) = q_{ji}(x),\ 1 \le i,j \le 2,\quad x \in \Omega$$

and satisfy the following ellipticity condition:

$$\sum_{i,j=1}^{2} q_{ij}(x)\xi_i\xi_j \ge c_1|\xi|^2 \quad \ldots \quad (1\text{-}14)$$

for any $x \in \Omega$ and any real number $\xi \in R^2$ with $|\xi|^2 = \xi_1^2+\xi_2^2$. Then, a bilinear form $a(u,v)$ on $V \times V$, where here $V \subset H^1(\Omega)$, can be derived by applying a formal integration by parts on the inner-product term $(Au,v)_{L_2(\Omega)}$, ie,

$$a(u,v) = \int_\Omega \left[-\sum_{i,j=1}^{2} \frac{\partial}{\partial x_i}\left(q_{ij}(x)\frac{\partial u}{\partial x_j}\right)\right] v(x)\,dx = -\sum_{i,j=1}^{2} \int_\Omega \frac{\partial}{\partial x_i}\left(q_{ij}(x)\frac{\partial u}{\partial x_j}\right) v(x)\,dx =$$

$$= \sum_{i,j=1}^{2} \int_\Omega q_{ij}(x)\frac{\partial u}{\partial x_j}\frac{\partial v}{\partial x_i}\,dx - \sum_{i,j=1}^{2} \int_\Omega \frac{\partial}{\partial x_i}\left(q_{ij}(x)\frac{\partial u}{\partial x_j}\right) v(x)\,dx =$$

$$= \sum_{i,j=1}^{2} \int_\Omega q_{ij}(x)\frac{\partial u}{\partial x_j}\frac{\partial v}{\partial x_i}\,dx,$$

since the second term is always equal to zero because of the boundary condition (1-13). Thus,

$$a(u,v) = \sum_{i,j=1}^{2} \int_{\Omega} q_{ij}(x) \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} dx \qquad \ldots (1\text{-}15)$$

from which, together with the ellipticity condition (1-14), we immediately get:

$$|a(u,u)| \geq C_1 \sum_{i=1}^{2} \int_{\Omega} \left(\frac{\partial u}{\partial x_i}\right)^2 dx = C_1 \|u\|^2_{H^1(\Omega)} \qquad \ldots (1\text{-}16)$$

On the other hand, from (1-15), we have that:

$$|a(u,u)| \leq \max_{1\leq i,j\leq 2} \max_{x\in\Omega} |q_{ij}(x)| \sum_{i,j=1}^{2} \int_{\Omega} \frac{\partial u}{\partial x_j} \frac{\partial u}{\partial x_i} dx = C_2 \|u\|^2_{H^1(\Omega)} \qquad \ldots (1\text{-}17)$$

where the constant

$$C_2 = \max_{1\leq i,j\leq 2} \max_{x\in\Omega} |q_{ij}(x)| .$$

Finally, by combining the inequalities (1-16) and (1-17), we obtain:

$$C_1 \|u\|^2_{H^1(\Omega)} \leq |a(u,u)| \leq C_2 \|u\|^2_{H^1(\Omega)}$$

which is the double inequality (1-11)

1.2.6 The Finite Element Approximation $u_h(x)$

For the implementation of the Ritz technique in the Finite

Element Method, instead, the infinite-dimensional space V is replaced by a finite-dimensional subspace

$$S_h \subset V$$

or, more precisely, by a sequence of finite-dimensional subspaces $S_h, h \to 0$, spanned by basis functions which normally are piecewise polynomials. Then, by a choice of local parameters we understand that each trial function $v_h \in S_h$ is expressed in terms of its nodal parameters which are the unknowns q_j of the discrete final system. Further, each of these nodal parameters is nothing more than the value, at a given node, say z_j, of either the function v_h itself or one of its derivatives (eg see [24]), ie,

$$q_j = D_j v_h(z_j)$$

with the operator D_j being of zero order in the case where the parameter q_j is just the function value. To each of these nodal parameters q_j we associate a basis function $\phi_j(x)$ which is defined through the following condition:

$$D_j \phi_j(z_i) = \delta_{ij}$$

which simply states that at the node z_j the value of the function $D_j\phi_j(x)$ is one and zero at all the others. Thus, with respect to the nodal points z_j and the operators D_j, the basis functions $\phi_j(x)$ constitute an interpolating basis for the trial space S_h; ie, any function $v_h \in S_h$ can be expressed by a combination:

$$v_h(x) = \sum_j q_j \phi_j(x)$$

and, furthermore, they must be chosen to satisfy the following conditions:

1. They are piecewise continuous over the entire domain Ω.
2. They have compact support over Ω or, more precisely, their support is decreased as the number of the nodal parameters q_j is increased.
3. They satisfy the essential boundary conditions of the problem at hand, and
4. $\int_\Omega \phi_i(x)\,\phi_j(x)\,dx = 0$
   for most of the pairs (i,j).

Over the new subspace S_h now, the quadratic functional (1-10) is replaced by the following:

$$F(v_h) = a(v_h, v_h) - 2(f, v_h)_{L_2(\Omega)} \quad \ldots (1\text{-}18)$$

and if the domain Ω is decomposed into a number, say M, of finite elements e_i, such that

$$\Omega = \bigcup_{i=1}^{M} e_i$$

the space integral on the right-hand side of (1-18) can be split up into M parts:

$$F(v_h) = \sum_{i=1}^{M} [a(v_h, v_h)_{e_i} - 2(f, v_h)_{L_2(e_i)}] \quad \ldots (1\text{-}19)$$

and the minimization of the above functional over the space S_h will give us the optimal parameters Q_j for which the approximate solution

$$u_h = \sum_j Q_j \phi_j(x)$$

is obtained. In more detail, the basic computation in the evaluation of the energy functional (1-18) is carried out, at first, over each element e_i separately:

$$[F(u_h)]_{e_i} = [a(u_h, u_h)]_{e_i} - 2(f, u_h)_{L_2(e_i)} \quad . \ . \ . \ (1\text{-}20)$$

or, by substituting the expression of the trial function:

$$u_h = \sum_j q_j^{e_i} \phi_j^{e_i}(x)$$

over each element e_i at a time into (1-20), we get:

$$[F(u_h)]_{e_i} = a\left(\sum_j q_j^{e_i}\phi_j^{e_i}(x), \sum_j q_j^{e_i}\phi_j^{e_i}(x)\right) - 2\left(f, \sum_j q_j^{e_i}\phi_j^{e_i}(x)\right)_{L_2(e_i)} =$$

$$= (q^T)^{e_i} K^{e_i} q^{e_i} - 2(q^T)^{e_i} F^{e_i}$$

Thus, there are two main problems, first to compute the element stiffness matrices K^{e_i} and the element load vectors F^{e_i} within each element e_i and second to assemble them over the entire domain Ω:

$$F(u_h) = \sum_{i=1}^{M} [F(u_h)]_{e_i} = \sum_{i=1}^{M} (q^T)^{e_i} K^{e_i} q^{e_i} - 2\sum_{i=1}^{M} (q^T)^{e_i} F^{e_i} = q^T K q - 2q^T F$$

to obtain the functional $F(v_h)$ which is a function of the parameters $q = (q_1, q_2, \ldots, q_n)$, where n here denotes the number of unconstrained nodes of the partition of the domain Ω and, consequently, the dimension of the subspace S_h. Then, the first derivative of $F(v_h)$ with respect to each parameter q_j, $1 \leq j \leq n$, is put equal to zero leading to a system of n equations in the n unknowns q_j, $1 \leq j \leq n$:

$$KQ = F \qquad \ldots \quad (1\text{-}21)$$

The equations (1-21) are linear if the problem is linear, and non-linear if the problem is non-linear. Because each point of the subdivision is coupled only to its neighbouring points, the matrix K in (1-21) is a sparse and symmetric matrix with all its non-zero elements being banded around the main diagonal.

1.3 CONVERGENCE OF THE METHOD

The main problem, which from the mathematical point of view lies at the very centre of the theory of the Finite Element Method, is that of giving an estimate for the error between the exact solution u and its finite element approximation u_h, or, since the function u minimizes the quadratic functional (1-10) over the infinite-dimensional space V and the function u_h minimizes the quadratic functional (1-18) over the finite-dimensional subspace $S_h \subset V$, between the minimizing function over V and the minimizing function over S_h.

As a first step in this direction, we shall first prove that the approximate solution $u_h \in S_h$ satisfies the following equation:

$$a(u_h,v_h) = (f,v_h)_{L_2(\Omega)} \text{, for all the functions } v_h \in S_h \quad \ldots \quad (1\text{-}22)$$

or, by considering the whole of the space V instead of the subspace S_h, that:

$$a(u,v) = (f,v)_{L_2(\Omega)} \text{, for all the functions } v \in V \quad \ldots \quad (1\text{-}23)$$

We say that the equation (1-23) expresses the vanishing of the first variation of the functional $F(v)$ in any direction $v \in V$, while equation (1-22) expresses the vanishing of the first variation of $F(v_h)$ in the direction of the particular function $v_h \in S_h$.

Indeed, since the function u_h minimizes (1-18) over S_h, we can write:

$$F(u_h) \leq F(u_h+\varepsilon v_h) \text{, for all } v_h \in S_h \text{ and any scalar } \varepsilon \in R$$

Then, since:

$$F(u_h+\varepsilon v_h) = a(u_h+\varepsilon v_h, u_h+\varepsilon v_h) - 2(f, u_h+\varepsilon v_h)_{L_2(\Omega)} =$$

$$= a(u_h,u_h) + 2\varepsilon a(u_h,v_h) + \varepsilon^2 a(v_h,v_h) - 2(f,u_h)_{L_2(\Omega)} - 2\varepsilon(f,v_h)_{L_2(\Omega)} =$$

$$= F(u_h) + 2\varepsilon[a(u_h,v_h) - (f,v_h)_{L_2(\Omega)}] + \varepsilon^2 a(v_h,v_h)$$

and since the perturbed functional $F(u_h+\varepsilon v_h)$ is quadratic in ε and attains its minimum value for $\varepsilon = 0$, we have:

$$\left.\frac{dF(u_h+\varepsilon v_h)}{d\varepsilon}\right|_{\varepsilon=0} = 0$$

or

$$a(u_h,v_h) = (f,v_h)_{L_2(\Omega)}, \text{ for all } v_h \in S_h$$

which is exactly the equation (1-22). The second equation, (1-23) is nothing more than a natural generalization of the first.

1.3.1 The Minimum Principle

An important characterization of the approximate solution $u_h \in S_h$ usually known as the *minimum principle* and which constitutes the foundation of the entire convergence analysis is the following: Suppose that the function u minimizes the functional (1-10) over the full admissible space $V \subset H^m(\Omega)$ and let S_h be any finite-dimensional subspace of V. Then,

$$a(u-u_h,u-u_h) = \min_{v_h \in S_h} a(u-v_h,u-v_h) \qquad \ldots (1\text{-}24)$$

where u_h again denotes the finite element approximation to u.

Indeed, for any two functions $u,v \in V$, we have:

$$F(u) - F(v) = a(u,u) - 2(f,u)_{L_2(\Omega)} - a(v,v) + 2(f,v)_{L_2(\Omega)} =$$

$$= a(u,u) - a(v,v) + 2(f,v-u)_{L_2(\Omega)},$$

or, by using the equality (1-23), that:

$$F(u) - F(v) = a(u,u) - a(v,v) + 2a(v,v-u) = a(u,u) - a(v,v) + 2a(v,v) - 2a(v,u) =$$

$$= a(u,u) - 2a(v,u) + a(v,v).$$

Thus,

$$F(u) - F(v) = a(u-v,u-v), \text{ for any functions } u,v \in V \quad . \; . \; . \; (1\text{-}25)$$

from (1-25) now, and for $v \equiv v_h$ and $v \equiv u_h$, we get:

$$\min_{u_h \in S_h} a(u-u_h,u-u_h) \leq a(u-u_h,u-u_h) = F(u_h) - F(u) \leq F(u_h) - F(u) =$$

$$= a(u-u_h,u-u_h)$$

Therefore,

$$\min_{u_h \in S_h} a(u-u_h,u-u_h) \leq a(u-u_h,u-u_h) \leq a(u-u_h,u-u_h)$$

from which we obtain:

$$\min_{u_h \in S_h} a(u-u_h,u-u_h) = a(u-u_h,u-u_h)$$

which completes the proof.

Next, if we subtract the following equation from the equation (1-22):

$$a(u,v_h) = (f,v_h)_{L_2(\Omega)}, \text{ for all } v_h \in S_h$$

which is the same as equation (1-23) applied for any function $v_h \in S_h$, we get:

$$a(u-u_h, v_h) = 0, \text{ for all } v_h \in S_h \qquad \ldots \quad (1\text{-}26)$$

This is the most remarkable result which has been obtained so far. Equation (1-26) states that, with respect to the energy inner-product $a(u,v)$, the finite element solution u_h is the projection of the exact solution u onto the space S_h, or, what is the same thing, that the error $u-u_h$ is orthogonal to the subspace S_h. As a result, *the problem of convergence in the Finite Element Method becomes a problem in the approximation theory and our main task is to estimate the distance between the function* u *and the subspace* S_h.

Nevertheless, it is not necessary to work directly with the finite element solution u_h. Instead, it is sufficient to consider any conveniently derived polynomial from the finite-dimensional space S_h which is as close to the exact solution u as possible. Then, since, by virtue of the minimum principle (1-24), the approximate solution u_h is always closer to u than any other function of the space S_h, we can immediately obtain a first upper bound for the difference between the function u and its finite element approximation u_h in terms of the energy norm. For such a conveniently derived polynomial. however, close enough to the exact solution u, we choose its interpolating

polynomial and, thus, *we are faced with a problem in approximation theory of giving an upper bound for the error between the function* u *and its interpolating polynomial.* The construction of such an interpolating polynomial, as will be seen in the sequel, is always possible for any function u which assumes any order of continuity over the entire domain Ω.

1.3.2 Consistency Plus Stability Implies Convergence

Let us consider once more the equation (1-26) applied to the function $u_h \in S_h$, viz:

$$a(u-u_h, u_h) = 0, \quad u_h \in S_h$$

or

$$a(u, u_h) = a(u_h, u_h)$$

Then, by using this result, we obtain:

$$a(u-u_h, u-u_h) = a(u,u) - a(u,u_h) - a(u_h,u) + a(u_h,u_h) =$$

$$= a(u,u) - a(u_h,u_h)$$

Thus,

$$a(u-u_h, u-u_h) = a(u,u) - a(u_h,u_h) \qquad \ldots \quad (1\text{-}27)$$

and equation (1-27) expresses the Pythagorean theorem for the Finite Element Method: the energy of the error is always equal to the error in the energy. Then, since:

$$a(u-u_h, u-u_h) \geq 0$$

from (1-27), we get

$$a(u_h, u_h) \leq a(u,u) \qquad \ldots \quad (1\text{-}28)$$

so that the energy in the finite element approximation u_h is always bounded by that of the exact solution u. Therefore, if we identify the inequality (1-28) with the stability condition of the method and the approximability - given that u can be approximated by the subspace S_h - with the consistency condition, then, by virtue of the fundamental principle in Numerical Analysis, ie, consistency plus stability implies convergence and conversely, the convergence in the Finite Element Method does occur within the subspace S_h and its order immediately follows from the minimum principle (1-24).

As far as the finite-dimensional subspace S_h is concerned, hereafter, we make the following two basic assumptions:

1. S_h is of general degree k-1; ie, it contains within each of its elements all the polynomials of degree less than or equal to k-1, and
2. for any subdivision of the domain Ω into finite elements, a uniformity condition is supposed to be satisfied by them, as the space parameter $h \to 0$.

This last condition imposed on the finite elements themselves is merely a geometrical condition which avoids degenerate

elements. Then, under these two conditions, and by using the results of the approximation theory of the next part, we shall prove that:

$$a(u-u_h,u-u_h) \leq Ch^{2(k-m)}|u|^2_{H^k(\Omega)} \quad \ldots \quad (1-29)$$

for any function $u \in H^k(\Omega)$, where m in (1-29) denotes the order of the highest derivative which is involved in the bilinear form $a(u,v)$, C is some numerical constant which does not depend on the function u and the parameter h and $|\cdot|_{H^k(\Omega)}$ is either a semi-norm or a norm. If we are able to estimate a bound for the quantity $|u|_{H^k(\Omega)}$ in terms of the second part of the original differential equation then, we have in (1-29) an a priori error bound which only depends on the data of the problem at hand. Furthermore, since the minimum principle (1-24) holds irrespective of homogeneous or inhomogeneous essential boundary conditions, the same is true for the error estimate (1-29) and it depends only on the order to which the solution u can be approximated by the trial space S_h composed, by construction, of piecewise polynomials.

Next, from the estimate (1-29) combined with the ellipticity condition (1-16), we get the following result concerning, this time, the H^m-norm:

$$\|u-u_h\|_{H^m(\Omega)} = O(h^{k-m}) \quad \ldots \quad (1-30)$$

A less straightforward problem, however, is to give an

estimate for the difference between u and u_h in terms of a different norm $\|\cdot\|_{H^s(\Omega)}$, with s being smaller or larger than m. In this case the application of the so-called *Nitsche trick* gives for the H^s-norm the following rate of convergence:

$$\|u-u_h\|_{H^s(\Omega)} = O(h^{k-s}+h^{2(k-m)}) \qquad \ldots \quad (1\text{-}31)$$

Nevertheless, in almost all practical applications of the method, the first exponent in (1-31) governs the rate of convergence and this is in agreement with the results of the approximation theory upon which the convergence of the Finite Element Method depends.

PART II

THE GENERAL MULTIVARIATE APPROXIMATION PROBLEM

2.1 FINITE ELEMENTS AND APPROXIMATING SUBSPACES

The expansion of the Finite Element Method has reached such a point nowadays that it has become one of the most popular and effective methods for the numerical solution of partial differential equations, particularly for elliptic equations. At the same time, it is a well-known fact that the main reason for the success of the method is reflected by its capability of dealing with complex geometrical regions by using arbitrarily shaped simple elements. Nevertheless, for reasons of simplicity, we restrict our attention at present to polyhedral type domains only and we shall examine in a later section of this Part the effect which an arbitrary curved domain has on the general approximation problem. Therefore, suppose that we are given a bounded and open subset Ω of the n-dimensional Euclidean space R^n whose n-polygonal boundary we denote by $\partial\Omega$. The implementation of the Finite Element Method starts with a partition (or subdivision or general triangulation):

$$\Delta:\{e_i\}_{i\in I} \qquad \ldots \quad (2\text{-}1)$$

of the domain Ω into a finite number of pieces $e_i \subset \Omega$, which we always call finite elements. We denote by Ω_h the interior of the union of these elements and by $\partial\Omega_h$ its respective boundary. The parameter h simply refers to the mesh spacing introduced by the particular partition (2-1). Note that, for our particular choice of the domain Ω the boundaries $\partial\Omega$ and $\partial\Omega_h$ coincide. Then, we say that a family T_h of such elements $e_i \subset \Omega$ constitute any admissible triangulation of the domain Ω, if and only if the following two basic conditions are satisfied:

(i) $\bigcup_{e_i \in T_h} e_i = \bar{\Omega}$

where $\bar{\Omega}$ is the closed domain resulting from the combination of Ω and $\partial\Omega$, and

(ii) If e_i and $e_j \in T_h$, then, either $e_i = e_j$, or $e_i \cap e_j = \emptyset$, or e_i and e_j have a common vertex or side.

Over each element e_i, a finite number of points a_i - usually called nodes or nodal points - are specified, some of which are common to several adjacent elements of the given triangulation. These points a_i constitute a set of interpolating points which in a unique fashion, and over each element e_i at a time, define an appropriate interpolating polynomial with a certain degree of continuity over the entire domain Ω. This order of continuity is required for the space S_h to be a subspace of the energy space $H^m(\Omega)$, ie, to be more

precise, for the case where:

$$S_h \subset H^m(\Omega)$$

the trial functions $v_h \in S_h$ are required to have m-1 continuous derivatives between interelement boundaries since, then, the m-th derivative will at most have a finite jump between adjacent elements and, therefore, it is possible for the energy to be found over the entire domain Ω by adding the separate contributions from within each element. Nevertheless, since the interpolating points a_i are most frequently associated with several adjacent elements, the basis functions $\phi_i(x)$ corresponding to these points, one or more with each point, constitute only parts of the complete basis function associated with such a point of the triangular network. In order to obtain the complete basis function at any node, we have to add up all the appropriate parts associated with the elements adjacent to the node. Before giving a brief description of some of the approximating subspaces S_h, we note that the subspace S_h is decomposed, with respect to the particular subdivision of the domain Ω satisfying the conditions (i) and (ii), into a finite number of subspaces such that:

$$S_h(\bar{\Omega}) = \bigcup_{e_i \in \Gamma_h} S_h(\bar{e}_i; \Gamma_h),$$

where, again, by $\bar{e}_i$ we understand the closed element e_i resulting from the combination of e_i and its boundary, and we are only concerned with the construction of such

subspaces $S_h(\bar{e}_i;T_h)$ for several types of elements e_i. As far as the dimension of the space S_h is concerned, which is exactly the number N of the free parameters of the trial functions $v_h \in S_h$, always coincides with the number of the unconstrained nodes. A node in the boundary where the function v_h is required to vanish, for example, or to equal to some other prescribed value, is constrained and will not count to the dimension of the subspace. With the natural boundary conditions assigned to the boundary nodes, however, the situation is completely different. There is no constraint on the trial functions $v_h \in S_h$ and the dimension of the subspace S_h equals the total number of the interior and boundary nodes.

There are two main categories of elements into which the given domain Ω can be divided, viz: either n-simplices (triangles for n=2, tetrahedra for n=3, etc) or unit n-hypercubes (unit squares for n=2, unit cubes for n=3, etc).

2.1.1 The Simplicial Finite Element

For the first, let (n+1) points $a_i, 1 \leq i \leq n+1$ be given in R^n with co-ordinates:

$$a_{1i}, a_{2i}, \ldots, a_{ni};\ 1 \leq i \leq n+1$$

and suppose that the matrix:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1,n+1} \\ a_{21} & a_{22} & \cdots & a_{2,n+1} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{n,n+1} \\ 1 & 1 & \cdots & 1 \end{bmatrix} \qquad \ldots (2\text{-}2)$$

is non-singular.

It is well-known that, if $x \in R^n$ with co-ordinates $(x_1, x_2, \ldots, x_n)$, it is uniquely represented in terms of the barycentric co-ordinates:

$$P_i, 1 \leq i \leq n+1$$

through the formula:

$$x = \sum_{i=1}^{n+1} P_i a_i, \quad \text{where} \quad \sum_{i=1}^{n+1} P_i = 1 \qquad \ldots (2\text{-}3)$$

Then, from (2-3) combined with (2-2), we can get:

$$x_i = \sum_{j=1}^{n+1} P_j a_{ij}, \quad \sum_{j=1}^{n+1} P_j = 1, \quad 1 \leq i \leq n \qquad \ldots (2\text{-}4)$$

which is a system of (n+1) linear equations and from which the vector:

$$(P_1, P_2, \ldots, P_{n+1})^T$$

corresponding to any given point:

$$(x_1, x_2, \ldots, x_n)^T$$

may be determined. The result is a linear, but generally *non-homogeneous, function of* x:

$$P_i = \sum_{j=1}^{n} b_{ij}x_j + b_{i,n+1},\ 1 \le i \le n+1 \qquad \ldots (2\text{-}5)$$

where the matrix $B = (b_{ij})$ in (2-5) is exactly the inverse matrix of (2-2). We call the closed convex hull $\bar{S}_n$ of the (n+1) points a_i, $1 \le i \le n+1$ (ie, the set of points of R^n with barycentric co-ordinates satisfying $0 \le p_i \le 1$, $1 \le i \le n+1$), the n-simplex generated by these points. The points a_i, $1 \le i \le n+1$, themselves constitute the vertices of the simplex. Furthermore, we define:

a. The barycenter G of $\bar{S}_n$ as that point of the simplex whose barycentric co-ordinates are all equal, and, therefore equal to $1/n+1$.

b. An m-dimensional face of $\bar{S}_n$ as the m-simplex, $1 \le m \le n-1$, generated by (m+1) vertices of $\bar{S}_n$. For an example of a 1-dimensional face we refer to the edge of the triangle.

c. If:

$$a_{k_i},\ 1 \le i \le m+1$$

are (m+1) members of the set $\{a_i\}_{1 \le i \le n+1}$, we define by:

$$H_m = \{x \in R^n:\ x = \sum_{i=1}^{m+1} P_i a_{k_i},\ \sum_{i=1}^{m+1} P_i = 1\}$$

a hyperplane in R^n of dimension m. This hyperplane contains the m-face defined by (b).

d. Let k be a fixed positive integer and let Z_k denote the following set of numbers:

$$\{0, 1/k, 2/k, \ldots, k-1/k, 1\}$$

Then, with the n-simplex $\bar{S}_n$ defined above, we associate the discrete set of points:

$$S_n(k) = \{x \in R^n : x = \sum_{i=1}^{n+1} P_i a_i, P_i \in Z_k, \sum_{i=1}^{n+1} P_i = 1\} \quad \ldots \quad (2-6)$$

which we call the k-th order principal lattice of the simplex $\bar{S}_n$. It contains exactly:

$$\binom{k+n}{k}$$

members which, in the practical applications of the Finite Element Method, constitute the interpolating points of the simplex.

Given a function u(x), with $x = (x_1, x_2, \ldots, x_n)$ an n-variable, defined over a (k-1)-st order principal lattice associated with the simplex $\bar{S}_n$ and assuming a certain degree of continuity over the closed element $\bar{S}_n$, the approximation problem can be described in a few words as follows: The values of the function u(x) are interpolated at the points a_i of the set $S_n(k-1)$. This is the general Lagrange interpolation problem. In addition to the values of the function u(x) at the points $a_i \in S_n(k-1)$, the values of some of its partial derivatives at several points of the set $S_n(k-1)$ are also interpolated. This is the general Hermite interpolation problem. In either case, however, given a function u(x) which is defined at a finite number of points $a_i \in S_n(k-1)$ and assumes a certain degree of contiuity over the simplex $\bar{S}_n$, its (Lagrange or Hermite) interpolating polynomial of general degree (k-1) is given through the formula:

$$u_n^{(k-1)}(x) = \sum_{i_1+i_2+\ldots+i_n \leq k-1} C_{i_1,i_2,\ldots,i_n} x_1^{i_1} \cdot x_2^{i_2} \ldots x_n^{i_n} \quad \ldots (2\text{-}7)$$

where the coefficients $C_{i_1,i_2,\ldots,i_n}$ in (2-7) are uniquely determined through the interpolating constraints. Then, by introducing the barycentric co-ordinates in our analysis, we consider the following special cases for the polynomial (2-7):

1. k=2. The interpolating polynomial is linear:

$$u_n^{(1)}(x) = \sum_{i_1+i_2+\ldots+i_n \leq 1} C_{i_1,i_2,\ldots,i_n} x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n} = \sum_{i=1}^{n+1} u(\alpha_i) \phi_i^{(1)}(x) \quad \ldots (2\text{-}8)$$

where $u(\alpha_i)$, $1 \leq i \leq n+1$, are the values of the function $u(x)$ at the points $\alpha_i \in S_n(1)$ - ie, the vertices of the simplex $\bar{S}_n$ - and $\phi_i^{(1)}(x)$, $1 \leq i \leq n+1$, are the basis functions associated with the points α_i, $1 \leq i \leq n+1$, and which, however, coincide with the barycentric co-ordinates of the point $x \in R^n$ with respect to the (n+1) vertices of the simplex. Obviously:

$$\phi_i^{(1)}(\alpha_j) = \delta_{ij}, \quad 1 \leq i, j \leq n+1$$

and, therefore, they constitute a class of linear basis functions over the simplex $\bar{S}_n$. They are related to the Cartesian co-ordinates through the formula:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ 1 \end{bmatrix} = A \cdot \begin{bmatrix} \phi_1^{(1)} \\ \phi_2^{(1)} \\ \vdots \\ \phi_n^{(1)} \\ \phi_{n+1}^{(1)} \end{bmatrix} \quad \ldots (2\text{-}9)$$

where A is the same as the matrix (2-2) and, since it has been assumed to be non-singular, we can get:

$$\begin{bmatrix} \phi_1^{(1)} \\ \phi_2^{(1)} \\ \vdots \\ \phi_n^{(1)} \\ \phi_{n+1}^{(1)} \end{bmatrix} = A^{-1} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ 1 \end{bmatrix} \qquad \ldots \quad (2\text{-}10)$$

The two transformation formulae (2-9) and (2-10) between Cartesian and barycentric co-ordinates (since $\phi_i^{(1)}(x) = p_i, 1 \leq i \leq n+1$) are of exceptional importance since they may transform any element of an arbitrary (but straight) shape into its standard form - it is always easier to work in terms of the right triangle, eg, than one of arbitrary shape - and vice versa. The interpolating polynomial (2-8) is defined by P.G. Ciarlet and W. Wagschal [6] as the interpolating polynomial of type I, with the interpolating polynomial of type II being the following:

2. $k=3$. The interpolating polynomial is quadratic:

$$u_n^{(2)}(x) = \sum_{i_1+i_2+\ldots i_n \leq 2} C_{i_1,i_2,\ldots,i_n} x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n} =$$

$$= \sum_{i=1}^{n+1} u(\alpha_i)\phi_i^{(2)}(x) + \sum_{\substack{i,j=1 \\ i \neq j}}^{n+1} u(\alpha_{ij})\phi_{ij}^{(2)}(x) \qquad \ldots \quad (2\text{-}11)$$

where $u(\alpha_i)$ are the values of the function $u(x)$ at the vertices α_i, $1 \le i \le n+1$, of the simplex $\bar{S}_n$ and $u(\alpha_{ij})$ the values of the function at the mid-points of the edges (α_i, α_j) generated by the vertices α_i and α_j, $1 \le i, j \le n+1$, with the convention that we always have:

$$\alpha_{ij} = \alpha_{ji}, \quad 1 \le i, j \le n+1$$

Note that the set of points:

$$\{\alpha_i\} \cup \{\alpha_{ij}\}$$

constitute a second-order principal lattice for the simplex $\bar{S}_n$. Furthermore, the basis functions in (2-11) are given, in terms of the linear basis functions $\phi_i^{(1)}(x)$, $1 \le i \le n+1$, as follows:

$$\phi_i^{(2)}(x) = \phi_i^{(1)}(2\phi_i^{(1)}-1), \quad 1 \le i \le n+1$$
$$\phi_{ij}^{(2)}(x) = 4\phi_i^{(1)}\phi_j^{(1)}, \quad 1 \le i, j \le n+1, \ i \ne j \qquad \ldots \quad (2\text{-}12)$$

3. k=4. The interpolating polynomial is cubic:

$$u_n^{(3)}(x) = \sum_{i_1+i_2+\ldots+i_n \le 3} C_{i_1,i_2,\ldots,i_n} x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n} = \sum_{i=1}^{n+1} u(\alpha_i)\phi_i^{(3)}(x) +$$

$$+ \sum_{\substack{i,j=1 \\ i \ne j}}^{n+1} u(\alpha_{iij})\phi_{iij}^{(3)}(x) + \sum_{\substack{i,j,k=1 \\ i \ne j \ne k}}^{n+1} u(\alpha_{ijk})\phi_{ijk}^{(3)}(x) \qquad \ldots \quad (2\text{-}13)$$

where the set of points:

$$\{\alpha_i\} \cup \{\alpha_{iij}\} \cup \{\alpha_{ijk}\}$$

constitute a third-order principal lattice ror the simplex $\bar{S}_n$; ie,

$$\alpha_{iij} = (2\alpha_i+\alpha_j)/3,\ 1 \le i,\ j \le n+1,\ i \ne j$$

$$\alpha_{ijk} = (\alpha_i+\alpha_j+\alpha_k)/3,\ 1 \le i,\ j,k \le n+1,\ i \ne j \ne k$$

and the basis functions can be given, in terms of the linear basis functions $\phi_i^{(1)}(x)$, $1 \le i \le n+1$, as follows:

$$\phi_i^{(3)}(x) = \tfrac{1}{2}\phi_i^{(1)}(3\phi_i^{(1)}-1)(3\phi_i^{(1)}-2),\ 1 \le i \le n+1$$

$$\phi_{iij}^{(3)}(x) = 9/2\phi_i^{(1)}\phi_j^{(1)}(3\phi_i^{(1)}-1),\ 1 \le i,\ j \le n+1,\ i \ne j \quad \ldots\ (2\text{-}14)$$

$$\phi_{ijk}^{(3)}(x) = 27\phi_i^{(1)}\phi_j^{(1)}\phi_k^{(1)},\ 1 \le i,\ j,k \le n+1,\ i \ne j \ne k$$

Nevertheless, within each 2-face of the simplex generated by the vertices α_i, α_j and α_k, $1 \le i,\ j,k \le n+1$, $i \ne j \ne k$, and with the point α_{ijk} being its barycenter, the values $u(\alpha_{ijk})$ of the function $u(x)$ at the points:

$$\alpha_{ijk} = (\alpha_i+\alpha_j+\alpha_k)/3,\ 1 \le i,\ j,k \le n+1,\ i \ne j \ne k$$

in (2-13) can be replaced by a linear combination of the form:

$$u(\alpha_{ijk}) = \tfrac{1}{4}[u(\alpha_{iij}) + u(\alpha_{jji}) + u(\alpha_{jjk}) + u(\alpha_{kkj}) + u(\alpha_{kki}) + u(\alpha_{iik})] -$$

$$+ 1/6[u(\alpha_i) + u(\alpha_j) + u(\alpha_k)] \quad . \quad . \quad . \quad (2\text{-}15)$$

This technique, widely used by the engineers, merely eliminates the internal nodes and, even though in this case the order of accuracy of the Finite Element Method is decreased by one, it is quite often applied in practice since the computing time which is needed is considerably decreased. For example, following M. Zlamal [30], consider the case where n=2, ie, the simplex $\bar{S}_n$ becomes a triangle, and no boundary conditions are prescribed for the trial functions. Also assume that the domain Ω is a unit square which has been decomposed into $2n^2$ triangles by dividing it, first, into n^2 squares of equal sides 1/n and, then, every such square into two triangles. The total number of vertices constructed so far by this triangulation equals $(n+1)^2$. On the other hand, the computer time needed to solve the final linear system KQ = F is proportional to Nw^2, where N is the number of equations in the system and 2w + 1 the band width of the stiffness matrix K. If polynomials with ten parameters over each triangle - which correspond to the three vertices, the three mid-points of the edges and the centroid of the triangle - are used, we have that:

$$N = 3(n+1)^2 + 2n^2$$

$$w = 5(n+1)$$

in which case the computer time needed approximately amounts to:

$$Nw^2 = [3(n+1)^2+2n^2]25(n+1)^2 \cong 125n^4$$

On the contrary, if polynomials with nine parameters are used - the tenth parameter corresponding to the centroid of the triangle having been eliminated by a linear combination similar to that given by the formula (2-15) - over each triangle, we have:

$$N = 3(n+1)^2$$

$$w = 3n+5$$

for which case:

$$Nw^2 = 3(n+1)^2(3n+5)^2 \cong 27n^4$$

which means that by eliminating the internal node we can save approximately 78% of the computer time.

However, since the dimension of the subspace S_h grows enormously fast with k, an interesting problem arises from the possibility of imposing further constraints on the interpolating function without destroying either the approximation properties or the simplicity of the local basis. Thus, for $k = 4$, we consider the following cubic Hermite interpolating polynomial which can uniquely be determined through the interpolating constraints:

(i) $u_n^{(3)}(\alpha_i) = u(\alpha_i)$, $1 \le i \le n+1$

(ii) $u_n^{(3)}(\alpha_{ijk}) = u(\alpha_{ijk})$, $\alpha_{ijk} = (\alpha_i+\alpha_j+\alpha_k)/3$, $1 \le i, j, k \le n+1$, $i \ne j \ne k$

(iii) $Du_n^{(3)}(\alpha_i) = Du(\alpha_i)$, $1 \le i \le n+1$.

Then, the interpolating polynomial - although rarely used

because of its complexity - is given through the formula (see A.P. Mitchell [14]):

$$u_n^{(3)}(x) = \sum_{i=1}^{n+1} u(\alpha_i)\,(3\phi_i^{(1)2} - 2\phi_i^{(1)3}) +$$

$$+ 1/6 \sum_{\substack{i,j,k=1 \\ i\neq j\neq k}}^{n+1} [27u(\alpha_{ijk}) - 7(u(\alpha_i)+u(\alpha_j)+u(\alpha_k))]\cdot\phi_i^{(1)}\phi_j^{(1)}\phi_k^{(1)} +$$

$$+ \sum_{\substack{i,j=1 \\ i\neq j}}^{n+1} Du(\alpha_i)\,(\overline{\alpha_i\alpha_j})\,\phi_i^{(1)2}\phi_j^{(1)} - \sum_{\substack{i,j,k=1 \\ i\neq j\neq k}}^{n+1} Du(\alpha_i)\,(\overline{\alpha_i\alpha_j})\,\phi_i^{(1)}\phi_j^{(1)}\phi_k^{(1)}$$

. . . (2-16)

where $(\overline{\alpha_i\alpha_j})$ denotes the length of the edge generated by the vertices α_i and α_j, $1 \le i,j \le n+1$, $i \neq j$. Again, the values $u(\alpha_{ijk})$ of the function $u(x)$ corresponding to the internal nodes α_{ijk}, $1 \le i,j,k \le n+1$, $i \neq j \neq k$, can be eliminated by a substitution, in terms of the others, of the following form:

$$u(\alpha_{ijk}) = 1/3[u(\alpha_i)+u(\alpha_j)+u(\alpha_k)] - 1/6[Du(\alpha_i)\cdot(\alpha_i-\alpha_{ijk}) + Du(\alpha_j)\cdot(\alpha_j-\alpha_{ijk}) +$$

$$+ Du(\alpha_k)\cdot(\alpha_k-\alpha_{ijk})] \quad . \; . \; . \quad (2-17)$$

and a considerable amount of computer time can again be saved.

2.1.2 The Hypercubic Finite Element

As far as the second main category of elements mentioned earlier is concerned, ie, the unit n-hypercube $\bar{\Pi}_n$, let $k \geq 1$ be any integer and let N_k denote the following set of numbers:

$$N_k = \{0,1,2\ldots,k\}$$

Then, in an exactly analogous way as for the n-simplex described earlier, we call the closed convex hull of the set:

$$\Pi_n(k) = \{x \in R^n : x=(\frac{i_1}{k},\frac{i_2}{k},\ldots,\frac{i_n}{k}),\ i_j \in N_k,\ 1 \leq j \leq n\} \quad \ldots \quad (2\text{-}18)$$

the unit n-hypercube $\bar{\Pi}_n$ of R^n. The set of points $\Pi_n(k)$, defined by (2-18 ins exactly $(k+1)^2$ members which, as for the set he simplex $\bar{S}_n$, in the practical applications of the ite Element Method constitute the discrete set of interpolating points for the hypercubic element. Thus, suppose that a function $u(x)$, $x=(x_1,x_2,\ldots,x_n) \in R^n$ is an n-variable, is defined over the set of points $\Pi_n(k-1)$ associated with the hypercube $\bar{\Pi}_n$ and assumes a certain degree of continuity over the closed element $\bar{\Pi}_n$. Then, we define its (Lagrange or Hermite) interpolating polynomial of general degree (k-1) as the unique polynomial of that degree which interpolates the values of the function $u(x)$ together, perhaps, with the values of some of its partial derivatives at several points $a_i \in \Pi_n(k-1)$. This polynomial

in its general form is given through the formula:

$$u_n^{(k-1)}(x) = \sum_{i_1,i_2,\ldots,i_n \leq k-1} b_{i_1,i_2,\ldots,i_n} x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n} \quad \ldots \quad (2\text{-}19)$$

Finally, we only mention here that, the technique of elimination of the internal nodes can be applied in the same way for the hypercubic element, as it was for the n-simplex, leading to the serendipity family of elements frequently used in engineering applications.

2.1.3 Finite Elements in the Plane

In the case of a planar domain Ω, the triangle or 2-dimensional simplex is the most widely used finite element in practice. Among other reasons justifying its great popularity are the following two:

1. Any artibrarily curved domain in the space R^2 can be approximated by a polygon which, in turn, can always be divided into a finite number of triangles, and

2. the boundary of any curved domain can always better be approached by using a refined mesh of triangles.

Of course, there are also some advantages as far as the rectangular element is concerned. It can be used for the interior of the domain, where there are fewer of them than triangles, and it seems that an appropriate mixture of triangular and rectangular elements can produce an

excellent subdivision of any domain in the plane. The crucial point, however, when such a mixture of finite elements is involved, is to make sure that the required degree of continuity of the interpolating function across the junction is secured. In analogy with the interpolating polynomial (2-7) defined over the simplex $\bar{S}_n$, over the triangle T we have the polynomial:

$$u_2^{(k-1)}(x,y) = \sum_{i+j=0}^{k-1} c_{ij} x^i y^j \qquad \ldots \quad (2\text{-}20)$$

which simply interpolates the values - and probably those of some of its partial derivatives - of a function u(x,y) which is well-defined over a (k-1)-st order principal lattice $S_2(k-1)$ associated with the triangle T. For the particular value of k=2, however, the polynomial (2-20) is uniquely determined by its values at the three vertices of the triangle and the corresponding linear trial subspace S_h exactly coincides with that proposed by Courant [9] some 35 years ago. Then, the quadratic and cubic polynomials immediately follow from the polynomial (2-20) and for the values of k=3 and k=4, respectively, in analogy to those described earlier for the simplicial element. However, for any particular value of the parameter k, the interpolating polynomial (2-20) reduces to a polynomial of degree (k-1) in one variable s measured along the edge of the triangle. This feature is in common with all the triangles of the given triangular network, and, therefore, the interpolating function (2-20) is of

class C^0 over the entire domain Ω. *Nevertheless, instead* of interpolating the function $u(x,y)$ at a large number of points, and thus increasing the dimension of the subspace S_h, it is possible to impose further constraints on the interpolating function without destroying the accuracy of the approximation. Thus, for k=4, we are faced with the Hermite cubic interpolating polynomial which is uniquely *determined* by the values of the function $u(x,y)$ as well as those of its first-order partial derivatives at the three vertices of the triangle. This polynomial, also being of class C^0 over the domain Ω, cannot be used for the solution of fourth order equations where a C^1 continuity for the trial functions is required. For the construction of such an interpolating polynomial we demand that, not only the function be continuous between adjacent triangles but also, its normal derivatives as well. This gives rise to the quintic polynomial which is very useful in practice and which is determined by the values of the function $u(x,y)$ and those of its first- and second-order partial derivatives at the three vertices of the triangle as well as the values of its normal derivatives at the mid-points of the edges. Finally, note that by using the following transformation formula:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} \qquad \ldots \text{(2-21)}$$

the polynomial (2-20) can be transformed into an equivalent

polynomial whose basis functions are expressed in terms of the barycentric co-ordinates (p_1,p_2,p_3) of the point $x \in R^2$ with respect to the three vertices $\alpha_1 = (x_1,y_1)$, $\alpha_2 = (x_2,y_2)$ and $\alpha_3 = (x_3,y_3)$ of the triangle T.

On the other hand, again over a planar domain Ω, and with the interpolating polynomial (2-19) defined over the unit hypercube $\bar{\Pi}_n$, for the unit square $S = [0,1] \times [0,1]$ we have the following polynomial:

$$u_2^{(k-1)}(x,y) = \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} b_{ij} x^i y^j \qquad \ldots \quad (2\text{-}22)$$

which, again, interpolates the values - and perhaps those of some of its partial derivatives - of a function $u(x,y)$ which is well-defined over a specified set of points $\alpha_i \in \Pi_2(k-1)$ given through the formula (2-18). Note that any rectangular element $[a,b] \times [c,d]$ can be transformed into the unit square element $[0,1] \times [0,1]$ through the following transformation:

$$x \to \frac{x-a}{b-a}, \quad y \to \frac{y-c}{d-c} \qquad \ldots \quad (2\text{-}23)$$

For the particular value of k=2, the interpolating polynomial (2-22) is a bilinear function - ie, it is linear with respect to each one of the two variables whenever the other is kept fixed - and is uniquely determined by its values at the four vertices of the square. This in analogy to the linear case described over the triangle,

corresponds to the simplest construction of the bilinear trial subspace S_h, whereas, the values of k=3 and k=4 give rise to the biquadratic and bicubic polynomials respectively. However, again, instead of interpolating the function u(x,y) at a large number of points we may impose further constraints on the interpolating function provided that by doing so we do not destroy the accuracy of the approximation. This, for the value of k=4, gives a Hermite bicubic interpolating polynomial which is uniquely determined by the values of the function u(x,y) and those of its partial derivatives:

$$\frac{\partial u(x,y)}{\partial x}, \quad \frac{\partial u(x,y)}{\partial y}, \quad \frac{\partial^2 u(x,y)}{\partial x \partial y}$$

at the four vertices of the square. Let us comment a little further on that polynomial. Suppose that the domain Ω is of rectangular type and has been divided into a finite number of rectangular elements which, in turn, through the transformation formulae (2-23) can be transformed into unit squares. Then, for any internal node, each basis function has a support of four squares over the entire subdivision, whereas, for any boundary node - not a corner node its support is only two elements. Furthermore, for a node in the corner, the support of the corresponding basis function is only one element. Since continuity of the first derivative is also secured between interelement boundaries, the Hermite bicubic interpolating polynomial belongs to the class of functions $C^{1,1}$. In general, we

say that a function u(x,y) belongs to the class $C^{i,j}$, if the derivatives:

$$\frac{\partial^{k+\ell} u(x,y)}{\partial x^k \partial y^\ell} \quad , \quad 0 \le k \le i, \quad 0 \le \ell \le j$$

are continuous over the entire region. An alternative interpolating polynomial, however, to the Hermite bicubic function is provided by the bicubic spline function. It is the tensor product of two cubic splines in the one-dimensional space and its support is sixteen elements, provided that the associated node is not on the boundary or adjacent to the boundary. The continuity of the bicubic spline is of class $C^{2,2}$ instead of $C^{1,1}$ of the Hermite bicubic. Furthermore, for a $C^{4,4}$ continuous polynomial we may consider the tensor product of the two quintic splines in one dimension. Its support now has been increased considerably to thirty-six elements and this fact makes the biquintic spline rather difficult to handle. We point out, however, that whereas the support of the splines increases with the order of the spline - it is $4k^2$ elements for a spline of degree 2k-1 -, the support of the Hermite function remains unchanged at the four elements irrespective of the order of the polynomial. The difference, of course, is that the splines give greater continuity, viz:

$$C^{2(k-1),2(k-1)}$$

instead of $C^{k-1,k-1}$ of the Hemite function.

2.1.4 The Quadrilateral Element and the Isoparametric Technique

The important situation which arises in considering the quadrilateral element is that, the continuity which was achieved by the interpolating polynomial between adjacent rectangular elements does not, in general, hold between arbitrary quadrilaterals. For example, consider the bilinear function defined through the formula (2-22) and k=2. If the two quadrilaterals are joined by a line:

$$y = mx + b$$

then, along that edge, the bilinear function reduces to a quadratic and, thus, cannot uniquely be determined from the values of the function $\phi(x,y)$ at the two vertices only. It reduces to a linear polynomial along that edge if and only if the edge is horizontal or vertical. The possibility, however, of achieving a C^0 continuity for the interpolating function over an overall quadrilateral network becomes a challenging mathematical problem and a practical technique employed for its solution is the following: change the co-ordinates in such a way that the quadrilateral becomes a rectangle - or rather a unit square - and the interpolating functions in the new co-ordinates are admissible. This is the well-known isoparametric technique and it merely consists of choosing piecewise polynomials such that:

(i) they define the co-ordinate transformation, and

(ii) the same polynomials can be used as interpolating functions over each element.

For a particular example, consider the quadrilateral Q with four nodes placed at its four vertices $a_1 = (x_1,y_1)$, $a_2 = (x_2,y_2)$, $a_3 = (x_3,y_3)$ and $a_4 = (x_4,y_4)$. Then, the following linear mapping:

$$x = x_1 + (x_2-x_1)\xi + (x_3-x_1)\eta + (x_1-x_2-x_3+x_4)\xi\eta$$
$$y = y_1 + (y_2-y_1)\xi + (y_3-y_1)\eta + (y_1-y_2-y_3+y_4)\xi\eta \quad \ldots (2\text{-}24)$$

transforms the square S, with vertices $a_1' = (0,0)$, $a_2' = (1,0)$, $a_3' = (1,1)$ and $a_4' = (0,1)$ in the (ξ,η)-plane, into the quadrilateral Q. It is an easy matter then to check out, from eqs (2-24), that the boundaries (see Fig 1 below) of Q and S correspond; ie,

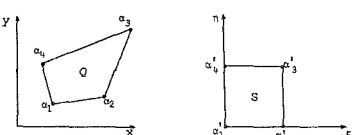

Fig 1:

$\xi = 0,1$ along the sides a_1a_4, a_2a_3 and $\eta = 0,1$ along the sides a_1a_2 and a_4a_3 respectively. Although such a correspondence between the boundaries actually occurs, it is also necessary to show that the mapping (2-24) is invertible so that each point (x,y) in Q corresponds to one and only one point (ξ,η) in S. This, in turn, can be shown by merely proving that the Jacobian matrix of the transformation (2-24) is non-zero inside the square S. This

Jacobian is:

$$J(\xi,\eta) = \det\begin{bmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial \eta} \\ \frac{\partial y}{\partial \xi} & \frac{\partial y}{\partial \eta} \end{bmatrix} = \det\begin{bmatrix} x_2-x_1+(x_1-x_2-x_3+x_4)\eta & x_3-x_1+(x_1-x_2-x_3+x_4)\xi \\ y_2-y_1+(y_1-y_2-y_3+y_4)\eta & y_3-y_1+(y_1-y_2-y_3+y_4)\xi \end{bmatrix}$$

and G. Strang and G. Fix in [24] have shown that the necessary and sufficient condition for the non-vanishing of the Jacobian inside S is that the quadrilateral is convex. Finally, since the same mapping (2-24) gives the interpolating polynomial over the quadrilateral Q, from (2-24), we can easily get:

$$x = (1-\xi)(1-\eta)x_1 + \xi(1-\eta)x_2 + \eta(1-\xi)x_3 + \xi\eta x_4$$

$$y = (1-\xi)(1-\eta)y_1 + \xi(1-\eta)y_2 + \eta(1-\xi)y_3 + \xi\eta y_4$$

from which we obtain:

$$u_2^{(1)}(x,y) = (1-\xi)(1-\eta)u(\alpha_1) + \xi(1-\eta)u(\alpha_2) + \eta(1-\xi)u(\alpha_3) + \xi\eta u(\alpha_4) =$$

$$= \sum_{i=1}^{4} u(\alpha_i)\phi_i(\xi,\eta)$$

where

$$\phi_1(\xi,\eta) = (1-\xi)(1-\eta) \qquad \phi_3(\xi,\eta) = \eta(1-\xi)$$

$$\phi_2(\xi,\eta) = \xi(1-\eta) \qquad \phi_4(\xi,\eta) = \xi\eta.$$

2.2 MULTIVARIATE POINTWISE APPROXIMATION

We shall make an attempt, in this section, to derive some error bounds for piecewise polynomial approximation over a domain $\Omega \subset R^n$ of the following type: Ω is a polyhedral type domain which through a partition:

$$\Delta : \{\bar{S}_n^\nu\}_{\nu \in I} \qquad \ldots (2\text{-}25)$$

has been decomposed into a finite number of non-degenerate contiguous n-simplices $\bar{S}_n^\nu$, $\nu \in I$, such that:

$$\bar{\Omega} = \bigcup_{\nu \in I} \bar{S}_n^\nu$$

Then, error bounds for piecewise polynomial approximation over the polyhedral domain Ω can be derived on the basis of error bounds for polynomial approximation over a non-degenerate simplex. Thus, with every n-simplex $\bar{S}_n^\nu$, $\nu \in I$, of the partition Δ associate the following discrete set of interpolating points:

$$S_n^\nu(k-1) = \{x \in R^n : x = \sum_{i=1}^{n+1} P_i^\nu \alpha_i^\nu, P_i^\nu \in Z_{k-1}, \sum_{i=1}^{n+1} P_i^\nu = 1\}, \nu \in I \quad \ldots (2\text{-}26)$$

where $Z_{k-1} = \{0,1/k,\ldots,k-2/k,1\}$ and P_i^ν, $1 \le i \le n+1$, $\nu \in I$, denote the barycentric co-ordinates of any point $x = (x_1, x_2, \ldots, x_n) \in R^n$ with respect to the (n+1) vertices α_i^ν, $1 \le i \le n+1$, of the simplex $\bar{S}_n^\nu$, $\nu \in I$. The set (2-26) exactly defines a (k-1)-st order principal lattice of the simplex, which was again introduced in an earlier stage, and contains precisely:

$$\binom{k+n-1}{k-1}$$

members. Next, the following set of functions:

$$P_{k-1}(\bar{S}_n^\nu) = \{p(x) : \bar{S}_n^\nu \to R : p(x) = \sum_{|i| \le k-1} C_i x^i, \; x \in \bar{S}_n^\nu, \; \nu \in I\} \qquad \ldots (2\text{-}27)$$

$$\Delta : \{\bar{S}_n^\nu\}_{\nu \in I} \qquad \ldots (2\text{-}25)$$

has been decomposed into a finite number of non-degenerate contiguous n-simplicos $\bar{S}_n^\nu$, $\nu \in I$, such that:

$$\bar{\Omega} = \bigcup_{\nu \in I} \bar{S}_n^\nu$$

Then, error bounds for piecewise polynomial approximation over the polyhedral domain $\bar{\Omega}$ can be derived on the basis of error bounds for polynomial approximation over a non-degenerate simplex. Thus, with every n-simplex $\bar{S}_n^\nu$, $\nu \in I$, of the partition Δ associate the following discrete set of interpolating points:

$$S_n^\nu(k\text{-}1) = \{x \in R^n : x = \sum_{i=1}^{n+1} P_i^\nu a_i^\nu, P_i^\nu \in Z_{k-1}, \sum_{i=1}^{n+1} P_i^\nu = 1\}, \nu \in I \quad \ldots (2\text{-}26)$$

where $Z_{k-1} = \{0, 1/k, \ldots, k-2/k, 1\}$ and P_i^ν, $1 \le i \le n+1$, $\nu \in I$, denote the barycentric co-ordinates of any point $x = (x_1, x_2, \ldots, x_n) \in R^n$ with respect to the (n+1) vertices a_i^ν, $1 \le i \le n+1$, of the simplex $\bar{S}_n^\nu$, $\nu \in I$. The set (2-26) exactly defines a (k-1)-st order principal lattice of the simplex, which was again introduced in an earlier stage, and contains precisely:

$$\binom{k+n-1}{k-1}$$

members. Next, the following set of functions:

$$P_{k-1}(\bar{S}_n^\nu) = \{p(x) : \bar{S}_n^\nu \to R : p(x) = \sum_{|i| \le k-1} C_i x^i,\ x \in \bar{S}_n^\nu,\ \nu \in I\} \qquad \ldots (2\text{-}27)$$

where $i = (i_1, i_2, \ldots, i_n)$ and $|i| = i_1+i_2+\ldots+i_n$ is the usual multi-index notation, defines the class of all the polynomials of degree less than or equal to (k-1) in n-variables $x = (x_1, x_2, \ldots, x_n) \in R^n$ restricted over the simplex $\bar{S}_n^\nu$, $\nu \in I$. It is within that class of functions that we are searching for the interpolating polynomial defined over the simplex $\bar{S}_n^\nu$, $\nu \in I$, and its dimension - since the dimension of the subspace restricted over the simplex $\bar{S}_n^\nu$, $\nu \in I$, coincides with the number N of all unconstrained nodes of the simplex - is given by:

$$N = \dim P_{k-1}(\bar{S}_n^\nu) = \binom{k+n-1}{k-1} \qquad \ldots \quad (2\text{-}28)$$

Consider now a function:

$$u(x) \in C^k(\bar{S}_n^\nu), \quad \nu \in I$$

which is well-defined over the discrete set of points $S_n^\nu(k-1)$, $\nu \in I$, where:

$$C^k(\bar{S}_n^\nu) = \{u(x) : u(x) \in C(\bar{S}_n^\nu);\ D^\alpha u(x) \in C(\bar{S}_n^\nu), \text{ for all } |\alpha| \leq k, \nu \in I\} \qquad \ldots \quad (2\text{-}29)$$

and $C(\bar{S}_n^\nu)$ is the set of all real-valued functions u(x) which are continuous over the simplex $\bar{S}_n^\nu$, $\nu \in I$. The unique function:

$$u_n^{(k-1)}(x) \in P_{k-1}(\bar{S}_n^\nu), \quad \nu \in I$$

which assumes the same values as the function u(x) at the N discrete points of the set $S_n^\nu(k-1)$, $\nu \in I$, as well as

those of some of its derivatives as the derivatives of the function $u(x)$ at several points of the same set, defines its Hermite interpolating polynomial over the simplex $\bar{S}_n^\nu$, $\nu \in I$. The case of the Lagrange approximation clearly consists of a special case of the Hermite approximation problem where only the function values are interpolated. Therefore, considering an analysis of the more general Hermite approximation problem over the simplex $\bar{S}_n^\nu$, $\nu \in I$, the Lagrange case is essentially included in the analysis and we shall only briefly outline some of the main points as far as the Lagrange approximation problem is concerned. In either case, however, given a function $u(x)$ which assumes a C^k order of continuity over the simplex $\bar{S}_n^\nu$, $\nu \in I$, we shall prove the following error estimate:

$$\max_{\substack{x \in \bar{S}_n^\nu \\ \nu \in I \\ |\alpha|=m}} |D^\alpha u(x) - D^\alpha u_n^{(k-1)}(x)| \leq C \max_{\substack{x \in \bar{S}_n^\nu \\ \nu \in I \\ |\beta|=k}} |D^\beta u(x)| h_\nu^{k-m}, \quad 0 \leq m \leq k-1 \quad \ldots \quad (2\text{-}30)$$

where h_ν is a geometrical parameter closely associated with the simplex $\bar{S}_n^\nu$, $\nu \in I$, and C is some numerical constant which does not depend upon the discrete set of points $S_n^\nu(k-1)$, $\nu \in I$, defined by (2-26).

2.2.1 Fréchet Differentiation

Since we have to deal here with the multivariate approximation problem based on multivariate analysis, an extensive use of the Fréchet differential calculus will naturally be

unavoidable, so we first have to give some basic notations, definitions and results concerning this more general concept of differentiation:

Suppose that E and F are two real normed linear spaces and A is a non-empty open subset of E. Then, if u denotes a mapping of A into F:

$$u : A \to F$$

it is said to be differentiable at a point $\alpha \in A$, if and only if there exists a mapping T of E into F:

$$T : E \to F$$

which satisfies the following condition: for any $\varepsilon > 0$ there exists a $\delta > 0$ such that:

$$\|u(x) - u(\alpha) - T(x-\alpha)\| \leq \varepsilon \|x-\alpha\|, \text{ for all } x \in A \quad \ldots \quad (2\text{-}31)$$

whenever:

$$\|x - \alpha\| \leq \delta.$$

Furthermore, we say that the mapping u is differentiable on A if and only if it is differentiable at each point $\alpha \in A$. If u is differentiable at a point $\alpha \in A$, then, there exists a unique linear transformation T from E into F which satisfies (2-31). We call this unique linear transformation T the Fréchet derivative or Fréchet differential of the function u at the point $\alpha \in A$ and denote it by $Du(\alpha)$. Its application to a point $x \in E$ is written, for sake of simplicity, as:

$$Du(\alpha).x$$

by:

$$\|\mathcal{D}u(a).x\| = \sup_{\|x\|_E \leq 1} |\mathcal{D}u(a).x| \qquad \ldots \quad (2\text{-}32)$$

where $\|\cdot\|_E$ denotes the norm-operator over the space E, we define the bound or norm of the operator $\mathcal{D}u(a)$. Also, we denote by:

$$L(E;F)$$

the class of all the bounded and linear mappings from E into F.

We consider now the special case where $E \equiv R^n$ and $F \equiv R$. Then, the Fréchet derivative $\mathcal{D}u(a)$ of the function u at a point $a \in R^n$ is the unique linear transformation from R^n into R, ie,

$$\mathcal{D}u(a) \in L(R^n;R)$$

such that its application to a point $\xi = (\xi_1,\xi_2,\ldots,\xi_n) \in R^n$ gives the following real number:

$$\mathcal{D}u(a).(\xi_1 \quad .,\xi_n) \in R$$

Then, in analogy to the formula (2-32), the norm of the operator $\mathcal{D}u(a)$ is defined by:

$$\|\mathcal{D}u(a)\| = \sup_{\substack{|\xi_i| \leq 1 \\ 1 \leq i \leq n}} |\mathcal{D}u(a).(\xi_1,\xi_2,\ldots,\xi_n)| \qquad \ldots \quad (2\text{-}33)$$

Nevertheless, the following alternative definition for the norm of the operator $\mathcal{D}u(a)$, to that given by (2-33), will be useful in the sequel:

$$\|\mathcal{D}u(\alpha)\| = \sup_{\substack{\xi\in R^n \\ \xi\neq 0}} \frac{|\mathcal{D}u(\alpha).\xi|}{\|\xi\|} \qquad \ldots \quad (2\text{-}34)$$

where, for convenience, we take the norm $\|\cdot\|$ over the space R^n in (2-34) to be the usual maximum norm:

$$\|\xi\| = \max\{|\xi_1|, |\xi_2|, \ldots, |\xi_n|\}$$

For any point $\xi = (\xi_1, \xi_2, \ldots, \xi_n) \in R^n$, however, the following relation exists between the Fréchet derivative of a function u and its usual partial derivatives:

$$\mathcal{D}u(\alpha).\xi = \sum_{i=1}^{n} \xi_i D_i u(\alpha) \qquad \ldots \quad (2\text{-}35)$$

where $D_i u(\alpha)$, $1 \le i \le n$, denotes the usual partial derivative of the first order of the function u in the direction of the i-th co-ordinate. Moreover, if the space R^n is equipped with its canonical basis $(e_1, e_2, \ldots, e_n)$, we have:

$$\mathcal{D}u(\alpha).(e_i) = D_i u(\alpha), \quad 1 \le i \le n \qquad \ldots \quad (2\text{-}36)$$

Then, from (2-36), we get:

$$\|\mathcal{D}u(\alpha).(e_i)\| = |D_i u(\alpha)| \le \|\mathcal{D}u(\alpha)\|.\|e_i\|$$

or

$$\|\mathcal{D}u(\alpha)\| \ge |D_i u(\alpha)|, \quad 1 \le i \le n \qquad \ldots \quad (2\text{-}37)$$

On the other hand, from (2-35), we have:

$$|\mathcal{D}u(\alpha).\xi| \le \sum_{i=1}^{n} |\xi_i| \, |D_i u(\alpha)| \le \|\xi\| \sum_{i=1}^{n} |D_i u(\alpha)|$$

or

$$\frac{|\mathcal{D}u(\alpha).\xi|}{\|\xi\|} \leq \sum_{i=1}^{n} |D_i u(\alpha)|$$

and, by taking the supremum over the space R^n, we obtain:

$$\|\mathcal{D}u(\alpha)\| \leq \sum_{i=1}^{n} |D_i u(\alpha)| \leq C_1(n) \max_{1\leq i\leq n} |D_i u(\alpha)| \quad . \ . \ . \ (2\text{-}38)$$

where, of course, the constant $C_1(n) = n$. Thus, from inequalities (2-37) and (2-38), we get the following useful double inequality:

$$|D_i u(\alpha)| \leq \|\mathcal{D}u(\alpha)\| \leq C_1(n) \max_{1\leq i\leq n} |D_i u(\alpha)|$$

Likewise, the k-th order Fréchet derivative of the function u at a point $\alpha \in R^n$ is defined as the unique linear transformation of the Cartesian product space $R^n \times R^n \times \ldots \times R^n$ (k times) into R, ie,

$$\mathcal{D}^k u(\alpha) \in L_k(R^n;R) \ : \ (R^n)^k \to R$$

such that its application to a point $(\xi^1,\xi^2,\ldots,\xi^k) \in (R^n)^k$, with $\xi^i \in R^n$, $1 \leq i \leq k$, gives the following real number:

$$\mathcal{D}^k u(\alpha).(\xi^1,\xi^2,\ldots,\xi^k)$$

with the convention:

$$\mathcal{D}^k u(\alpha).(\xi)^k$$

whenever $\xi^i = \xi$, $1 \leq i \leq k$.

Again, the norm of the operator $\mathcal{D}^k u(\alpha)$ is defined by the formula:

$$\| D^k u(\alpha) \| = \sup_{\substack{\| \xi^i \| \leq 1 \\ 1 \leq i \leq n}} | D^k u(\alpha) . (\xi^1, \xi^2, \ldots, \xi^k) | \quad \ldots (2\text{-}39)$$

or, alternatively, by:

$$\| D^k u(\alpha) \| = \sup_{\substack{\xi \in (R^n)^k \\ \xi \neq 0}} \frac{| D^k u(\alpha) . (\xi^1, \xi^2, \ldots, \xi^k) |}{\| (\xi^1, \xi^2, \ldots, \xi^k) \|} \quad \ldots (2\text{-}40)$$

Then, if the vectors $\xi^i \in R^n$, $1 \leq i \leq k$, have components $(\xi_1^i, \xi_2^i, \ldots, \xi_n^i)$, $1 \leq i \leq k$, with respect to some co-ordinate system in R^n, the following formula holds, corresponding to that defined by (2-35):

$$D^k u(\alpha) . (\xi^1, \xi^2, \ldots, \xi^k) = \sum_{\lambda_1=1}^{n} \sum_{\lambda_2=1}^{n} \cdots \sum_{\lambda_k=1}^{n} \xi_{\lambda_1}^1 \xi_{\lambda_2}^2 \cdots \xi_{\lambda_k}^k D_{\lambda_1} D_{\lambda_2} \cdots D_{\lambda_k} u(\alpha) \quad \ldots (2\text{-}41)$$

Furthermore, if the element $(\xi^1, \xi^2, \ldots, \xi^k) \in (R^n)^k$ is such that:

$$\| \xi^i \| \leq 1, \; 1 \leq i \leq k$$

then,

$$D^k u(\alpha) . (\xi^1, \xi^2, \ldots, \xi^k) = D^r u(\alpha) \quad \ldots (2\text{-}42)$$

where $r = (r_1, r_2, \ldots, r_n)$ and $|r| = \sum_{i=1}^{n} r_i = k$.
From (2-42) now, we get:

$$| D^r u(\alpha) | = \| D^k u(\alpha) . (\xi^1, \xi^2, \ldots, \xi^k) \| \leq \| D^k u(\alpha) \| . \| \xi \|$$

where:

$$\| \xi \| = \max_{1 \leq i \leq n} \| \xi^i \|$$

or, since $\|\xi^i\| \leq 1$, $1 \leq i \leq n$, we obtain:

$$|D^r u(a)| \leq \|D^k u(a)\|, \quad |r| = k \qquad \ldots (2\text{-}43)$$

On the other hand, from (2-41), we have:

$$|D^k u(a).(\xi^1,\xi^2,\ldots,\xi^k)| \leq$$

$$\leq \sum_{\lambda_1=1}^{n} \sum_{\lambda_2=1}^{n} \cdots \sum_{\lambda_k=1}^{n} |\xi^1_{\lambda_1}\xi^2_{\lambda_2}\cdots\xi^k_{\lambda_k}| \cdot |D_{\lambda_1}D_{\lambda_2}\cdots D_{\lambda_k}u(a)| \leq$$

$$\leq \|(\xi^1,\xi^2,\ldots,\xi^k)\| \sum_{\lambda_1=1}^{n} \sum_{\lambda_2=1}^{n} \cdots \sum_{\lambda_k=1}^{n} |D_{\lambda_1}D_{\lambda_2}\cdots D_{\lambda_k}u(a)| \qquad \ldots (2\text{-}44)$$

where by $\|(\xi^1,\xi^2,\ldots,\xi^k)\|$ we define the following maximum norm over the Cartesian product space $(\mathbb{R}^n)^k$:

$\|(\xi^1,\xi^2,\ldots,\xi^k)\| = \max\{|\xi^1_{\sigma\lambda_1}\xi^2_{\sigma\lambda_2}\cdots\xi^k_{\sigma\lambda_k}|$, for all the possible permutations $\sigma : \lambda_i \to \sigma\lambda_i$ of the set of indices $\{1,2,\ldots,n\}$, with $1 \leq i \leq k\}$.

Therefore, from (2-44), we have:

$$\frac{|D^k u(a).(\xi^1,\xi^2\ldots\xi^k)|}{\|(\xi^1,\xi^2,\ldots,\xi^k)\|} \leq \sum_{\lambda_1=1}^{n} \sum_{\lambda_2=1}^{n} \cdots \sum_{\lambda_k=1}^{n} |D_{\lambda_1}D_{\lambda_2}\cdots D_{\lambda_k}u(a)| \qquad \ldots (2\text{-}45)$$

However, since:

$$\sum_{\lambda_1=1}^{n} \sum_{\lambda_2=1}^{n} \cdots \sum_{\lambda_k=1}^{n} |D_{\lambda_1}D_{\lambda_2}\cdots D_{\lambda_k}u(a)| = \sum_{|\mu|=k} \frac{k!}{\mu_1!\mu_2!\ldots\mu_n!}|D^\mu u(a)| \leq$$

$$\leq \sum_{|\mu|=k} \left(\frac{k!}{\mu_1!\mu_2!\ldots\mu_n!}\right) \max_{|\mu|=k} |D^\mu u(a)|$$

or s.

$$\sum_{|\mu|=k} \frac{k!}{\mu_1!\mu_2!\cdots\mu_n!} = n^k$$

we have:

$$\sum_{\lambda_1=1}^{n}\sum_{\lambda_2=1}^{n}\cdots\sum_{\lambda_k=1}^{n} |D_{\lambda_1}D_{\lambda_2}\cdots D_{\lambda_k}u(\alpha)| \le n^k \max_{|\mu|=k} |D^\mu u(\alpha)| \quad \ldots (2\text{-}46)$$

Thus, by taking the supremum of (2-45) over the space R^n and combining the result with (2-40) and (2-46), we obtain:

$$\|D^k u(\alpha)\| \le C_k(n) \max_{|\mu|=k} |D^\mu u(\alpha)| \quad \ldots (2\text{-}47)$$

where the constant $C_k(n) = n^k$.

Finally, from the inequalities (2-43) and (2-47), we get:

$$|L^{\cdots}u| \le \|D^k u(\alpha)\| \le C_k(n) \max_{|\mu|=k} |D^\mu u(\alpha)| \quad \ldots (2\text{-}48)$$

where $r = (r_1, r_2, \ldots, r_n)$ and $|r| = k$. The double inequality (2-48) is very useful in relating the Fréchet derivatives to the usual composite partial derivatives and vice versa.

2.2.2 Multivariate Hermite Interpolation

Given an n-simplex $\bar{S}_n^\nu$, $\nu \in I$, defined through a partition Δ of the given polyhedral domain Ω, we begin our analysis on approximation with the general multivariate Hermite

approximation problem, the final goal being that of giving an estimate for the error in:

$$\max_{\substack{x\in\bar{S}_n^\nu \\ \nu\in I \\ |\alpha|=m}} |D^\alpha u(x) - D^\alpha u_n^{(k-1)}(x)|, \text{ for any integer } m \text{ with } 0\le m\le k-1, \quad \ldots \quad (2\text{-}49)$$

where the function $u(x) \in C^k(\bar{S}_n^\nu)$ and $u_n^{(k-1)}(x) \in P_{k-1}(\bar{S}_n^\nu)$ is its unique Hermite interpolating polynomial of general degree (k-1). It is essential that the function u(x) is well-defined on a discrete set S of interpolating points, always associated with the particular simplex $\bar{S}_n^\nu$, $\nu \in I$, under consideration, which can be defined as follows: Since the Hermite polynomial $u_n^{(k-1)}(x)$ interpolates, not only the values of the function u(x) at the points $\alpha_i^\nu \in S_n^\nu(k-1)$, $\nu \in I$, but also the values of some of its derivatives at several - not necessarily all - points of the same set as well, let us introduce one more superscript and denote by:

$$S_n^{\nu,0}(k-1) \equiv S_n^\nu(k- \qquad I$$

the usual (k-1)-st order principal lattice associated to the simplex $\bar{S}_n^\nu$, $\nu \in I$, which is defined by the formula (2-26). It contains exactly:

$$No = \binom{k+n-1}{k-1}$$

points. Furthermore, denote by:

$$S_n^{\nu,\mu}(k-1),\ 1\le\mu\le\lambda,\ 1\le\lambda\le k-1,\ \nu\in I$$

the λ sets of points which in almost all practical applications constitute only subsets of the original set $S_n^{\nu,0}(k-1)$, $\nu \in I$. For any $1 \le \mu \le \lambda$ the set $S_n^{\nu,\mu}(k-1)$, $\nu \in I$, contains those points of the set $S_n^{\nu,0}(k-1)$, $\nu \in I$, on which the values of the partial derivatives of order μ of the function $u(x)$ are interpolated. Denote the number of those points by:

$$N_\mu, \quad 1 \le \mu \le \lambda$$

Then, the set S of interpolating points, mentioned above, is defined by the following set theoretic union:

$$S = \bigcup_{\mu=0}^{\lambda} S_n^{\nu,\mu}(k-1), \quad \nu \in I \qquad \ldots \quad (2\text{-}50)$$

Following the same argument as P.G. Ciarlet and P.A. Raviart in [7], the general Hermite interpolation problem is defined as follows:

With every point:

$$a_i^{\nu,\mu} \in S_n^{\nu,\mu}(k-1), \quad 1 \le i \le N_\mu, \quad 0 \le \mu \le \lambda, \quad \nu \in I$$

we associate a subspace:

$$A_i^{\nu,\mu} \subset (R^n)^\mu, \quad 1 \le i \le N_\mu, \quad 0 \le \mu \le \lambda, \quad \nu \in I$$

with the convention that:

$$A_i^{\nu,0} \equiv R, \quad 1 \le i \le N_0, \text{ whenever } \mu = 0.$$

Then, we say, by definition, that the set (2-50) constitutes a (k-1)-unisolvent set, if and only if, given a set of linear transformations:

$$R_i^\mu \in L_\mu(R^n;R),\ 1 \le i \le N_\mu,\ 0 \le \mu \le \lambda$$

with the convention that:

$$R_i^0 \in L_0(R^n;R) \equiv R,\ 1 \le i \le N_0$$

there exists one and only one polynomial $p(x)$ of degree $(k-1)$ such that:

$$D^\mu\, p(\alpha_i^{\nu,\mu})\cdot(\xi^1,\xi^2,\ldots,\xi^\mu) = R_i^\mu\cdot(\xi^1,\xi^2,\ldots,\xi^\mu) \qquad \ldots\ (2\text{-}51)$$

for all

$$(\xi^1,\xi^2,\ldots,\xi^\mu) \in A_i^{\nu,\mu},\ 1 \le i \le N_\mu,\ 0 \le \mu \le \lambda,\ \nu \in I$$

Therefore, according to the above definition, the Hermite interpolating polynomial:

$$u_n^{(k-1)}(x) \in P_{k-1}(\tilde{S}_n^\nu),\ \nu \in I$$

of degree $(k-1)$ is the unique polynomial of that degree which satisfies the following interpolating constraints - in analogy to those defined by (2-51):

$$D^\mu u(\alpha_i^{\nu,\mu})\cdot(\xi^1,\xi^2,\ldots,\xi^\mu) = D^\mu u_n^{(k-1)}(\alpha_i^{\nu,\mu})\cdot(\xi^1,\xi^2,\ldots,\xi^\mu) \qquad \ldots\ (2\text{-}52)$$

where $(\xi^1,\xi^2,\ldots,\xi^\mu) \in A_i^{\nu,\mu} \subset (R^n)^\mu$, $1 \le i \le N_\mu$, $0 \le \mu \le \lambda$, $\nu \in I$, and with the convention that:

$$u(\alpha_i^{\nu,0}) = u_n^{(k-1)}(\alpha_i^{\nu,0}),\ 1 \le i \le N_0,\ \text{whenever } \mu = 0.$$

The introduction of the subspaces:

$$A_i^{\nu,\mu} \subset (R^n)^\mu,\ 1 \le \mu \le \lambda,\ \nu \in I$$

instead of the entire Euclidean-product spaces:

$$(R^n)^\mu,\ 1 \le \mu \le \lambda$$

becomes a necessity to cope with the situation which arises in the majority of the practical applications. Indeed, in many practical cases only a specific kind of partial derivatives has to be interpolated - eg the values of the derivatives u_{xx} and u_{yy}, but not that of the cross derivative u_{xy} - and, thus, the introduction of those subsets is completely warranted.

Suppose now that for every such subspace $A_i^{\nu,\mu}$, $1 \le i \le N_\mu$, $1 \le \mu \le \lambda$, $\nu \in I$, the following vectors:

$$\xi_{i1}^{\nu,\mu}, \xi_{i2}^{\nu,\mu}, \ldots, \xi_{i\nu_i}^{\nu,\mu}$$

form a basis. Then, the Hermite interpolating polynomial of degree (k-1) is given by the formula:

$$u_n^{(k-1)}(x) = \sum_{i=1}^{N_0} u(a_i^{\nu,0}) \phi_i^{\nu,0}(x) + \sum_{i=1}^{N_1} \sum_{\ell=1}^{\nu_i} [Du(a_i^{\nu,1}) \cdot (\xi_{i\ell}^{\nu,1})] \phi_{i\ell}^{\nu,1}(x) + \ldots$$

$$+ \sum_{i=1}^{N_\lambda} \sum_{\ell=1}^{\nu_i} [D^\lambda u(a_i^{\nu,\lambda}) \cdot (\xi_{i\ell}^{\nu,\lambda})] \phi_{i\ell}^{\nu,\lambda}(x) \qquad \ldots \quad (2\text{-}53)$$

with:

$$u_n^{(k-1)}(x) \in P_{k-1}(\bar{S}_n^\nu),\ \nu \in I$$

and

$$N = \dim P_{k-1}(\bar{S}_n^\nu) = N_0 + \sum_{\mu=1}^{\lambda} \sum_{i=1}^{N_\mu} \nu_i$$

Some of the interpolating points of the set S, defined by (2-50), constitute multiple nodes, by definition of the Hermite problem, and as far as the basis functions of the formula (2-53) are concerned, they satisfy the following conditions:

$$\phi_i^{\nu,0}(x) \in P_{k-1}(\bar{S}_n^{\nu}),\ 1 \le i \le N_0,\ \nu \in I$$

$$\phi_i^{\nu,0}(\alpha_j^{\nu,0}) = \delta_{ij},\ 1 \le i,j \le N_0,\ \nu \in I$$

$$D^{\mu}\phi_i^{\nu,0}(\alpha_j^{\nu,\mu}) \cdot (\xi_{i\ell}^{\nu,\mu}) = 0,\ 1 \le i \le N_0,\ 1 \le j \le N_{\mu},\ 1 \le \ell \le \nu_j,\ 1 \le \mu \le \lambda,\ \nu \in I \qquad (2\text{-}54)$$

$$\phi_{i\ell}^{\nu,\mu}(x) \in P_{k-1}(\bar{S}_n^{\nu}),\ 1 \le i \le N_{\mu},\ 1 \le \ell \le \nu_i,\ 1 \le \mu \le \lambda,\ \nu \in I$$

$$\phi_{i\ell}^{\nu,\mu}(\alpha_j^{\nu,0}) = 0,\ 1 \le i \le N_{\mu},\ 1 \le j \le N_0,\ 1 \le \ell \le \nu_i,\ 1 \le \mu \le \lambda,\ \nu \in I$$

$$D^{\mu}\phi_{i\ell}^{\nu,\mu}(\alpha_j^{\nu,\mu}) \cdot (\xi_{j\rho}^{\mu}) = \delta_{ij}\delta_{\ell\rho},\ 1 \le i,j \le N_{\mu},\ 1 \le \ell,\rho \le \nu_i,\ 1 \le \mu \le \lambda,\ \nu \in I.$$

However, as for the one- and two-variable approximation problems where the Taylor's formula plays a significant role in deriving several upper error bounds between the function and its interpolating polynomial, exactly the same applies with the general multivariate approximation problem and, therefore, it makes sense to try and get first the appropriate such multivariate formula. Thus, for a function f(x) which assumes a C^k order of continuity over an interval domain I, we can write the following Taylor's expansion formula with an integral remainder:

$$f(t) = \sum_{j=0}^{k-1} \frac{(t-\alpha)^j}{j!} f^{(j)}(\alpha) + \int_{\alpha}^{t} \frac{(t-\tau)^{k-1}}{(k-1)!} f^{(k)}(\tau)d\tau \qquad \ldots (2\text{-}55)$$

where t and α are two distinct points of the interval I and $\tau \in [\alpha,t]$. For the particular points $t=1$ and $\alpha=0$, the formula (2-55) gives:

$$f(1) = \sum_{j=0}^{k-1} \frac{1}{j!} f^{(j)}(0) + \int_{0}^{1} \frac{(1-\tau)^{k-1}}{(k-1)!} f^{(k)}(\tau)d\tau, \ \tau \in [0,1] \qquad \ldots (2\text{-}56)$$

Then, for any point:

$$\alpha_i^{\nu,0} \in S_n^{\nu,0}(k-1), \ 1 \le i \le N_0, \ \nu \in I \qquad \ldots (2\text{-}57)$$

and any point:

$$x \in \bar{S}_n^{\nu} - S_n^{\nu,0}(k-1), \ \nu \in I \qquad \ldots (2\text{-}58)$$

we have:

$$x + t(\alpha_i^{\nu,0} - x) \in \bar{S}_n^{\nu}, \ 1 \le i \le N_0, \ \nu \in I, \ t \in [0,1]$$

since the simplex $\bar{S}_n^{\nu}$, $\nu \in I$, constitutes a convex region. Next, define the following multivariate function:

$$f(t) = u(x+t(\alpha_i^{\nu,0} - x)) \qquad \ldots (2\text{-}59)$$

with the points $\alpha_i^{\nu,0}$, $1 \le i \le N_0$, $\nu \in I$, and x defined as in (2-57) and (2-58) respectively. Then, by differentiating the function (2-59) j-times, we get:

$$\frac{d^j f(t)}{dt^j} = \frac{d^j[u(x+t(\alpha_i^{\nu,0} - x))]}{dt^j} = \sum_{|r|=j} (\alpha_i^{\nu,0} - x)_1^{r_1} \cdot (\alpha_i^{\nu,0} - x)_2^{r_2} \ldots (\alpha_i^{\nu,0} - x)_n^{r_n} D^r u(x+t(\alpha_i^{\nu,0} - x)),$$

for all the integers j such that $0 \le j \le k-1$. However, since:

$$\sum_{|r|=j} (\alpha_i^{\nu,0}-x)_1^{r_1}\cdot(\alpha_i^{\nu,0}-x)_2^{r_2}\cdots(\alpha_i^{\nu,0}-x)_n^{r_n} D^r u(x+t(\alpha_i^{\nu,0}-x)) =$$

$$= D^j u(x+t(\alpha_i^{\nu,0}-x))\cdot(\alpha_i^{\nu,0}-x)^j$$

we get the following result:

$$\frac{d^j f(t)}{dt^j} = D^j u(x+t(\alpha_i^{\nu,0}-x))\cdot(\alpha_i^{\nu,0}-x)^j, \text{ for all } 0 \le j \le k-1 \quad \ldots \quad (2\text{-}60)$$

Finally, from (2-56) combined with (2-60), we obtain the following multivariate Taylor formula:

$$u(\alpha_i^{\nu,0}) = \sum_{j=1}^{k-1} \frac{1}{j!} D^j u(x)\cdot(\alpha_i^{\nu,0}-x)^j + \int_0^1 \frac{(1-t)^{k-1}}{(k-1)!} D^k u(x+t(\alpha_i^{\nu,0}-x))\cdot(\alpha_i^{\nu,0}-x)\,dt$$

$$\ldots \quad (2\text{-}61)$$

for any point $\alpha_i^{\nu,0} \in S_n^{\nu,0}(k-1)$, $1 \le i \le N_0$, $\nu \in I$, and any point $x \in \bar{S}_n^{\nu} - S_n^{\nu,0}(k-1)$, $\nu \in I$, $t \in [0,1]$.

We shall use the formula (2-61) in order to find an estimate for the difference:

$$D^m u(x) - D^m u_n^{(k-1)}(x), \text{ for any } 0 \le m \le k-1$$

Then, by using the inequalities (2-48), we can easily obtain an equivalent estimate for the difference:

$$D^\alpha u(x) - D^\alpha u_n^{(k-1)}(x), \quad |\alpha| = m, \ 0 \le m \le k-1.$$

involving, this time, the usual composite partial derivatives.

Consider now a function:

$$u(x) \in C^k(\bar{S}_n^\nu), \quad \nu \in I$$

and its unique Hermite interpolating polynomial defined by the formula (2-53). Then, for any point:

$$x \in S_n^\nu - S_n^{\nu,0}(k-1), \quad \nu \in I$$

where S_n^ν denotes the interior of the n-simplex $\bar{S}_n^\nu$, $\nu \in I$, and any integer m with $0 \le m \le k-1$, we shall first prove the following important result:

$$D^m u_n^{(k-1)}(x) = D^m u(x) + \sum_{i=1}^{N_0} \left[\int_0^1 \frac{(1-t)^{k-1}}{(k-1)!} D^k u(x+t(a_i^{\nu,0}-x)).(a_i^{\nu,0}-x)^k dt \right] . D^m \phi_i^{\nu,0}(x) +$$

$$+ \sum_{i=1}^{N_1} \sum_{\ell=1}^{\nu_i} \left[\int_0^1 \frac{(1-t)^{k-2}}{(k-2)!} D^k u(x+t(a_i^{\nu,1}-x)).(\xi_{i\ell}^{\nu,1},(a_i^{\nu,1}-x)^{k-1}) dt \right] . D^m \phi_{i\ell}^{\nu,1}(x) + \ldots$$

$$+ \sum_{i=1}^{N_\lambda} \sum_{\ell=1}^{\nu_i} \left[\int_0^1 \frac{(1-t)^{k-\lambda-1}}{(k-\lambda-1)!} D^k u(x+t(a_i^{\nu,\lambda}-x)).(\xi_{i\ell}^{\nu,\lambda},(a_i^{\nu,\lambda}-x)^{k-\lambda}) dt \right] . D^m \phi_{i\ell}^{\nu,\lambda}(x)$$

. . . (2-62)

Indeed, from the formula (2-53) and for any integer m with $0 \le m \le k-1$, ν have:

$$D^m u_n^{(k-1)}(x) = \sum_{i=1}^{N_0} u(a_i^{\nu,0}) D^m \phi_i^{\nu,0}(x) + \sum_{i=1}^{N_1} \sum_{\ell=1}^{\nu_i} [Du(a_i^{\nu,1}).(\xi_{i\ell}^{\nu,1})] . D^m \phi_{i\ell}^{\nu,1}(x) + \ldots$$

$$\ldots + \sum_{i=1}^{N_\lambda} \sum_{\ell=1}^{\nu_i} [D^\lambda u(a_i^{\nu,\lambda}).(\xi_{i\ell}^{\nu,\lambda})] . D^m \phi_{i\ell}^{\nu,\lambda}(x) \quad . . . \quad (2-63)$$

Furthermore, from (2-61) and for any point $a_i^{\nu,\mu} \in S$, $1 \le i \le N_\mu$, $0 \le \mu \le \lambda$, $\nu \in I$, we have:

$$u(a_i^{\nu,\mu}) = \sum_{j=0}^{k-1} \frac{1}{j!} D^j u(x)\cdot(a_i^{\nu,\mu}-x)^j + \int_0^1 \frac{(1-t)^{k-1}}{(k-1)!} D^k u(x+t(a_i^{\nu,\mu}-x))\cdot(a_i^{\nu,\mu}-x)^k dt$$

where $t \in [0,1]$, $1 \le i \le N_\mu$, $0 \le \mu \le \lambda$, $\nu \in I$, and from which we obtain:

$$u(a_i^{\nu,0}) = \sum_{j=0}^{k-1} \frac{1}{j!} D^j u(x)\cdot(a_i^{\nu,0}-x)^j + \int_0^1 \frac{(1-t)^{k-1}}{(k-1)!} D^k u(x+t(a_i^{\nu,0}-x))\cdot(a_i^{\nu,0}-x)^k dt$$

where $t \equiv t_i^{\nu,0}$, $t_i^{\nu,0} \in [0,1]$, $1 \le i \le N_0$, $\nu \in I$

$$Du(a_i^{\nu,1})\cdot(\xi_{i\ell}^{\nu,1}) = Du(x)\cdot(\xi_{i\ell}^{\nu,1}) + D^2u(x)\cdot(\xi_{i\ell}^{\nu,1},(a_i^{\nu,1}-x)) + \ldots$$

$$\ldots + \frac{1}{(k-2)!} D^{k-1}u(x)\cdot(\xi_{i\ell}^{\nu,1},(a_i^{\nu,1}-x)^{k-2}) + \int_0^1 \frac{(1-t)^{k-2}}{(k-2)!} D^k u(x+t(a_i^{\nu,1}-x))\cdot$$

$$\cdot(\xi_{i\ell}^{\nu,1},(a_i^{\nu,1}-x)^{k-1})dt = \sum_{j=1}^{k-1} \frac{1}{(j-1)!} D^j u(x)\cdot(\xi_{i\ell}^{\nu,1},(a_i^{\nu,1}-x)^{j-1}) +$$

$$+ \int_0^1 \frac{(1-t)^{k-2}}{(k-2)!} D^k u(x+t(a_i^{\nu,1}-x))\cdot(\xi_{i\ell}^{\nu,1},(a_i^{\nu,1}-x)^{k-1})dt$$

where $t \equiv t_{i\ell}^{\nu,1}$, $t_{i\ell}^{\nu,1} \in [0,1]$, $1 \le i \le N_1$, $1 \le \ell \le \nu_i$, $\nu \in I$.

etc ... , and:

$$D^m u(a_i^{\nu,\lambda})\cdot(\xi_{i\ell}^{\nu,\lambda}) = D^\lambda u(x)\cdot(\xi_{i\ell}^{\nu,\lambda}) + D^{\lambda+1}u(x)\cdot(\xi_{i\ell}^{\nu,\lambda},(a_i^{\nu,\lambda}-x)) + \ldots$$

$$\ldots + \frac{1}{(k-\lambda-1)!} D^{k-1}u(x)\cdot(\xi_{i\ell}^{\nu,\lambda},(a_i^{\nu,\lambda}-x)^{k-\lambda-1}) +$$

$$+ \int_0^1 \frac{(1-t)^{k-\lambda-1}}{(k-\lambda-1)!} D^k u(x+t(a_i^{\nu,\lambda}-x))\cdot(\xi_{i\ell}^{\nu,\lambda},(a_i^{\nu,\lambda}-x)^{k-\lambda})dt = \sum_{j=\lambda}^{k-1} \frac{1}{(j-\lambda)!} D^\lambda u(x)\cdot$$

$$\cdot(\xi_{i\ell}^{\nu,\lambda},(a_i^{\nu,\lambda}-x)^{j-\lambda}) + \int_0^1 \frac{(1-t)^{k-\lambda-1}}{(k-\lambda-1)!} D^k u(x+t(a_i^{\nu,\lambda}-x))\cdot(\xi_{i\ell}^{\nu,\lambda},(a_i^{\nu,\lambda}-x)^{k-\lambda})dt$$

where $t \equiv t_{i\ell}^{\nu,\lambda}$, $t_{i\ell}^{\nu,\lambda} \in [0,1]$, $1 \le i \le N_\lambda$, $1 \le \ell \le \nu_i$, $\nu \in I$.

By substituting these results into the formula (2-63), we get:

$$D^m u_n^{(k-1)}(x) = \sum_{j=0}^{k-1} \frac{1}{j!} \sum_{i=1}^{N_0} [D^j u(x)\cdot(\alpha_i^{\nu,0}-x)^j] D^m \phi_i^{\nu,0}(x) +$$

$$+ \sum_{j=1}^{k-1} \frac{1}{(j-1)!} \sum_{i=1}^{N_1} \sum_{\ell=1}^{\nu_i} [D^j u(x)\cdot(\xi_{i\ell}^{\nu,1},(\alpha_i^{\nu,1}-x)^{j-1}) D^m \phi_{i\ell}^{\nu,1}(x) + \ldots$$

$$\ldots + \sum_{j=\lambda}^{k-1} \frac{1}{(j-\lambda)!} \sum_{i=1}^{N_\lambda} \sum_{\ell=1}^{\nu_i} [D^j u(x)\cdot(\xi_{i\ell}^{\nu,\lambda},(\alpha_i^{\nu,\lambda}-x)^{j-\lambda}) D^m \phi_{i\ell}^{\nu,\lambda}(x) +$$

$$+ \sum_{i=1}^{N_0} [\int_0^1 \frac{(1-t)^{k-1}}{(k-1)!} D^k u(x+t(\alpha_i^{\nu,0}-x)).(\alpha_i^{\nu,0}-x)^k dt] D^m \phi_i^{\nu,0}(x) +$$

$$+ \sum_{i=1}^{N_1} \sum_{\ell=1}^{\nu_i} [\int_0^1 \frac{(1-t)^{k-2}}{(k-2)!} D^k u(x+t(\alpha_i^{\nu,1}-x)).(\xi_{i\ell}^{\nu,1},(\alpha_i^{\nu,1}-x)^{k-1}) dt] D^m \phi_{i\ell}^{\nu,1}(x) + \ldots$$

$$\ldots + \sum_{i=1}^{N_\lambda} \sum_{\ell=1}^{\nu_i} [\int_0^1 \frac{(1-t)^{k-\lambda-1}}{(k-\lambda-1)!} D^k u(x+t(\alpha_i^{\nu,\lambda}-x)).(\xi_{i\ell}^{\nu,\lambda},(\alpha_i^{\nu,\lambda}-x)^{k-\lambda}) dt] D^m \phi_{i\ell}^{\nu,\lambda}(x)$$

. . . (2-64)

Therefore, the crucial point of proving the formula (2-62) is to show, using the property:

$$u(x) = u_n^{(k-1)}(x), \text{ whenever } u(x) \in P_{k-1}(\bar{S}_n^\nu), \ \nu \in I \qquad . . . \ (2\text{-}65)$$

that:

$$\frac{1}{j!} \sum_{i=1}^{N_0} [D^j u(x)\cdot(\alpha_i^{\nu,0}-x)^j] D^m \phi_i^{\nu,0}(x) + \ldots$$

$$\ldots + \frac{1}{(j-\lambda)!} \sum_{i=1}^{N_\lambda} \sum_{\ell=1}^{\nu_i} [D^j u(x)\cdot(\xi_{i\ell}^{\nu,\lambda},(\alpha_i^{\nu,\lambda}-x)^{j-\lambda}) D^m \phi_{i\ell}^{\nu,\lambda}(x) = \begin{cases} D^m u(x), & \text{for } j=m \\ 0, & \text{for } j \neq m \end{cases}$$

. . . (2-66)

Indeed, for any point $x \in S_n^\nu - S$, $\nu \in I$, we have:

$$\mathcal{D}^j u(x) \in L_j(R^n;R)$$

Thus, for any integer j such that $0 \leq j \leq k-1$,

$$\mathcal{D}^j u(x)\cdot(\alpha-x)^j \qquad \ldots (2\text{-}67)$$

is a polynomial in α of degree less than or equal to $(k-i)$, where α here denotes any point of the set S defined by (2-50). Therefore,

$$\mathcal{D}^j u(x)\cdot(\alpha-x)^j \in P_{k-1}(\bar{S}_n^\nu), \quad \nu \in I$$

and since the Hermite interpolation can also be defined as being the projection of the continuous space:

$$C^\lambda(\bar{S}_n^\nu), \quad \nu \in I$$

into the space:

$$P_{k-1}(\bar{S}_n^\nu), \quad \nu \in I$$

where always:

$$P_{k-1}(\bar{S}_n^\nu) \subset C^\lambda(\bar{S}_n^\nu), \quad \nu \in I$$

the polynomial (2-67) coincides with its Hermite interpolating polynomial defined through the formula (2-53), ie

$$\sum_{i=1}^{N_0} \mathcal{D}^j u(x)\cdot(\alpha_i^\nu - x)^j \phi_i^{\nu,0}(\alpha) + \ldots + \sum_{i=1}^{N_\lambda} \sum_{\ell=1}^{\nu_i} \{\mathcal{D}^\lambda[\mathcal{D}^j u(x)\cdot(\xi_{i\ell}^{\nu,\lambda} - x)^j]\}\, \phi_{i\ell}^{\nu,\lambda}(\alpha) =$$

$$= \mathcal{D}^j u(x)\cdot(\alpha-x)^j$$

or

$$\sum_{i=1}^{N_0} D^j u(x)\cdot(\alpha_i^{\nu,0}-x)^j \phi_i^{\nu,0}(\alpha) + \ldots + [j(j-1)\ldots(j-\lambda+1)]\cdot$$

$$\cdot \sum_{i=1}^{N_\lambda}\sum_{k=1}^{\nu_i} [D^j u(x)\cdot(\xi_{ik}^{\nu,\lambda},(\alpha_i^{\nu,\lambda}-x)^{j-\lambda})]\phi_{ik}^{\nu,\lambda}(\alpha) = D^j u(x)\cdot(\alpha-x)^j$$

Then, for any integer m, with $0 \le m \le k-1$, we have:

$$\sum_{i=1}^{N_0} [D^j u(x)\cdot(\alpha_i^{\nu,0}-x)^j] D^m \phi_i^{\nu,0}(\alpha) + \ldots$$

$$\ldots + [j(j-1)\ldots(j-\lambda+1)] \sum_{i=1}^{N_\lambda}\sum_{k=1}^{\nu_i} [D^j u(x)\cdot(\xi_{ik}^{\nu,\lambda},(\alpha_i^{\nu,\lambda}-x)^{j-\lambda})]\, D^m \phi_{ik}^{\nu,\lambda}(\alpha) =$$

$$= [j(j-1)\ldots(j-m+1)] D^j u(x)\cdot(\alpha-x)^{j-m} \qquad \ldots (2\text{-}68)$$

From (2-68) now, and for $\alpha = x$, we have:

$$\sum_{i=1}^{N_0} [D^j u(x)\cdot(\alpha_i^{\nu,0}-x)^j] D^m \phi_i^{\nu,0}(x) + \ldots + [j(j-1)\ldots(j-\lambda+1)]\cdot$$

$$\cdot \sum_{i=1}^{N_\lambda}\sum_{k=1}^{\nu_i} [D^j u(x)\cdot(\xi_{ik}^{\nu,\lambda},(\alpha_i^{\nu,\lambda}-x)^{j-\lambda})] D^m \phi_{ik}^{\nu,\lambda}(x) = j! D^m u(x)$$

when and only when $j = m$. Thus,

$$\frac{1}{j!}\sum_{i=1}^{N_0} [D^j u(x)\cdot(\alpha_i^{\nu,0}-x)^j] D^m \phi_i^{\nu,0}(x) + \ldots + \frac{j(j-1)\ldots(j-\lambda+1)}{j!}\cdot$$

$$\cdot \sum_{i=1}^{N_\lambda}\sum_{k=1}^{\nu_i} [D^j u(x)\cdot(\xi_{ik}^{\nu,\lambda},(\alpha_i^{\nu,\lambda}-x)^{j-\lambda})] D^m \phi_{ik}^{\nu,\lambda}(x) =$$

$$= \frac{1}{j!} \sum_{i=1}^{N_0} [D^j u(x) \cdot (a_i^{\nu,0} - x)^j] D^m \phi_i^{\nu,0}(x) + \ldots$$

$$\ldots + \frac{1}{(j-\lambda)!} \sum_{i=1}^{N_\lambda} \sum_{\ell=1}^{\nu_i} [D^j u(x) \cdot (\xi_{i\ell}^{\nu,\lambda}, (a_i^{\nu,\lambda} - x)^{j-\lambda})] D^m \phi_{i\ell}^{\nu,\lambda}(x) = D^m u(x).$$

when and only when $j = m$. For $j \neq m$ it can easily be seen, from (2-68), that:

$$\frac{1}{j!} \sum_{i=1}^{N_0} [D^j u(x) \cdot (a_i^{\nu,0} - x)^j] D^m \phi_i^{\nu,0}(x) + \ldots$$

$$\ldots + \frac{1}{(j-\lambda)!} \sum_{i=1}^{N_\lambda} \sum_{k=1}^{\nu_i} [D^j u(x) \cdot (\xi_{i\ell}^{\nu,\lambda}, (a_i^{\nu,\lambda} - x)^{j-\lambda})] \cdot D^m \phi_{i\ell}^{\nu,\lambda}(x) = 0$$

which completes the proof for (2-66). Therefore, as an immediate consequence of the result (2-66), we obtain the result:

$$\sum_{j=\lambda}^{k-1} \frac{1}{j!} \sum_{i=1}^{N_0} [D^j u(x) \cdot (a_i^{\nu,0} - x)^j] D^m \phi_i^{\nu,0}(x) + \ldots$$

$$\ldots + \sum_{j=\lambda}^{k-1} \frac{1}{(j-\lambda)!} \sum_{i=1}^{N_\lambda} \sum_{\ell=1}^{\nu_i} [D^j u(x) \cdot (\xi_{i\ell}^{\nu,\lambda}, (a_i^{\nu,\lambda} - x)^{j-\lambda})] D^m \phi_{i\ell}^{\nu,\lambda}(x) = D^m u(x)$$

from which, together with (2-64), we obtain the required result (2-62).

From this point on it is an easy matter to derive an upper bound for the error in

$$D^m u(x) - D^m u_n^{(k-1)}(x)$$

for any integer $0 \leq m \leq k-1$, but, nevertheless, we wish to do that under the more general notion - suggested in [7] - of equivalent sets of interpolating points defined as follows: Suppose that $\bar{S}_n^\nu$, $\nu \in I$, is any simplex defined through a partition:

$$\Delta : \{\bar{S}_n^\nu\}, \quad \nu \in I$$

of the domain Ω, and:

$$S = \bigcup_{\mu=0}^{\lambda} S_n^{\nu,\mu}(k-1), \quad \nu \in I$$

is its associated set of interpolating points which in a unique fashion defines the Hermite (piecewise) interpolating polynomial of the formula (2-53). Furthermore, with the simplex $\bar{S}_n^\nu$, $\nu \in I$, we associate a set of basis functions defined by the conditions (2-54) as well as the following two geometrical parameters:

$$\begin{aligned} h_\nu &= \text{diameter of the simplex } \bar{S}_n^\nu, \ \nu \in I \\ \rho_\nu &= \text{diameter of the inscribed sphere in } \bar{S}_n^\nu, \ \nu \in I \end{aligned} \qquad \ldots \quad (2\text{-}69)$$

For practical purposes, however, we consider a family of all possible partitions:

$$\{\Delta_h\}, \quad h \in H$$

where H denotes a collection of positive parameters h, of the domain Ω and we say that it defines a regular family, if there exists a constant $a > 0$ such that the two geometrical parameters h_ν and ρ_ν, $\nu \in I$, defined in (2-69), satisfy the following inequality:

$$h_\nu \leq a\rho_\nu, \quad \text{for all} \quad \nu \in I \qquad \ldots \quad (2\text{-}70)$$

After that, we choose, once and for all, an n-simplex $\hat{\hat{S}}_n$, and let:

$$\hat{S} = \bigcup_{\mu=0}^{\lambda} \hat{S}_n^\mu(k\text{-}1)$$

be its associated (k-1)-unisolvent set of interpolating points. An analogous set of basis functions, to those defined through the conditions (2-54), are associated with the simplex $\hat{\hat{S}}_n$ as well as the following two geometrical parameters:

$$\begin{aligned} \hat{h} &= \text{diameter of the simplex } \hat{\hat{S}}_n \\ \hat{\rho} &= \text{diameter of the inscribed sphere in } \hat{\hat{S}}_n \end{aligned} \qquad \ldots \quad (2\text{-}71)$$

Next, we define an affine (one-to-one and onto) transformation:

$$x = R_\nu \hat{x} + r_\nu, \quad \nu \in I, \; R_\nu \in L(R^n), \; r_\nu \in R^n \qquad \ldots \quad (2\text{-}72)$$

mapping the simplex $\hat{\hat{S}}_n$ into the simplex $\hat{S}_n^\nu$, $\nu \in I$, and is such that the image of the set $\hat{S}$ under this transformation is exactly the set S. Furthermore, suppose that to each point:

$$a_i^{\nu,\mu}, \quad 1 \leq i \leq N_\mu, \quad 1 \leq \mu \leq \lambda, \quad \nu \in I$$

- respectively $\hat{a}_i^\mu$, $1 \leq i \leq N_\mu$, $1 \leq \mu \leq \lambda$ - is associated a subset:

$$A_i^{\nu,\mu} \subset (R^n)^\mu \quad - \text{resp.} \quad \hat{A}_i^\mu \subset (R^n)^\mu \quad -$$

Then, we say that the two sets $\hat{S}$ and S are equivalent if and only if the following two conditions are satisfied:

(i) $a_i^{\nu,\mu} = R_\nu \hat{a}_i^\mu + r_\nu$, $R_\nu \in L(\mathbb{R}^n)$, $r_\nu \in \mathbb{R}^n$, $\nu \in I$

(ii) $A_i^\mu = \{(\xi^1, \xi^2, \ldots, \xi^\mu) \in (\mathbb{R}^n)^\mu : \xi^i = R_\nu \hat{\xi}^i,$ . . . (2-73)

for all $(\hat{\xi}^1, \hat{\xi}^2, \ldots, \hat{\xi}^\mu) \in \hat{A}_i^\mu$, $1 \le i \le N_\mu$, $1 \le \mu \le \lambda$, $\nu \in I\}$

Finally, for the two equivalent sets of interpolating points S and $\hat{S}$, the following result (see P.G. Ciarlet and P.A. Raviart [7]) is of great importance in deriving the error bounds:

$$\|R_\nu\| \le \frac{h_\nu}{\hat{\rho}} \text{ and } \|R_\nu^{-1}\| \le \frac{\hat{h}}{\rho_\nu}, \quad R_\nu \in L(\mathbb{R}^n),\ r_\nu \in \mathbb{R}^n,\ \nu \in I \qquad \ldots (2\text{-}74)$$

where $\|\cdot\|$ denotes the usual Euclidean vector norm in $\mathbb{R}^n$. Then:

THEOREM I

Let a function $u(x) \in C^k(\bar{S}_n^\nu)$, $\nu \in I$, be given and $S = \bigcup_{\mu=0}^{\lambda} S_n^{\nu,\mu}(k\text{-}1)$, $\nu \in I$, be a (k-1)-Unisolvent set of interpolating points associated with the simplex $\bar{S}_n^\nu$, $\nu \in I$, defined by a partition of the domain $\bar{\Omega}$ such that the inequality (2-70) is satisfied. Then, if $v_n^{(k-1)}(x)$ is its unique Hermite interpolating polynomial defined by the formula:

$$u_n^{(k-1)}(x) = \sum_{i=1}^{N_0} u(\alpha_i^{\nu,0})\phi_i^{\nu,0}(x) + \sum_{i=1}^{N_1}\sum_{\ell=1}^{\nu_i} [Du(\alpha_i^{\nu,1})\cdot(\xi_{i\ell}^{\nu,1})]\phi_{i\ell}^{\nu,1}(x) + \ldots$$

$$\ldots + \sum_{i=1}^{N_\lambda}\sum_{\ell=1}^{\nu_i} [D^\lambda u(\alpha_i^{\nu,\lambda})\cdot(\xi_{i\ell}^{\nu,\lambda})]\phi_{i\ell}^{\nu,\lambda}(x),$$

we have:

$$\max_{\substack{x\in\bar{S}_n^\nu \\ \nu\in I \\ |\alpha|=m}} |D^\alpha u(x) - D^\alpha u_n^{(k-1)}(x)| \le Ch_\nu^{k-m} \cdot \max_{\substack{x\in\bar{S}_n^\nu \\ \nu\in I \\ |\beta|=k}} |D^\beta u(x)| \qquad \ldots \quad (2\text{-}75)$$

for all the integers m such that $0 \le m \le k-1$, where the numerical constant C is the same for all the equivalent (k-1)-Unisolvent sets of interpolating points S associated with the simplex $\bar{S}_n^\nu$, $\nu \in I$, of the partition Δ and is computed, once and for all, in a (k-1)-Unisolvent set $\hat{S}$ which is equivalent to S.

PROOF

Indeed, from (2-62), we immediately get:

$$\|D^m u(x) - D^m u_n^{(k-1)}(x)\| \le \sum_{i=1}^{N_0}\int_0^1 [\frac{(1-t)^{k-1}}{(k-1)!}\|D^k u(x+t(\alpha_i^{\nu,0}-x))\cdot(\alpha_i^{\nu,0}-x)^k\|]dt \cdot$$

$$\cdot \|D^m\phi_i^{\nu,0}(x)\| +$$

$$+ \sum_{i=1}^{N_1}\sum_{\ell=1}^{\nu_i}\int_0^1 [\frac{(1-t)^{k-2}}{(k-2)!}\|D^k u(x+t(\alpha_i^{\nu,1}-x))\cdot(\xi_{i\ell}^{\nu,1},(\alpha_i^{\nu,1}-x)^{k-1})\|]dt \cdot \|D^m\phi_{i\ell}^{\nu,1}(x)\| + \ldots$$

$$\ldots+\sum_{i=1}^{N_\lambda}\sum_{\ell=1}^{\nu_i}\int_0^1\left[\frac{(1-t)^{k-\lambda-1}}{(k-\lambda-1)!}\|D^k u(x+t(a_i^{\nu,\lambda}-x))\cdot(\xi_{i\ell}^{\nu,\lambda},(a_i^{\nu,\lambda}-x)^{k-\lambda})\|\right]dt.\|D^m\phi_{i\ell}^{\nu,\lambda}(x)\| \leq$$

$$\leq \frac{h_\nu^k}{k!}\cdot\sup_{x\in\bar{S}_n^\nu}\|D^k u(x)\|\cdot\sum_{i=1}^{N_0}\|D^m\phi_i^{\nu,0}(x)\|+\frac{h_\nu^{k-1}}{(k-1)!}\sup_{x\in\bar{S}_n^\nu}\|D^k u(x)\|\cdot\sum_{i=1}^{N_1}\sum_{\ell=1}^{\nu_i}\|\xi_{i\ell}^{\nu,1}\|\cdot\|D^m\phi_{i\ell}^{\nu,1}(x)\| +$$

$$\nu\in I \qquad\qquad \nu\in I$$

$$\ldots+\frac{h_\nu^{k-\lambda}}{(k-\lambda)!}\sup_{x\in\bar{S}_n^\nu}\|D^k u(x)\|\cdot\sum_{i=1}^{N_\lambda}\sum_{\ell=1}^{\nu_i}\|\xi_{i\ell}^{\nu,\lambda}\|\cdot\|D^m\phi_{i\ell}^{\nu,\lambda}(x)\| \qquad \ldots (2\text{-}76)$$

$$\nu\in I$$

Then, for the vectors $(\xi_{i1}^{\nu,\mu}, \xi_{i2}^{\nu,\mu},\ldots, \xi_{i\nu_i}^{\nu,\mu})$, $\nu\in I$, $1\leq\mu\leq\lambda$, which form a basis for the subspace $A_i^{\nu,\mu}$, $1\leq i\leq N_\mu$, $1\leq\mu\leq\lambda$, $\nu\in I$, and by recalling the conditions (2-73), we have:

$$\xi_{i\ell}^{\nu,\mu} = R_\nu^\mu\hat{\xi}_{i\ell}^\mu, \quad 1\leq i\leq N_\mu,\ 1\leq\mu\leq\lambda,\ 1\leq\ell\leq\nu_i,\ \nu\in I$$

or

$$\|\xi_{i\ell}^{\nu,\mu}\| \leq \|R_\nu\|^\mu\,\|\hat{\xi}_{i\ell}^\mu\| \leq \frac{h_\nu^\mu}{\hat{\rho}^\mu}\|\hat{\xi}_{i\ell}^\mu\|, \quad 1\leq\mu\leq\lambda,\ \nu\in I.$$

Thus, from (2-76), we get:

$$\|D^m u(x) - D^m u_n^{(k-1)}(x)\| \leq \frac{h_\nu^k}{k!}\sup_{x\in\bar{S}_n^\nu}\|D^k u(x)\|\cdot\sum_{i=1}^{N_0}\|D^m\phi_i^{\nu,0}(x)\| +$$

$$\nu\in I$$

$$+\frac{h_\nu^k}{(k-1)!}\frac{1}{\hat{\rho}}\cdot\sup_{x\in\bar{S}_n^\nu}\|D^k u(x)\|\cdot\sum_{i=1}^{N_1}\sum_{\ell=1}^{\nu_i}\|D^m\phi_{i\ell}^{\nu,1}(x)\|\cdot\|\hat{\xi}_{i\ell}^1\| +\ldots$$

$$\nu\in I$$

$$\ldots + \frac{h_\nu^k}{(k-\lambda)!}\cdot\frac{1}{\hat{\rho}^\lambda}\cdot \sup_{\substack{x\in\bar{S}_n^\nu \\ \nu\in I}} \|D^k u(x)\| \cdot \sum_{i=1}^{N_\lambda}\sum_{\ell=1}^{\nu_i} \|D^m \phi_{i\ell}^{\nu,\lambda}(x)\| \cdot \|\hat{\xi}_{i\ell}^\lambda\| \qquad \ldots \quad (2\text{-}77)$$

Furthermore, for the two sets of points S and $\hat{S}$, their respective basis functions satisfy the following conditions:

$$\phi_i^{\nu,0}(x) = \hat{\phi}_i^0(R_\nu^{-1}(x-r_\nu)),\ 1\le i\le N_0,\ \nu\in I$$

and

$$\phi_{i\ell}^{\nu,\mu}(x) = \hat{\phi}_{i\ell}^\mu(R_\nu^{-1}(x-r_\nu)),\ 1\le i\le N_\mu,\ 1\le\mu\le\lambda,\ 1\le\ell\le\nu_i,\ \nu\in I.$$

Therefore, for any vectors:

$$(\xi^1,\xi^2,\ldots,\xi^m)\in(R^n)^m,\ \xi^i\in R^n,\ 1\le i\le m$$

we have:

$$D^m\phi_i^{\nu,0}(x)\cdot(\xi^1,\xi^2,\ldots,\xi^m) = D^m\hat{\phi}_i^0(R_\nu^{-1}(x-r_\nu))\cdot(R_\nu^{-1}\xi^1,R_\nu^{-1}\xi^2,\ldots,R_\nu^{-1}\xi^m)$$

and, (2-78)

$$D^m\phi_{i\ell}^{\nu,\mu}(x)\cdot(\xi^1,\xi^2,\ldots,\xi^m) = D^m\hat{\phi}_{i\ell}^\mu(R_\nu^{-1}(x-r_\nu))\,(R_\nu^{-1}\xi^1,R_\nu^{-1}\xi^2,\ldots,R_\nu^{-1}\xi^m)$$

$$1\le\mu\le\lambda,\ \nu\in I$$

Then,

$$\sup_{\substack{x\in\bar{S}_n^\nu \\ \nu\in I}}\|D^m\phi_i^{\nu,0}(x)\| \le \sup_{\substack{x\in\bar{S}_n^\nu \\ \nu\in I}}\|D^m\hat{\phi}_i^0(R_\nu^{-1}(x-r_\nu))\|\cdot\|R_\nu^{-1}\|^m,\ \nu\in I$$

and (2-79)

$$\sup_{\substack{x\in\bar{S}_n^\nu \\ \nu\in I}}\|D^m\phi_{i\ell}^{\nu,\mu}(x)\| \le \sup_{\substack{x\in\bar{S}_n^\nu \\ \nu\in I}}\|D^m\hat{\phi}_{i\ell}^{\nu,\mu}(R_\nu^{-1}(x-r_\nu))\|\cdot\|R_\nu^{-1}\|^m$$

$$1\le\mu\le\lambda,\ \nu\in I$$

and since the image of the simplex $\hat{S}_n$ under the transformation (2-72) is exactly the simplex S_n^ν, $\nu \in I$, we have:

$$\sup_{\substack{x\in S_n^\nu \\ \nu\in I}} \| D^m \phi_i^{\nu,0}(R_\nu^{-1}(x-r_\nu))\| = \sup_{\hat{x}\in\hat{S}_n} \| D^m \hat{\phi}_i^0(\hat{x})\| \tag{2-80}$$

and

$$\sup_{\substack{x\in S_n^\nu \\ \nu\in I}} \| D^m \phi_{i\ell}^{\nu,\mu}(R_\nu^{-1}(x-r_\nu))\| = \sup_{\hat{x}\in\hat{S}_n} \| D^m \hat{\phi}_{i\ell}^{\mu}(\hat{x})\|$$

$$1 \le \mu \le \lambda, \ \nu \in I$$

Next, from (2-77) combined with (2-78), (2-79) and (2-80), we get:

$$\| D^m u(x) - D^m u_n^{(k-1)}(x)\| \le \frac{h_\nu^k}{k!} \sup_{\substack{x\in S_n^\nu \\ \nu\in I}} \| D^k u(x)\| . \frac{\hat{n}^m}{\rho_\nu^m} \sum_{i=1}^{N_0} \sup_{\hat{x}\in\hat{S}_n} \| D^m \hat{\phi}_i^0(\hat{x})\| +$$

$$+ \frac{h_\nu^k}{(k-1)!} . \frac{1}{\rho} . \sup_{\substack{x\in S_n^\nu \\ \nu\in I}} \| D^k u(x)\| . \frac{\hat{n}^m}{\rho_\nu^m} \sum_{i=1}^{N_1} \sum_{\ell=1}^{\nu_i} \sup_{\hat{x}\in\hat{S}_n} \| D^m \hat{\phi}_{i\ell}^1(\hat{x})\| . \| \hat{\xi}_{i\ell}^\lambda\| + \ldots$$

$$\ldots + \frac{h_\nu^k}{(k-\lambda)!} . \frac{1}{\hat{\rho}^\lambda} . \sup_{\substack{x\in S_n^\nu \\ \nu\in I}} \| D^k u(x)\| . \frac{\hat{n}^m}{\rho_\nu^m} \sum_{i=1}^{N_1} \sum_{\ell=1}^{\nu_i} \sup_{\hat{x}\in\hat{S}_n} \| D^m \hat{\phi}_{i\ell}^\lambda(\hat{x})\| . \| \hat{\xi}_{i\ell}^\lambda\| \quad \ldots \tag{2-81}$$

Therefore, from (2-81) combined together with the regularity condition (2-70), we obtain:

$$\| D^m u(x) - D^m u_n^{(k-1)}(x)\| \le C h_\nu^{k-m} \sup_{\substack{x\in S_n^\nu \\ \nu\in I}} \| D^k u(x)\|, \text{ for all } 0 \le m \le k-1 \quad \ldots \tag{2-82}$$

where the numerical constant C is given explicitly by:

$$C = a^m \frac{h^m}{k!} \sum_{i=1}^{N_0} \sup_{\hat{x}\in\hat{S}_n} \|D^m \phi_i^0(\hat{x})\| + a^m \frac{h^m}{(k-1)!} \frac{1}{\rho} \sum_{i=1}^{N_1} \sum_{\ell=1}^{\nu_i} \sup_{\hat{x}\in\hat{S}_n} \|D^m \phi_{i\ell}^1(\hat{x})\| . \|\hat{\tau}_{i\ell}^1\| + \ldots$$

$$\ldots + a^m \frac{h^m}{(k-\lambda)!} \frac{1}{\rho^\lambda} \sum_{i=1}^{N_\lambda} \sum_{\ell=1}^{\nu_i} \sup_{\hat{x}\in\hat{S}_n} \|D^m \phi_{i\ell}^\lambda(\hat{x})\| . \|\hat{\tau}_{i\ell}^\lambda\| \qquad \ldots (2\text{-}83)$$

Finally, by combining the inequality (2-82) with the double inequality (2-48), we obtain:

$$\max_{\substack{x\in S_n^\nu \\ \nu\in I \\ |\alpha|=m}} |D^\alpha u(x) - D^\alpha u_n^{(k-1)}(x)| \le C\, h_\nu^{k-m} \max_{\substack{x\in S_n^\nu \\ \nu\in I \\ |\beta|=k}} |D^\beta u(x)| \qquad \ldots (2\text{-}84)$$

for all the integers m such that $0 \le m \le k-1$. The constant C in (2-84) is given by:

$$C = C_n(k).C$$

where $C_n(k) = n^k$ and the numerical constant C is explicitly given by (2-83). This completes the proof.

2.2.3 Multivariate Lagrange Interpolation

Once the general multivariate Hermite approximation problem has been analysed, it is an easy matter to emphasise some of the essential points by referring to the particular problem of the Lagrange (piecewise) interpolation. Thus,

let $\bar{S}_n^\nu$, $\nu \in I$, again be any simplex of the partition Δ of a given polyhedral domain Ω, such that the following condition:

$$\bar{\Omega} = \bigcup_{\nu \in I} \bar{S}_n^\nu$$

is satisfied. Also, let:

$$S_n^\nu(k-1), \quad \nu \in I$$

be its associated discrete set of interpolating points defined through the formula (2-26) and which contains exactly

$$N = \binom{k+n-1}{k-1}$$

members. Then, for a function $u(x)$ with a C^k order of continuity over the simplex $\bar{S}_n^\nu$, $\nu \in I$, its unique Lagrange interpolating polynomial is given by the formula:

$$u_n^{(k-1)}(x) = \sum_{i=1}^{N} u(\alpha_i^\nu)\phi_i^\nu(x) \qquad \ldots \quad (2\text{-}85)$$

where $u(\alpha_i^\nu)$, $1 \le i \le N$, $\nu \in I$, are the values of the function $u(x)$ at the N points $\alpha_i^\nu \in S_n^\nu(k-1)$, $\nu \in I$, of the simplex and $\phi_i^\nu(x)$, $1 \le i \le N$, $\nu \in I$, are the basis functions which correspond to those points and are such that:

$$\phi_i^\nu(x) \in P_{k-1}(\bar{S}_n^\nu), \ \nu \in I \text{ and } \phi_i^\nu(\alpha_j^\nu) = \delta ij, \ 1 \le i,j \le N, \nu \in I.$$

By using now the formula (2-85) together with the multivariate Taylor formula defined by (2-61) we, again, can easily get the following important result:

let $\bar{S}_n^\nu$, $\nu \in I$, again be any simplex of the partition Δ of a given polyhedral domain Ω, such that the following condition:

$$\bar{\Omega} = \bigcup_{\nu \in I} \bar{S}_n^\nu$$

is satisfied. Also, let:

$$S_r^\nu(k-1), \quad \nu \in I$$

be its associated discrete set of interpolating points defined through the formula (2-26) and which contains exactly

$$N = \binom{k+n-1}{k-1}$$

members. Then, for a function u(x) with a C^k order of continuity over the simplex $\bar{S}_n^\nu$, $\nu \in I$, its unique Lagrange interpolating polynomial is given by the formula:

$$u_n^{(k-1)}(x) = \sum_{i=1}^{N} u(\alpha_i^\nu)\phi_i^\nu(x) \qquad \ldots \quad (2\text{-}85)$$

where $u(\alpha_i^\nu)$, $1 \le i \le N$, $\nu \in I$, are the values of the function u(x) at the N points $\alpha_i^\nu \in S_n^\nu(k-1)$, $\nu \in I$, of the simplex and $\phi_i^\nu(x)$, $1 \le i \le N$, $\nu \in I$, are the basis functions which correspond to those points and are such that:

$$\phi_i^\nu(x) \in P_{k-1}(\bar{S}_n^\nu), \ \nu \in I \text{ and } \phi_i^\nu(\alpha_j^\nu) = \delta ij, \ 1 \le i,j \le N, \nu \in I.$$

By using now the formula (2-85) together with the multivariate Taylor formula defined by (2-61) we, again, can *easily get the following important result:*

$$D^m u_n^{(k-1)}(x) = D^m u(x) + \sum_{i=1}^{N} \left[\int_0^1 \frac{(1-t)^{k-1}}{(k-1)!} D^k u(x+t(a_i^\nu - x)) \cdot (a_i^\nu - x)^k dt \right] D^m \phi_i^\nu(x) \quad \ldots \quad (2\text{-}86)$$

for all the integers m such that $0 \leq m \leq k-1$. This is true since, by virtue of the more general result given by (2-66), we always have:

$$\frac{1}{j!} \sum_{i=1}^{N} [D^j u(x) \cdot (a_i^\nu - x)^j] D^m \phi_i^\nu(x) = \begin{cases} D^m u(x), & \text{for } j=m \\ 0, & \text{for } j \neq m \end{cases}$$

The general notion of equivalent sets of interpolating points, again, plays an important role in the particular problem of the Lagrange approximation and the starting point for the computation of an error bound for the difference of the sort (2-49) will again be the result (2-86). Therefore, suppose that we choose, once and for all, an n-simplex $\hat{S}_n$ of the partition Δ with a discrete set of interpolating points $\hat{S}_n(k-1)$ associated with it and a set of basis functions $\hat{\phi}_i(x)$, $1 \leq i \leq N$, corresponding to the points $\hat{a}_i \in \hat{S}_n(k-1)$, $1 \leq i \leq N$. Next, define an affine transformation:

$$x = R_\nu \hat{x} + r_\nu, \quad R_\nu \in L(R^n), \quad r_\nu \in R^n$$

mapping the simplex $\hat{S}_n$ into the simplex S_n^ν, $\nu \in I$, and which is such that the image of the set $\hat{S}_n(k-1)$ under this tranformation is exactly the set $S_n(k-1)$, $\nu \in I$. Then, over any simplex S_n^ν, $\nu \in I$, of the partition Δ such that the inequality (2-70) is satisfied, by following the same steps as we did in Theorem I, we can easily prove that:

$$\| \mathcal{D}^m u(x) - \mathcal{D}^m u_n^{(k-1)}(x) \| \leq C \, h_\nu^{k-m} \sup_{\substack{x \in \bar{S}_n^\nu \\ \nu \in I}} \| \mathcal{D}^k u(x) \| \qquad \ldots \quad (2\text{-}87)$$

where $u(x) \in C^k(\bar{S}_n^\nu)$, $u_n^{(k-1)}(x) \in P_{k-1}(\bar{S}_n^\nu)$ is its unique *lagrange interpolating polynomial defined* through the formula (2-85), h_ν, $\nu \in I$, is the geometrical parameter associated with the simplex $\bar{S}_n^\nu$, $\nu \in I$, and defined by (2-69) and C is the following numerical constant:

$$C = a^m \frac{\hat{h}^k}{k!} \sum_{i=1}^{N} \sup_{\hat{x} \in \hat{\bar{S}}_n} \| \mathcal{D}^m \hat{\phi}_i(\hat{x}) \| \qquad \ldots \quad (2\text{-}88)$$

where a is the parameter involved in the inequality (2-70) and $\hat{h}$ is, again, a geometrical parameter defined by (2-71). Finally, by combining the inequality (2-87) with the double inequality (2-48), we obtain:

$$\max_{\substack{x \in \bar{S}_n^\nu \\ \nu \in I \\ |\alpha| = m}} |D^\alpha u(x) - D^\alpha u_n^{(k-1)}(x)| \leq C \, h_\nu^{k-m} \max_{\substack{x \in \bar{S}_n^\nu \\ \nu \in I \\ |\beta| = k}} |D^\beta u(x)|, \text{ for any } 0 \leq m \leq k-1 \qquad \ldots \quad (2\text{-}89)$$

with the numerical constant C given by:

$$C = C_n(k) . C$$

where $C_n(k) = n^k$ and C is the constant explicitly given by (2-88).

We only note here that the quantities:

$$\sup_{\hat{x} \in \hat{\bar{S}}_n} \| \mathcal{D}^m \hat{\phi}_i(\hat{x}) \|$$

in (2-88) can always be bounded in terms of the quantities:

$$\sup_{\hat{x}\in\hat{\bar{S}}_n} \| \hat{\phi}_i(\hat{x}) \|$$

which involve only the functions $\hat{\phi}_i(\hat{x})$, $1 \le i \le N$, instead of their derivatives. Indeed, from (2-48), we have that:

$$\sup_{\hat{x}\in\hat{\bar{S}}_n} \| D^m \hat{\phi}_i(\hat{x}) \| \le C_n(k) \cdot \max_{\hat{x}\in\hat{\bar{S}}_n} | D^\alpha \hat{\phi}_i(\hat{x}) | \quad \ldots \quad (2\text{-}90)$$

and since the function $\hat{\phi}_i(\hat{x})$, $1 \le i \le N$, is a polynomial of degree $(k-1)$ over the simplex $\hat{\bar{S}}_n$ - which constitute a compact convex subset of the space R^n - Markov's generalized inequality gives:

$$\max_{\substack{\hat{x}\in\hat{\bar{S}}_n \\ |\alpha|=m}} | D^\alpha \hat{\phi}_i(\hat{x}) | \le \frac{4^{|\alpha|} \cdot (k-1)^2 (k-2)^2 \ldots (k-|\alpha|)^2}{\hat{\rho}^{|\alpha|}} \cdot \max_{\hat{x}\in\hat{\bar{S}}_n} | \hat{\phi}_i(\hat{x}) | \quad \ldots \quad (2\text{-}91)$$

where the parameter $\hat{\rho}$ *is the same* as in (2-71). Thus, from (2-90) and (2-91), we obtain:

$$\sup_{\hat{x}\in\hat{\bar{S}}_n} \| D^m \hat{\phi}_i(\hat{x}) \| \le C_n(k) \cdot \frac{4^{|\alpha|} (k-1)^2 (k-2)^2 \ldots (k-|\alpha|)^2}{\hat{\rho}^{|\alpha|}} \cdot \max_{\hat{x}\in\hat{\bar{S}}_n} | \hat{\phi}_i(\hat{x}) |$$

and the numerical constant (2-88), instead, is given by:

$$C = \frac{k^k}{k!} C_n(k) \cdot \frac{4^{|\alpha|} (k-1)^2 (k-2)^2 \ldots (k-|\alpha|)^2}{\hat{\rho}^{|\alpha|}} \sum_{i=1}^{N} \max_{\hat{x}\in\hat{\bar{S}}_n} | \hat{\phi}_i(\hat{x}) | .$$

2.3 MULTIVARIATE SOBOLEV APPROXIMATION

2.3.1 Sobolev spaces

For a more general approximation shceme, than that described in the previous section, we need an introduction to the Sobolev spaces as well as to some useful results from Functional Analysis. Therefore, in addition to what has been said about the function space $H^k(\Omega)$ defined by (1-4) where k is any nonnegative integer and Ω any bounded and open subset of the space R^n, for any integer p with $1 \leq p < \infty$, we denote by:

$$H_p^k(\Omega)$$

the following Sobolev space:

$$H_p^k(\Omega) = \{u(x) : u(x) \in L_p(\Omega);\ D^\alpha u(x) \in L_p(\Omega),\ \text{for all } |\alpha| \leq k\} \quad \ldots \quad (2\text{-}92)$$

ie, $H_p^k(\Omega)$ is the space of functions which together with their generalized derivatives of up to the order k-th inclusive belong to the space $L_p(\Omega)$, where

$$L_p(\Omega) = \{u(x) : [\int_\Omega [u(x)]^p dx]^{1/p} < \infty\} \quad \ldots \quad (2\text{-}93)$$

For example, for p=2 and k=0, the space $H_p^k(\Omega)$ defined by (2-92) coincides with the space $L_2(\Omega)$ defined by (1-2), while for p=2 and any nonnegative integer k, the space $H_p^k(\Omega)$ exactly coincides with the space $H^k(\Omega)$ defined by (1-4). We equip the space (2-92) with a norm:

$$\|u\|^p_{H^k_p(\Omega)} = \sum_{|\alpha| \le k} \|D^\alpha u\|^p_{L_p(\Omega)}, \quad u \in H^k_p(\Omega) \qquad \ldots (2\text{-}94)$$

where, for any function $u \in L_p(\Omega)$:

$$\|u\|^p_{L_p(\Omega)} = \int_\Omega [u(x)]^p dx.$$

is the norm associated with the space $L_p(\Omega)$, $1 \le p < \infty$. Furthermore, corresponding to the norm (2-94), we define the following semi-norm:

$$|u|^p_{H^k_p(\Omega)} = \sum_{|\alpha| = k} \|D^\alpha u\|^p_{L_p(\Omega)}, \quad u \in H^k_p(\Omega) \qquad \ldots (2\text{-}95)$$

Then, it can easily be seen, from (2-94) combined with (2-95), that:

$$\|u\|^p_{H^k_p(\Omega)} = \sum_{i=0}^{k} |u|^p_{H^i_p(\Omega)}, \quad u \in H^k_p(\Omega) \qquad \ldots (2\text{-}96)$$

2.3.2 Sobolev Lemma

We recall here some important results from the imbedding theory over the Sobolev spaces $H^k_p(\Omega)$ which will be of great value in what we are going to say in our analysis hereafter. (An extensive analysis of that is given by Smirnov: A course in higher mathematics, vol IV).

Suppose that we have a function $u(x)$ which is defined over the Sobolev space $H^k_p(\Omega)$, for any integers $k \ge 1$ and $1 \le p < \infty$ such that:

$$pk > n \qquad \ldots (2\text{-}97)$$

where n denotes the dimension of the space R^n. Then, the function u(x) is continuous over the closed domain $\bar{\Omega}$; ie,

$$u(x) \in C(\bar{\Omega})$$

and the following inequality holds:

$$\|u\|_{C(\bar{\Omega})} \leq M \|u\|_{H_p^k(\Omega)} \qquad \ldots (2\text{-}98)$$

where by:

$$\|u\|_{C(\bar{\Omega})} = \max_{x \in \bar{\Omega}} |u(x)|$$

we define a norm over the continuous space $C(\bar{\Omega})$ and M is some numerical constant which does not depend on the function u(x). More generally, suppose that we have a function $u(x) \in H_p^k(\Omega)$ and m is some natural number such that:

$$p(k-m) > n \qquad \ldots (2\text{-}99)$$

Then, the function:

$$u(x) \in C^m(\bar{\Omega})$$

and the following inequality holds:

$$\|u\|_{C^m(\bar{\Omega})} \leq M \|u\|_{H_p^k(\Omega)} \qquad \ldots (2\text{-}100)$$

where, again, by:

$$\|u\|_{C^m(\bar{\Omega})} = \max_{\substack{|\alpha| = m \\ x \in \bar{\Omega}}} |D^\alpha u(x)|$$

we define a norm over the space $C^m(\bar{\Omega})$ and M is some numerical constant which does not depend on the function u(x). The inequality (2-100), with the inequality (2-99) being a special case of the first, is the well-known in Functional Analysis *Sobolev lemma* and simply relates the continuity of the function u(x) to the finite energy of the derivatives. We note here, however, that the Sobolev lemma is stated for star-shaped domains Ω and, since any convex region is star-shaped with respect to any of its points, it is also applicable for the several domains considered in the practical applications of the Finite Element Method.

2.3.3 The Quotient Space

Consider the following finite-dimensional subspace:

$$P_{k-1}(\Omega) \subset H_p^k(\Omega)$$

of all the polynomials of degree less than or equal to (k-1) defined over the domain Ω. We shall use the subspace $P_{k-1}(\Omega)$ to define an equivalence relation in $H_p^k(\Omega)$ as follows: two elements u_1 and u_2 in $H_p^k(\Omega)$ are said to be equivalent modulo $P_{k-1}(\Omega)$, if the difference u_1-u_2 is in $P_{k-1}(\Omega)$ and we write:

$$u_1 \equiv u_2 \ (\text{mod. } P_{k-1}(\Omega)).$$

It is an easy matter to verify that this is indeed an equivalence relation, ie, it has the usual properties

which characterize an equivalence relation: reflexivity, symmetry and transitivity. Thus, the space $H_p^k(\Omega)$ is divided into a number of mutually disjoint equivalence classes of functions, two functions being in the same equivalence class if and only if they are equivalent modulo $P_{k-1}(\Omega)$. We denote the set of all such equivalence classes by:

$$H_p^k(\Omega)/P_{k-1}(\Omega)$$

In order to answer the question of what is the structure of these equivalence classes, let u be an element function of the space $H_p^k(\Omega)$. The equivalent class containing u is by definition the set of all the elements υ such that:

$$\upsilon \equiv u(\text{mod.}P_{k-1}(\Omega)),$$

that is,

$$[u] = \{\upsilon : \upsilon \equiv u(\text{mod.}P_{k-1}(\Omega))\}$$

But, since:

$$\{\upsilon : \upsilon \equiv u(\text{mod.}P_{k-1}(\Omega))\} = \{\upsilon : \upsilon - u \in P_{k-1}(\Omega)\} = \{\upsilon : \upsilon - u = g, \text{ for some}$$

$$g \in P_{k-1}(\Omega)\} = \{\upsilon : \upsilon = u+g, \text{ for some } g \in P_{k-1}(\Omega)\} = \{u+g : g \in P_{k-1}(\Omega)\}$$

we can write that:

$$[u] = \{u+g : g \in P_{k-1}(\Omega)\} \equiv u+P_{k-1}(\Omega) \qquad \ldots \quad (2\text{-}101)$$

where the last notation is understood to signify the set of all sums of u and elements of $P_{k-1}(\Omega)$.

Therefore, a new linear space is constructed which we

denote by:

$$H_p^k(\Omega)/P_{k-1}(\Omega)$$

and call *the quotient space* of the space $H_p^k(\Omega)$ with respect to $P_{k-1}(\Omega)$. The manner in which the equivalence classes are added and scalarly multiplied - so that we indeed have a linear space - is as follows:

(i) $[u] + [v] = [u+v]$

(ii) $\alpha.[u] = [\alpha.u]$, for any scalar α

or, by using the notation (2-101):

(i)' $(u+P_{k-1}(\Omega)) + (v+P_{k-1}(\Omega)) = (u+v) + P_{k-1}(\Omega)$

(ii)' $\alpha.(u+P_{k-1}(\Omega)) = \alpha u + P_{k-1}(\Omega)$, for any scalar α

Then, the origin in $H_p^k(\Omega)/P_{k-1}(\Omega)$ is the equivalence class $0 + P_{k-1}(\Omega) = P_{k-1}(\Omega)$, and the negative of $u + P_{k-1}(\Omega)$ is $(-u) + P_{k-1}(\Omega)$.

2.3.4 The Quotient Norm

For any element:

$$[u] = u + P_{k-1}(\Omega) \in H_p^k(\Omega)/P_{k-1}(\Omega)$$

of the quotient space, we define:

$$\|[u]\|_{H_p^k(\Omega)/P_{k-1}(\Omega)} = \|u+P_{k-1}(\Omega)\|_{H_p^k(\Omega)/P_{k-1}(\Omega)} = \inf_{v \in P_{k-1}(\Omega)} \|u+v\|_{H_p^k(\Omega)} \quad \ldots \quad (2\text{-}102)$$

Then, formula (2-102) introduces a norm into the quotient

space $H^k_p(\Omega)/P_{k-1}(\Omega)$. Indeed, we have:

a. $\|[u_1]+[u_2]\|_{H^k_p(\Omega)/P_{k-1}(\Omega)} = \|(u_1+P_{k-1}(\Omega)) + (u_2+P_{k-1}(\Omega))\|_{H^k_p(\Omega)/P_{k-1}(\Omega)} =$

$= \|(u_1+u_2)+P_{k-1}(\Omega)\|_{H^k_p(\Omega)/P_{k-1}(\Omega)} = \inf_{v\in P_{k-1}(\Omega)} \|u_1+u_2+v\|_{H^k_p(\Omega)} =$

$= \inf_{v,v'\in P_{k-1}(\Omega)} \|u_1+u_2+v+v'\|_{H^k_p(\Omega)} = \inf_{v,v'\in P_{k-1}(\Omega)} \|(u_1+v)+(u_2+v')\|_{H^k_p(\Omega)} \le$

$\le \inf_{v,v'\in P_{k-1}(\Omega)} [\|u_1+v\|_{H^k_p(\Omega)} + \|u_2+v'\|_{H^k_p(\Omega)}] = \inf_{v\in P_{k-1}(\Omega)} \|u_1+v\|_{H^k_p(\Omega)} +$

$+ \inf_{v'\in P_{k-1}(\Omega)} \|u_2+v'\|_{H^k_p(\Omega)} = \|[u_1]\|_{H^k_p(\Omega)/P_{k-1}(\Omega)} + \|[u_2]\|_{H^k_p(\Omega)/P_{k-1}(\Omega)}$. Thus,

$\|[u_1]+[u_2]\|_{H^k_p(\Omega)/P_{k-1}(\Omega)} \le \|[u_1]\|_{H^k_p(\Omega)/P_{k-1}(\Omega)} + \|[u_2]\|_{H^k_p(\Omega)/P_{k-1}(\Omega)}$,

for any $[u_1]$, $[u_2] \in H^k_p(\Omega)/P_{k-1}(\Omega)$.

b. $\|\alpha.[u]\|_{H^k_p(\Omega)/P_{k-1}(\Omega)} = \|\alpha.(u+P_{k-1}(\Omega))\|_{H^k_p(\Omega)/P_{k-1}(\Omega)} =$

$= \|\alpha u+P_{k-1}(\Omega)\|_{H^k_p(\Omega)/P_{k-1}(\Omega)} = \inf_{v\in P_{k-1}(\Omega)} \|\alpha u+v\|_{H^k_p(\Omega)} =$

$= |\alpha| \cdot \inf_{v\in P_{k-1}(\Omega)} \|u+v\|_{H^k_p(\Omega)} = |\alpha| \cdot \|[u]\|_{H^k_p(\Omega)/P_{k-1}(\Omega)}$. Therefore,

$$\|\alpha.[u]\|_{H_p^k(\Omega)/P_{k-1}(\Omega)} = |\alpha|.\|[u]\|_{H_p^k(\Omega)/P_{k-1}(\Omega)}, \quad \text{for any}$$

$[u] \in H_p^k(\Omega)/P_{k-1}(\Omega)$ and any scalar $\alpha \in R$

c. Finally, it remains to show that if:

$$\|[u]\|_{H_p^k(\Omega)/P_{k-1}(\Omega)} = 0$$

then, $[u]$ is the zero element of the quotient space $H_p^k(\Omega)/P_{k-1}(\Omega)$. Indeed, suppose that:

$$\|[u]\|_{H_p^k(\Omega)/P_{k-1}(\Omega)} = 0, \quad \text{or that:}$$

$$\|u + P_{k-1}(\Omega)\|_{H_p^k(\Omega)/P_{k-1}(\Omega)} = 0. \qquad \ldots \quad (2\text{-}103)$$

Then, since the space $P_{k-1}(\Omega)$ is closed - it is a finite-dimensional space - (2-103) holds if and only if there exists a sequence:

$$\{v_n\} \in P_{k-1}(\Omega)$$

such that:

$$\|u+v_n\|_{H_p^k(\Omega)/P_{k-1}(\Omega)} \to 0, \quad \text{as} \quad n \to \infty$$

which implies that $u \in P_{k-1}(\Omega)$, and from which, in turn, we have that: $[u] = u + P_{k-1}(\Omega) = P_{k-1}(\Omega) =$ the zero element of the quotient space.

2.3.5 Equivalence of the Norm

Once the norm (2-102) has been introduced into the quotient space $H_p^k(\Omega)/P_{k-1}(\Omega)$ we shall next prove that this norm is equivalent to the semi-norm (2-95) defined over the space $H_p^k(\Omega)$; ie, we shall prove that:

$$|u|_{H_p^k(\Omega)} \leq \|[u]\|_{H_p^k(\Omega)/P_{k-1}(\Omega)} \leq C_1 |u|_{H_p^k(\Omega)} \quad \ldots \text{(2-104)}$$

where C_1 is some numerical constant which does not depend on the particular function u. The proof of the inequality (2-104) is based on the following two basic results which can be found in C. Morrey [16]:

(A) Suppose that $u \in H_p^k(\Omega)$. Then, there exists a unique polynomial:

$$p(x) \in P_{k-1}(\Omega)$$

such that:

$$\int_\Omega D^\alpha (u+p)\, dx = 0, \text{ for all } |\alpha| \leq k-1 \quad \ldots \text{(2-105)}$$

and

(B) Let Ω be a bounded domain which satisfies a strong cone condition - ie, we say that Ω satisfies a strong cone condition (Agmon [1]) if its boundary $\partial\Omega$ has a finite open covering:

$$\{O_i\},\ 1 \leq i \leq n$$

and corresponding cones:

$\{C_i\}$, $1 \le i \le n$

with vertices at the origin, such that:

$$(x+C_i) \subset \Omega$$

for any $x \in \Omega \cap O_i$. Then, if h denotes the diameter of the domain Ω, we have:

$$|u|_{H^i_p(\Omega)} \le C\, h^{k-i} |u|_{H^k_p(\Omega)}, \quad 0 \le i \le k-1 \qquad \ldots \quad (2\text{-}106)$$

for all functions $u \in H^k_p(\Omega)$ such that:

$$\int_\Omega D^\alpha u \, dx = 0, \text{ for all } |\alpha| \le k-1 \qquad \ldots \quad (2\text{-}107)$$

where C is some numerical constant which does not depend on the function u.

Then, for a proof of the inequality (2-104), consider the diameter h of the domain Ω to be equal to one. Thus, from (2-106) and for $h = 1$, we get:

$$|u|_{H^i_p(\Omega)} \le C\, |u|_{H^k_p(\Omega)}, \quad 0 \le i \le k-1 \qquad \ldots \quad (2\text{-}108)$$

for all the functions $u \in H^k_p(\Omega)$ such that (2-107) is satisfied. From the definition now of the H^k_p - norm together with (2-108), we have:

$$\|u+v\|^p_{H^k_p(\Omega)} = \sum_{i=0}^{k} |u+v|^p_{H^i_p(\Omega)} = \sum_{i=0}^{k-1} |u+v|^p_{H^i_p(\Omega)} + |u+v|^p_{H^k_p(\Omega)} \le$$

$$\le \sum_{i=0}^{k-1} C |u+v|^p_{H^k_p(\Omega)} + |u+v|^p_{H^k_p(\Omega)} = C' |u+v|^p_{H^k_p(\Omega)} \qquad \ldots \quad (2\text{-}109)$$

But since $v \in P_{k-1}(\Omega)$, we have:

$$|u+v|_{H_p^k(\Omega)} = |u|_{H_p^k(\Omega)}, \text{ for all } v \in P_{k-1}(\Omega)$$

and, from (2-109), we get:

$$\|u+v\|^p_{H_p^k(\Omega)} \leq C' \, |u|^p_{H_p^k(\Omega)}, \quad \text{or}$$

$$\|u+v\|_{H_p^k(\Omega)} \leq C_1 \, |u|_{H_p^k(\Omega)} \qquad \ldots \quad (2\text{-}110)$$

and, if we take the infimum over the space $P_{k-1}(\Omega)$ in (2-110), we obtain:

$$\inf_{v \in P_{k-1}(\Omega)} \|u+v\|_{H_p^k(\Omega)} \equiv \|[u]\|_{H_p^k(\Omega)/P_{k-1}(\Omega)} \leq C_1 \, |u|_{H_p^k(\Omega)}$$

Finally, for the left-hand inequality in (2-104), we have:

$$|u+v|_{H_p^k(\Omega)} = |u|_{H_p^k(\Omega)} \leq \inf_{v \in P_{k-1}(\Omega)} \|u+v\|_{H_p^k(\Omega)} = \|[u]\|_{H_p^k(\Omega)/P_{k-1}(\Omega)}$$

which completes the proof.

The more general approximation scheme which will be described hereafter, can briefly be outlined as follows: With each function u in the Sobolev space $H_p^k(\Omega)$ we associate a unique interpolation denoted by Pu, where P denotes some linear transformation with the sole assumption that it preserves all the polynomials of degree less

than or equal to (k-1) - Note that a similar condition is expressed by (2-65) for the case of Hermite interpolation -. Then, we can show that the following approximation estimate holds:

$$\|u-Pu\|_{H_p^m(\Omega)} \leq C\, h^{k-m} |u|_{H_p^k(\Omega)}, \text{ for any } 0 \leq m \leq k \qquad \ldots (2\text{-}111)$$

where h denotes the diameter of the domain Ω and C is some numerical constant which does not depend on the domain Ω. This new approach of the approximation problem which is unavoidable since the variational problem is posed on the Sobolev spaces $H_p^k(\Omega)$, is, of course, more elegant from the mathematical point of view and it contains the Lagrange and Hermite interpolation problems as special cases.

2.3.6 Bramble and Hilbert Lemma

Suppose now that $(H_p^k(\Omega))'$ denotes the (strong) dual space of the space $H_p^k(\Omega)$, for any integers $k \geq 1$ and $1 \leq p < \infty$, and that F(u) is a bounded linear functional on $H_p^k(\Omega)$; ie,

$$|F(u)| \leq K \|u\|_{H_p^k(\Omega)} \qquad \ldots (2\text{-}112)$$

The smallest constant K for which the inequality (2-112) holds, we call the norm of the functional F(u); it is given through the following formula:

$$\|F\|_{(H_p^k(\Omega))'} = \inf.\{K: |F(u)| \leq K \|u\|_{H_p^k(\Omega)}, \text{ for all } u \in H_p^k(\Omega)\} =$$

$$= \sup_{u \in H_p^k(\Omega)} \frac{|F(u)|}{\|u\|_{H_p^k(\Omega)}} \qquad \ldots (2\text{-}113)$$

Functionals which annihilate polynomials of a certain degree or less play an important role in this general approximation problem. Thus, the following Lemma, which was first introduced by Bramble and Hilbert [5], constitutes an effective tool for deriving approximation estimates:

Lemma:

Suppose that:

$$F \in (H_p^k(\Omega))'$$

is a bounded linear functional on $H_p^k(\Omega)$; ie,

$$|F(u)| \leq \|F\|_{(H_p^k(\Omega))'} \cdot \|u\|_{H_p^k(\Omega)} \qquad \ldots (2\text{-}114)$$

and is such that:

$$F(u) = 0, \text{ for all } u \in P_{k-1}(\Omega) \qquad \ldots (2\text{-}115)$$

Then,

$$|F(u)| \leq C_1 \|F\|_{(H_p^k(\Omega))'} \cdot |u|_{H_p^k(\Omega)} \qquad \ldots (2\text{-}116)$$

where C_1 is some numerical constant which does not depend on the function u (it is the same constant as in (2-104))

and the diameter h of the domain Ω is assumed to be equal to one.

Proof

Indeed, from the inequality (2-114) combined with (2-115), we have:

$$|F(u)| = |F(u+v)| \leq \|F\|_{(H_p^k(\Omega))'} \cdot \|u+v\|_{H_p^k(\Omega)}$$

Thus,

$$|F(u)| \leq \|F\|_{(H_p^k(\Omega))'} \cdot \inf_{v \in P_{k-1}(\Omega)} \|u+v\|_{H_p^k(\Omega)} = \|F\|_{(H_p^k(\Omega))'} \cdot \|[u]\|_{H_p^k(\Omega)/P_{k-1}(\Omega)}$$

or,

$$|F(u)| \leq C_1 \|F\|_{(H_p^k(\Omega))'} \cdot |u|_{H_p^k(\Omega)}$$

since $\|[u]\|_{H_p^k(\Omega)/P_{k-1}(\Omega)} \leq C_1 |u|_{H_p^k(\Omega)}$ by the inequality (2-104).

2.3.7 Multivariate Sobolev Interpolation

For any integer m with $0 \leq m \leq k$, consider now the space $H_p^m(\Omega)$ - of which the space $H_p^k(\Omega)$ is a subspace - and the following linear transformation:

$$P: H_p^k(\Omega) \to H_p^m(\Omega) : H_p^k(\Omega) \ni u \to Pu \in H_p^m(\Omega) \qquad \ldots \quad (2\text{-}117)$$

which is such that:

$$Pu=u, \text{ for all } u \in P_{k-1}(\Omega). \qquad \ldots (2\text{-}118)$$

ie, the mapping:

$$P \in L(H_p^k(\Omega);\ H_p^m(\Omega))$$

defined by (2-117) and (2-118) leaves invariant all the polynomials of degree less than or equal to (k-1).

Furthermore, for any element:

$$G \in (H_p^m(\Omega))'$$

define the following linear functional:

$$F : u \to F(u) = G(u-Pu) \qquad \ldots (2\text{-}119)$$

over the space $H_p^m(\Omega)$, $0 \le m \le k$, with its dual norm defined by:

$$\|G\|_{(H_p^m(\Omega))'} = \sup_{u \in H_p^m(\Omega)} \frac{|G(u-Pu)|}{\|u-Pu\|_{H_p^m(\Omega)}} \qquad \ldots (2\text{-}120)$$

Then, we can show that:

$$\|F\|_{(H_p^k(\Omega))'} \le \|G\|_{(H_p^m(\Omega))'} \cdot \|I-P\|_L \qquad \ldots (2\text{-}121)$$

where,

$$\|I-P\|_L = \sup_{u \in H_p^k(\Omega)} \frac{\|(I-P)u\|_{H_p^m(\Omega)}}{\|u\|_{H_p^k(\Omega)}} \qquad \ldots (2\text{-}122)$$

Indeed, from (2-120), we have:

$$|G(u-Pu)| \le \|G\|_{(H_p^m(\Omega))'} \cdot \|u-Pu\|_{H_p^m(\Omega)} \qquad \ldots (2\text{-}123)$$

and, since from (2-119):

$$|F(u)| = |G(u-Pu)|$$

from (2-123), we get:

$$|F(u)| \leq \|G\|_{(H_p^m(\Omega))'} \cdot \|u-Pu\|_{H_p^m(\Omega)}, \quad \text{or}$$

$$\frac{|F(u)|}{\|u\|_{H_p^k(\Omega)}} \leq \|G\|_{(H_p^m(\Omega))'} \cdot \frac{\|u-Pu\|_{H_p^m(\Omega)}}{\|u\|_{H_p^k(\Omega)}} \qquad \ldots \quad (2\text{-}124)$$

Therefore, by taking the supremum in both sides, of the inequality (2-124), over the space $H_p^k(\Omega)$, we have:

$$\sup_{u \in H_p^k(\Omega)} \cdot \frac{|F(u)|}{\|u\|_{H_p^k(\Omega)}} \leq \|G\|_{(H_p^m(\Omega))'} \cdot \sup_{u \in H_p^k(\Omega)} \frac{\|u-Pu\|_{H_p^m(\Omega)}}{\|u\|_{H_p^k(\Omega)}}$$

or,

$$\|F\|_{(H_p^k(\Omega))'} \leq \|G\|_{(H_p^m(\Omega))'} \cdot \|I-P\|_L$$

which completes the proof.

Next, a first bound for the difference:

$$\|u-Pu\|_{H_p^m(\Omega)}, \quad 0 \leq m \leq k$$

can easily be obtained and this is an immediate consequence of the inequality (2-121) together with the Bramble and Hilbert lemma and the following important result from Functional Analysis (see Taylor [26] theor. 4.3-B, p.186): Let u be an element of the space $H_p^m(\Omega)$ - actually it applies to any normed linear space - which is such that

$u \neq 0$. Then, there exists an element:

$$G \in (H_p^m(\Omega))'$$

such that:

$$\|G\|_{(H_p^m(\Omega))'} = 1 \quad \text{and} \quad \|u\|_{H_p^m(\Omega)} = G(u)$$

or, as a consequence, that:

$$\|u\|_{H_p^m(\Omega)} = \sup_{G \in (H_p^m(\Omega))'} |G(u)|$$

or

$$\|u\|_{H_p^m(\Omega)} = \sup_{G \in (H_p^m(\Omega))'} \frac{|G(u)|}{\|G\|_{(H_p^m(\Omega))'}} \qquad \ldots \quad (2\text{-}125)$$

Thus, from the lemma of Bramble and Hilbert combined together with the inequality (2-121), we get:

$$|F(u)| \leq \|G\|_{(H_p^m(\Omega))'} \cdot C_1 \cdot \|I-P\|_L \cdot |u|_{H_p^k(\Omega)}$$

or, since by definition $F(u) = G(u-Pu)$:

$$|G(u-Pu)| \leq \|G\|_{(H_p^m(\Omega))'} \cdot C_1 \cdot \|I-P\|_L \cdot |u|_{H_p^k(\Omega)}$$

Therefore,

$$\frac{|G(u-Pu)|}{\|G\|_{(H_p^m(\Omega))'}} \leq C_1 \|I-P\|_L \cdot |u|_{H_p^k(\Omega)}$$

and by taking the supremum of this inequality over the *dual space* $(H_p^m(\Omega))'$ *combined with the result* (2-125), we obtain:

$$\|u-Pu\|_{H_p^m(\Omega)} \leq C_1 . \|I-P\|_{L} . |u|_{H_p^k(\Omega)}, \text{ for any } 0 \leq m \leq k \quad \ldots (2-126)$$

where the constant C_1 is the s. as in (2-104) and the norm of the operator (I-P) is given by the formula (2-122).

Following the same steps, as for the pointwise approximation of the previous section, we shall give the final error bound for the difference:

$$\|u-Pu\|_{H_p^m(\Omega)}, \quad 0 \leq m \leq k$$

under the general notion of equivalent domains defined as follows: At first, with any domain Ω in R^n we associate the following two geometrical parameters:

$$\begin{aligned} h &= \text{diameter of the domain } \Omega \\ \rho &= \sup \{\text{diameter of the spheres contained in } \bar{\Omega}\} \end{aligned} \quad \ldots (2-127)$$

and let us suppose that there exists a constant a such that the following inequality is satisfied:

$$h < a\rho \quad \ldots (2-128)$$

in analogy to that given by (2-70). Next, we choose, once and for all, a domain:

$$\hat{\Omega}$$

and with it we associate, in a quite analogous way, the two geometrical parameters of (2-127). Then, we define the following affine transformation:

$$x = R\,\hat{x} + r, \quad R \in L(R^n), \; r \in R^n \quad \ldots (2-129)$$

mapping the domain Ω into the domain $\hat{\Omega}$ and is such that the image of $\hat{\Omega}$ under this transformation is exactly the domain Ω. Then, we say that the domains Ω and $\hat{\Omega}$ are equivalent under the mapping (2-129) and the following two important results hold (see P.G. Ciarlet and P.A. Raviart [7]):

1. With every function $u(x)$ *defined over* Ω, *we associate* a function $\hat{u}(\hat{x})$ defined over $\hat{\Omega}$ by letting:

$$\hat{u}(\hat{x}) = u(B\hat{x}+r), \text{ for every } \hat{x} \in \hat{\Omega} \qquad \ldots \quad (2\text{-}130)$$

Then, the mapping:

$$u \to \hat{u}$$

is an isomorphism between the spaces $W_p^m(\Omega)$ and $W_p^m(\hat{\Omega})$ for any integer m such that $0 \le m \le k$ and any integer $1 \le p < \infty$.

2. With every mapping:

$$P \in \mathcal{L}(W_p^k(\Omega);\ W_p^m(\Omega))$$

we associate a mapping:

$$\hat{P} \in \mathcal{L}(W_p^k(\hat{\Omega});\ W_p^m(\hat{\Omega}))$$

by letting:

$$\hat{P}\hat{u} = \widehat{Pu}, \text{ for each } u \in W_p^k(\Omega) \qquad \ldots \quad (2\text{-}131)$$

Then, if the mapping P leaves all the polynomials of degree less than or equal to (k-1) invariant; ie,

$$Pu = u, \text{ for all } u \in P_{k-1}(\Omega)$$

from (2-131), we have that:

$$\hat{P}\hat{u} = \hat{F}\hat{u} = \hat{u}, \text{ for all } \hat{u} \in P_{k-1}(\hat{n})$$

which means that also the mapping $\hat{P}$ leaves invariant all the polynomials of degree less than or equal to $(k-1)$.

By using next the inequality (from Fréchet differentiation):

$$|D^{\alpha}u(x)| \leq \|D^{k}u(x)\| \leq \sum_{|\mu|=k} \frac{k!}{\mu_1! \cdots \mu_n!} |D^{\mu}u(x)|, \quad |\alpha| = k \qquad \ldots (2\text{-}132)$$

we introduce a new norm over the space $H_p^k(\Omega)$ which is equivalent to the usual norm defined by the formula (2-94). This time, however, the Fréchet derivatives of the function u, instead of the usual composite partial derivatives, are involved. Indeed, by using the Hölder's inequality, we can write:

$$\sum_{|\mu|=k} \frac{k!}{\mu_1! \cdots \mu_n!} |D^{\mu}u(x)| \leq \left[\sum_{|\mu|=k} \left(\frac{k!}{\mu_1! \cdots \mu_n!}\right)^{q}\right]^{1/q} \cdot \left[\sum_{|\mu|=k} |D^{\mu}u(x)|^{p}\right]^{1/p} \qquad \ldots (2\text{-}133)$$

for any integers $p > 1$ and $q > 1$ such that:

$$1/p + 1/q = 1.$$

Therefore, the right-hand inequality in (2-132) can be written as:

$$\|D^{k}u(\alpha)\|^{p} \leq k_2 \sum_{|\mu|=k} |D^{\mu}u(x)|^{p} \qquad \ldots (2\text{-}134)$$

where the constant k_2 is:

$$k_2 = \left[\sum_{|\mu|=k} \left(\frac{k!}{\mu_1! \cdots \mu_n!}\right)^{q}\right]^{p/q} \qquad \ldots (2\text{-}135)$$

Furthermore, from the left-hand inequality in (2-132), we get:

$$|D^r u(x)|^p \leq \|D^k u(x)\|^p, \quad |r| = k$$

and

$$\sum_{|r|=k} |D^r u(x)|^p \leq k_1' \|D^k u(x)\|^p$$

with k_1' being equal to the number of all possible combinations of the partial derivatives for which we always have:

$$|r| = k \quad \text{and} \quad |r| = \sum_{i=1}^{n} r_i$$

that is,

$$k_1' = n^k \qquad \ldots \quad (2\text{-}136)$$

Thus,

$$k_1 \sum_{|r|=k} |D^r u(x)|^p \leq \|D^k u(x)\|^p \qquad \ldots \quad (2\text{-}137)$$

where $k_1 = 1/k_1'$ and k_1' is given by (2-136). By combining now the inequalities (2-134) and (2-137), we obtain:

$$k_1 \sum_{|r|=k} |D^r u(x)|^p \leq \|D^k u(x)\|^p \leq k_2 \sum_{|r|=k} |D^r u(x)|^p \qquad \ldots \quad (2\text{-}138)$$

and the following new norm and semi-norm can be introduced into the space $H_p^k(\Omega)$:

$$\|u\|^{*p}_{H_p^k(\Omega)} = \sum_{i=1}^{k} \|D^i u(x)\|^p_{L_p(\Omega)} \qquad \ldots \quad (2\text{-}139)$$

and

$$|u|^{*p}_{H^k_p(\Omega)} = \|\mathcal{D}^k u(x)\|^p_{L_p(\Omega)} \qquad \ldots \quad (2\text{-}140)$$

To prove that (2-139) is equivalent to the usual norm (2-94), we have: Integrating the double inequality (2-138) over the given domain Ω, we get:

$$k_1 \int_\Omega \sum_{|r|=k} |D^r u(x)|^p dx \le \int_\Omega \|\mathcal{D}^k u(x)\|^p dx \le k_2 \int_\Omega \sum_{|r|=k} |D^r u(x)|^p dx$$

Thus,

$$k_1 |u|^p_{H^k_p(\Omega)} \le |u|^{*p}_{H^k_p(\Omega)} \le k_2 |u|^p_{H^k_p(\Omega)} \qquad \ldots \quad (2\text{-}141$$

and

$$k_1 \|u\|^p_{H^k_p(\Omega)} \le \|u\|^{*p}_{H^k_p(\Omega)} \le k_2 \|u\|^p_{H^k_p(\Omega)} \qquad \ldots \quad (2\text{-}142)$$

which means, by definition, that the two norms are equivalent, where the constants k_1 and k_2 are given by (2-136) and (2-135) respectively. We conclude this subsection by giving the final theorem which gives an upper bound for the error between any function $u \in H^k_p(\Omega)$ and its approximation $Pu \in H^m_p(\Omega)$, for any $0 \le m \le k$:

THEOREM II

Let a function:

$$u \in H^k_p(\Omega), \quad k \ge 1, \quad 1 \le p < \infty$$

be given and suppose that the linear mapping:

$$P \in L(H^k_p(\Omega); H^m_p(\Omega)), \quad 0 \le m \le k$$

is such that:

$$Pu = u, \text{ for all } u \in P_{k-1}(\Omega).$$

Then, for any bounded and open subset Ω of R^n, for which the following condition:

$$h \leq a\rho$$

is satisfied, we have:

$$\|u-Pu\|_{H^m_p(\Omega)} \leq C\, h^{k-m}.|u|_{H^k_p(\Omega)}, \text{ for all } 0 \leq m \leq k \qquad \ldots (2\text{-}143)$$

where the numerical constant C is the same for all the equivalent domains Ω to $\hat{\Omega}$ and is computed, once and for all, for a domain $\hat{\Omega}$ which is equivalent to Ω.

PROOF

Indeed, since:

$$D^j u(x) \in L_j(R^n;R)$$

is a linear operator, from (2-130), we have:

$$D^j \hat{u}(\hat{x}).(\xi^1,\xi^2,\ldots,\xi^j) = D^j u(R\hat{x}+r).(R\xi^1,R\xi^2,\ldots,R\xi^j)$$

for any vector $(\xi^1,\xi^2,\ldots,\xi^j) \in (R^n)^j$, $\xi^i \in R^n$, $1 \leq i \leq j$

Thus,

$$\|D^j \hat{u}(\hat{x})\| \leq \|R\|^j.\|D^j u(R\hat{x}+r)\|$$

or, integrating over $\hat{\Omega}$:

$$\int_{\hat{\Omega}} \|D^j \hat{u}(\hat{x})\|^p d\hat{x} \leq \|R\|^{jp} \int_{\hat{\Omega}} \|D^j u(R\hat{x}+r)\|^p d\hat{x} \qquad \ldots (2\text{-}144)$$

Then, since for every $x \in \Omega$, $x = R\hat{x}+r$ with $\hat{x} \in \hat{\Omega}$, $R \in L(R^n)$ and $r \in R^n$, we have that:

$dx = |J|\,d\hat{x}$, where the Jacobian of the transformation: $|J| = |\det(R)|$

and, from (2-144), we get:

$$\int_{\hat{\Omega}} \|D^j \hat{u}(\hat{x})\|^p d\hat{x} \le \|R\|^{jp} |\det(R)|^{-1} \int_{\Omega} \|D^j u(x)\|^p dx$$

or

$$|\hat{u}|^*_{H^j_p(\hat{\Omega})} \le \|R\|^j |\det(R)|^{-1/p} |u|^*_{H^j_p(\Omega)} \qquad \ldots \quad (2\text{-}145)$$

By combining this inequality with that of (2-141), we get:

$$|\hat{u}|_{H^j_p(\hat{\Omega})} \le \left(\frac{k_2}{k_1}\right) \|R\|^j \, |\det(R)|^{-1/p} |u|_{H^j_p(\Omega)} \qquad \ldots \quad (2\text{-}146)$$

where the constants k_1 and k_2 are given by (2-136) and (2-135) respectively. Following the same argument as above, once more, from:

$$D^j u(x)\cdot(\xi^1,\xi^2,\ldots,\xi^j) = D^j \hat{u}(R^{-1}(x-r))\cdot(R^{-1}\xi^1, R^{-1}\xi^2,\ldots,R^{-1}\xi^j)$$

we get:

$$\|D^j u(x)\| \le \|R^{-1}\|^j \|D^j \hat{u}(R^{-1}(x-r))\|$$

and integrating over Ω:

$$\int_{\Omega} \|D^j u(x)\|^p dx \le \|R^{-1}\|^{jp} \int_{\Omega} \|D^j \hat{u}(R^{-1}(x-r))\|^p dx =$$

$$= \|R^{-1}\|^{jp} |\det(R)| \int_{\hat{\Omega}} \|D^j \hat{u}(\hat{x})\|^p d\hat{x}. \quad \text{Thus,}$$

$$|u|^*_{H^j_p(\Omega)} \le \|R^{-1}\|^j |\det(R)|^{1/p} |\hat{u}|^*_{H^j_p(\hat{\Omega})}, \quad \text{or}$$

$$|u|_{H^j_p(\Omega)} \le \left(\frac{k_2}{k_1}\right) \|R^{-1}\|^j |\det(R)|^{1/p} |\hat{u}|_{H^j_p(\hat{\Omega})} \qquad \ldots \quad (2\text{-}147)$$

Suppose now that the mapping R defined by (2-129) is such that:

$$\|R^{-1}\| \geq 1 \qquad \ldots \quad (2\text{-}148)$$

Then, from (2-147), we have:

$$\|u\|^p_{H^m_p(\Omega)} = \sum_{j=0}^{m} |u|^p_{H^j_p(\Omega)} \leq \left(\frac{k_2}{k_1}\right) |\det(R)| \sum_{j=0}^{m} \|R^{-1}\|^{jp} |\hat{u}|^p_{H^j_p(\hat{\Omega})} \leq$$

$$\leq \left(\frac{k_2}{k_1}\right) |\det(R)| \|R^{-1}\|^{mp} \sum_{j=0}^{m} |\hat{u}|^p_{H^j_p(\hat{\Omega})}$$

Therefore,

$$\|u\|^p_{H^m_p(\Omega)} \leq \left(\frac{k_2}{k_1}\right) |\det(R)| \|R^{-1}\|^{mp} \|\hat{u}\|^p_{H^m_p(\hat{\Omega})} \qquad \ldots \quad (2\text{-}149)$$

Next, from (2-149), we get:

$$\|u-Pu\|_{H^m_p(\Omega)} \leq \left(\frac{k_2}{k_1}\right)^{1/p} |\det(R)|^{1/p} \|R^{-1}\|^m \|\hat{u}-\hat{P}\hat{u}\|_{H^m_p(\hat{\Omega})} \qquad \ldots \quad (2\text{-}150)$$

and, since from inequality (2-126):

$$\|\hat{u}-\hat{P}\hat{u}\|_{H^m_p(\hat{\Omega})} \leq C_1 \|I-\hat{P}\|_{\hat{L}} |\hat{u}|_{H^k_p(\hat{\Omega})}$$

the inequality (2-150) can be written as:

$$\|u-Pu\|_{H^m_p(\Omega)} \leq \left(\frac{k_2}{k_1}\right)^{1/p} |\det(R)|^{1/p} \cdot C_1 \cdot \|R^{-1}\|^m \|I-\hat{P}\|_{\hat{L}} |\hat{u}|_{H^k_p(\hat{\Omega})} \qquad \ldots \quad (2\text{-}151)$$

Furthermore, from (2-146) and for $j = k$, we get:

$$|\hat{u}|_{H^k_p(\hat{\Omega})} \leq \left(\frac{k_2}{k_1}\right) \|R\|^k |\det(R)|^{-1/p} |u|_{H^k_p(\Omega)} \qquad \ldots \quad (2\text{-}152)$$

and, from (2-151) combined with (2-152), we obtain:

$$\|u-Pu\|_{H_p^m(\Omega)} \leq \left(\frac{k_2}{k_1}\right)^{1/p} |\det(R)|^{1/p} \|R^{-1}\|^m C_1 \|I-\hat{\Pi}\|_{\hat{L}} \cdot \left(\frac{k_2}{k_1}\right) \cdot \|R\|^k \cdot$$

$$\cdot |\det(R)|^{-1/p} |u|_{H_p^k(\Omega)}, \quad \text{or}$$

$$\|u-Pu\|_{H_p^m(\Omega)} \leq C_1 \left(\frac{k_2}{k_1}\right)^{(p+1)/p} \|R^{-1}\|^m \|R\|^k \|I-\hat{\Pi}\|_{\hat{L}} |u|_{H_p^k(\Omega)} \qquad \ldots (2\text{-}153)$$

Then, since:

$$\|R^{-1}\|^m \leq \frac{\hat{h}^m}{\rho^m} \quad \text{and} \quad \|R\|^k \leq \frac{h^k}{\hat{\rho}^k}$$

from (2-153), we have:

$$\|u-Pu\|_{H_p^m(\Omega)} \leq C \frac{h^k}{\rho^m} |u|_{H_p^k(\Omega)}, \quad \text{for all } 0 \leq m \leq k \qquad \ldots (2\text{-}154)$$

where:

$$C = C_1 \left(\frac{k_2}{k_1}\right)^{(p+1)/p} \frac{\hat{h}^m}{\hat{\rho}^k} \|I-\hat{\Pi}\|_{\hat{L}} \qquad \ldots (2\text{-}155)$$

Finally, from (2-154) together with (2-128), we obtain:

$$\|u-Pu\|_{H_p^m(\Omega)} \leq C\, h^{k-m} |u|_{H_p^k(\Omega)}, \quad \text{for all } 0 \leq m \leq k$$

where the numerical constant $C = aC$ and C is given by (2-155). Furthermore, C_1 is the constant involved in the inequality (2-116), k_1 and k_2 are given by (2-136) and (2-135) respectively and according to the formula (2-122):

$$\|I-\hat{P}\|_{\hat{L}} = \sup_{\hat{u}\in H_p^k(\hat{\Omega})} \frac{\|(I-\hat{P})\hat{u}\|_{H_p^m(\hat{\Omega})}}{\|\hat{u}\|_{H_p^k(\hat{\Omega})}}$$

2.4 APPROXIMATION BY CONVOLUTION

There still remains one important possibility which has to be examined as far as the approximation problem is concerned, viz: what happens if the function u (ie, the exact solution of the problem), which belongs to the Sobolev space $H_p^k(\Omega)$, $k \geq 1$, $1 \leq p < \infty$, does not satisfy the necessary continuity conditions for its interpolating polynomial to be determined?

In order to cope with this new situation a natural remedy arises from the ability to smooth (or molify) the given function u sufficiently before we apply, say, the Hermite approximation technique of the section 2.2. In this case, however, we have to distinguish between the following two basic types of error:

a. The error committed in replacing the given function u by its smoothed version, always denoted by Ju, and

b. The error committed in replacing the (smooth) function Ju by its interpolating polynomial, generally denoted by Pu.

Then, the overall error between the function u and its interpolate Pu, by using the triangle inequality, is given by:

$$\|u-Pu\| \leq \|u-Ju\| + \|Ju-Pu\| \qquad \ldots \quad (2\text{-}156)$$

in some norm $\|.\|$. Furthermore, if these two types of error are of the same order of accuracy, then, the function Pu will give an accurate approximation to u, even when the function u is not continuous. In this subsection, however, we will mainly be concerned with the error in:

$$\|u - Ju\|$$

since the second:

$$\|Ju-Pu\|$$

can be obtained by a straightforward application of the theorems I and II.

2.4.1 The Space $H_p^k(R^n)$

Let:

$$C^\infty(R^n)$$

denote the class of all real functions which are infinitely differentiable over the entire Euclidean space R^n, and

$$C_0^\infty(R^n)$$

be a subset of $C^\infty(R^n)$ consisting of those functions which have a compact support contained in R^n. Then, the completion of the space $C_0^\infty(R^n)$ under the norm:

$$\|u\|^p_{L_p(R^n)} = \int_{R^n} |u(x)|^p dx \qquad \ldots \quad (2\text{-}157)$$

defines the space $L_p(R^n)$. Furthermore, for any nonnegative integer k, the completion of the space $C_0^\infty(R^n)$ under the norm:

$$\|u\|^p_{H_p^k(R^n)} = \sum_{|\alpha|\le k} \|D^\alpha u\|^p_{L_p(R^n)} \qquad \ldots (2\text{-}158)$$

with $\|D^\alpha u\|^p_{L_p(R^n)}$ defined as in (2-157), defines the space $H_p^k(R^n)$.

2.4.2 The Convolution Operator

In order to smooth or molify a given function $u \in H_p^k(R^n)$ we make use of the technique of convolving the given function with a second sufficiently differentiable function. In general, by definition, if u(x) and v(x) are two functions in $L_p(R^n)$, the convolution of u and v, denoted by u*v, also belongs to the space $L_p(R^n)$ and is defined through the following integral:

$$u*v(x) = \int_{R^n} u(x-y)\,v(y)\,dy = \int_{R^n} v(x-y)\,u(y)\,dy = v*u(x) \qquad \ldots (2\text{-}159)$$

where, in the second integral above, we have replaced, for a fixed x, the variable y by x-y. Thus, for a function $u \in H_p^k(R^n)$ and a sufficiently differentiable function:

$$J(x) \in C_0^\infty(R^n)$$

their convolution is defined - by virtue of the definition formula (2-159) - by:

$$Ju = u*J(x) = \int_{R^n} u(x-y)\,J(y)\,dy \qquad \ldots (2\text{-}160)$$

The most important result about the convolution operator (2-160) is that, under certain conditions imposed on the smoothing function $J(x)$, the function Ju behaves very much like the function u - in the sense of the approximation Ju is, indeed, very close to u - but it is much smoother than u. This last observation is a consequence of the fact that, for any integer α:

$$D^{\alpha}Ju = D^{\alpha}(u*J(x)) = u*D^{\alpha}J(x)$$

and, since $J(x) \in C_0^{\infty}(R^n)$ is infinitel differentiable, the derivative of Ju always exist.

2.4.3 Application to the Approximation Problem

The problem of approximation by convolution having been considered by G. Strang in one of his published papers on approximation (see [23]) in a somewhat different way by using Fourier transforms, we should like to take the opportunity here, of giving, what we believe it to be, a more elegant mathematical treatment of this same problem by using the various results of the previous section over the Sobolev spaces.

Thus, consider an arbitrary function $u \in H_p^k(R^n)$ and let $J(x) \in C_0^{\infty}(R^n)$ be a smoothing function which has compact support in some set K - subset of R^n - and satisfies the following condition:

$$\int_{R^n} J(x)\,dx = 1 \qquad \ldots \quad (2\text{-}161)$$

Then, for any $h > 0$, we define the function:

$$J_h(x) = h^{-n} J(x/h) \qquad \ldots (2\text{-}162)$$

It is obvious, from (2-161) and (2-162), that by a change of variables, $y = x/h$, the function $J_h(x)$ defined by (2-162) also satisfies the condition (2-161); ie,

$$\int_{R^n} J_h(x)\,dx = h^{-n} \int_{R^n} J(x/h)\,dx = h^{-n} \int_{R^1} h^n J(y)\,dy = \int_{R^n} J(y)\,dy = 1$$

Next, by using the Hölder inequality and the Fubini theorem, we prove the following basic result:

$$\| D^\alpha J_h u \|_{L_p(R^n)} \le C\, h^{-|\alpha|} \|u\|_{L_p(R^n)}, \text{ for all } 0 \le |\alpha| \le k \qquad \ldots (2\text{-}163)$$

where the function $u \in H_p^k(R^n)$, the function $J_h u$ is given by the convolution formula:

$$J_h u = h^{-n} \int_{R^n} u(x-y) J(y/h)\,dy \qquad \ldots (2\text{-}164)$$

and the numerical constant C does not depend on the function u and the parameter h. Indeed, from (2-164), we have:

$$D^\alpha J_h u = h^{-n} \int_{R^n} u(x-y) D^\alpha J(y/h)\,dy$$

or

$$|D^\alpha J_h u| \le h^{-n} \int_{R^n} |u(x-y)\,(D^\alpha J(y/h))^{1/p}| \cdot |D^\alpha J(y/h)|^{1/q} dy$$

for any integers $p > 1$ and $q > 1$ such that:

$$1/p + 1/q = 1$$

Then, by applying the Hölder's inequality, we get:

Then, for any $h > 0$, we define the function:

$$J_h(x) = h^{-n} J(x/h) \qquad \ldots (2\text{-}162)$$

It is obvious, from (2-161) and (2-162), that by a change of variables, $y = x/h$, the function $J_h(x)$ defined by (2-162) also satisfies the condition (2-161); ie,

$$\int_{R^n} J_h(x)\,dx = h^{-n} \int_{R^n} J(x/h)\,dx = h^{-n} \int_{R^n} h^n J(y)\,dy = \int_{R^n} J(y)\,dy = 1$$

Next, by using the Hölder inequality and the Fubini theorem, we prove the following basic result:

$$\| D^\alpha J_h u \|_{L_p(R^n)} \le C\, h^{-|\alpha|} \|u\|_{L_p(R^n)}, \text{ for all } 0 \le |\alpha| \le k \qquad \ldots (2\text{-}163)$$

where the function $u \in H_p^k(R^n)$, the function $J_h u$ is given by the convolution formula:

$$J_h u = h^{-n} \int_{R^n} u(x-y) J(y/h)\,dy \qquad \ldots (2\text{-}164)$$

and the numerical constant C does not depend on the function u and the parameter h. Indeed, from (2-164), we have:

$$D^\alpha J_h u = h^{-n} \int_{R^n} u(x-y)\, D^\alpha J(y/h)\,dy$$

or

$$|D^\alpha J_h u| \le h^{-n} \int_{R^n} |u(x-y)\,(D^\alpha J(y/h))^{1/p}| \cdot |D^\alpha J(y/h)|^{1/q} dy$$

for any integers $p > 1$ and $q > 1$ such that:

$$1/p + 1/q = 1$$

Then, by applying the Hölder's inequality, we get:

$$|D^{\alpha}J_h u| \le h^{-n}\left(\int_{R^n} |u(x-y)|^p \cdot |D^{\alpha}J(y/h)|\,dy\right)^{1/p} \cdot \left(\int_{R^n} |D^{\alpha}J(y/h)|\,dy\right)^{1/q}$$

or

$$|D^{\alpha}J_h u|^p \le h^{-np}\left(\int_{R^n} |u(x-y)|^p \cdot |D^{\alpha}J(y/h|\,dy\right) \cdot \left(\int_{R^n} |D^{\alpha}J(y/h)|\,dy\right)^{p/q}$$

and, integrating the above inequality over R^n, we obtain:

$$\int_{R^n} |D^{\alpha}J_h u|^p dx \le \int_{R^n} h^{-np}\left[\left(\int_{R^n} |u(x-y)|^p \cdot |D^{\alpha}J(y/h)|\,dy\right) \cdot \left(\int_{R^n} |D^{\alpha}J(y/h)|\,dy\right)^{p/q}\right]dx$$

or, by using the Fubini theorem, that:

$$\| D^{\alpha}J_h u\|^p_{L_p(R^n)} \le h^{-np}\left(\int_{R^n} |u(x-y)|^p dx\right) \cdot \left(\int_{R^n} |D^{\alpha}J(y/h)|\,dy\right)^{1+p/q}$$

since:

$$\int_{R^n} |D^{\alpha}J_h u|^p dx = \| D^{\alpha}J_h u\|^p_{L_p(R^n)}$$

Therefore,

$$\| D^{\alpha}J_h u\|^p_{L_p(R^n)} \le h^{-np}\|u\|^p_{L_p(R^n)} \cdot \left(\int_{R^n} |D^{\alpha}J(y/h)|\,dy\right)^p \qquad \ldots \quad (2\text{-}165)$$

Then, since:

$$|D^{\alpha}J(y/h)| = h^{-|\alpha|}\,|(D^{\alpha}J)(y/h)|$$

from (2-165), we have:

$$\|D^{\alpha}J_h u\|_{L_p(R^n)} \le h^{-n}\|u\|_{L_p(R^n)} \cdot h^{-|\alpha|} \cdot C \cdot \int_{R^n} dy \le C h^{-|\alpha|}\|u\|_{L_p(R^n)}$$

Thus,

$$\| D^{\alpha}J_h u\|_{L_p(R^n)} \le C\, h^{-|\alpha|}\|u\|_{L_p(R^n)}$$

where $|(D^{\alpha}J)(y/h)| \leq C$ is an upper bound for the function $D^{\alpha}J$. This completes the proof for the inequality (2-163). We shall use this result in order to obtain an estimate for the error between the function u and its smoothed version Ju, constructed by convolution.

Another technical result which will be of great importance in the sequel, comes from a modification of the Bramble and Hilbert lemma to cope with the situation when the diameter of the domain Ω of R^n is equal to $h > 0$ and not equal to one as was previously assumed. Therefore: Suppose that:

$$F \in (H_p^k(\Omega))'$$

is a bounded linear functional on $H_p^k(\Omega)$; ie,

$$|F(u)| \leq \|F\|_{(H_p^k(\Omega))'} \|u\|_{H_p^k(\Omega)} \qquad \ldots \quad (2\text{-}166)$$

and such that:

$$F(v) = 0, \text{ for all } v \in P_{k-1}(\Omega)$$

Then,

$$|F(u)| \leq C_1 \|F\|_{(H_p^k(\Omega))'} h^k |u|_{H_p^k(\Omega)} \qquad \ldots \quad (2\text{-}167)$$

where, again, h denotes the diameter of the domain Ω and C_1 is some numerical constant which does not depend on the function u and the parameter h. For the proof of the inequality (2-167) it makes sense to introduce into the space $H_p^k(\Omega)$ the following norm:

$$\|u\|^p_{H^k_p(\Omega)} = \sum_{i=0}^{k} h^{ip} |u|^p_{H^i_p(\Omega)}$$

which is equivalent to the usual norm defined by the formula (2-94). As a result, the inequality (2-104) is written as:

$$h^k |u|_{H^k_p(\Omega)} \leq \|\Gamma u\|_{H^k_p(\Omega)/P_{k-1}(\Omega)} \leq C_1 h^k |u|_{H^k_p(\Omega)} \quad \ldots \quad (2\text{-}168)$$

which in turn gives:

$$|F(u)| = |F(u+v)| \leq \|F\|_{(H^k_p(\Omega))'} \|u+v\|_{H^k_p(\Omega)} \leq$$

$$\leq \|F\|_{(H^k_p(\Omega))'} \inf_{v \in P_{k-1}(\Omega)} \|u+v\|_{H^k_p(\Omega)} =$$

$$= \|F\|_{(H^k_p(\Omega))'} \|\Gamma u\|_{H^k_p(\Omega)/P_{k-1}(\Omega)} \leq C_1 \|F\|_{(H^k_p(\Omega))'} h^k |u|_{H^k_p(\Omega)},$$

where in the last inequality above, we have used the right-hand inequality of (2-168).

We consider now a decomposition of the space R^n into hypercubes Ω_i with sides of length equal to h; ie,

$$R^n = \cup_i \Omega_i$$

with $\Omega_i \cap \Omega_j = \emptyset$ or an edge of Ω_i, for any two successive indices $i \neq j$. With any such hypercube Ω_i and any point:

$$x \in \Omega_i$$

we associate the following set of points:

$$\omega_h^i = \{\Omega_j : |x-y| \leq h,\ y \in \Omega_j\},\ x \in \Omega_i \qquad \ldots \text{(2-169)}$$

and for the smoothing function:

$$J_h(x) \in C_0^\infty(R^n)$$

defined by the formula (2-162), we make the following two basic assumptions:

i. For any point $x \in \Omega_i$, the function $J_h(x)$ has compact support contained in ω_h^i, and

ii. $J_h(x)$ leaves invariant all the polynomials of degree less than or equal to (k-1); ie,

$$J_h u = u, \text{ for all } u \in P_{k-1}(\omega_h^i)$$

where by $P_{k-1}(\omega_h^i)$ we define the space of all the polynomials of degree less than or equal to (k-1) restricted over the subdomain ω_h^i.

Then, under these assumptions and by following a process analogous to that of the previous section for Sobolev interpolation, we shall prove the result:

$$\| D^\alpha (J_h u - u) \|_{L_p(R^n)} \leq C\, h^{k-|\alpha|} |u|_{H_p^k(R^n)} \qquad \ldots \text{(2-170)}$$

for all the integers α such that $0 \leq |\alpha| \leq k$, where C is some numerical constant which does not depend on the function u and the parameter h. Indeed, for the proof of the inequality (2-170), for any $x \in \Omega_i$ and any element:

$$T \in (L_p(\omega_h^i))'$$

we define the following linear functional:

$$F : H_p^k(\omega_h^i) \ni u \mapsto F(u) = T(D^\alpha(J_h u - u)), \text{ for all } 0 \le |\alpha| \le k \quad \ldots \quad (2\text{-}171)$$

over the space $L_p(\omega_h^i)$, and with its dual norm given by the following formula:

$$\|T\|_{(L_p(\omega_h^i))'} = \sup_{u \in H_p^k(\omega_h^i)} \frac{|T(D^\alpha(J_h u - u))|}{\|D^\alpha(J_h u - u)\|_{L_p(\omega_h^i)}}$$

Thus,

$$|T(D^\alpha(J_h u - u))| \le \|T\|_{(L_p(\omega_h^i))'} \cdot \|D^\alpha(J_h u - u)\|_{L_p(\omega_h^i)} \quad \ldots (2\text{-}172)$$

Then, since from (2-171):

$$|F(u)| = |T(D^\alpha(J_h u - u))|, \quad u \in H_p^k(\omega_h^i)$$

from (2-172), we get:

$$|F(u)| \le \|T\|_{(L_p(\omega_h^i))'} \cdot \|D^\alpha(J_h u - u)\|_{L_p(\omega_h^i)}, \text{ or}$$

$$\frac{|F(u)|}{\|u\|_{H_p^k(\omega_h^i)}} \le \|T\|_{(L_p(\omega_h^i))'} \frac{\|D^\alpha(J_h u - u)\|_{L_p(\omega_h^i)}}{\|u\|_{H_p^k(\omega_h^i)}} \quad \ldots \quad (2\text{-}173)$$

and by taking the supremum in both sides of (2-173) over the space $H_p^k(\omega_h^i)$, we obtain:

$$\sup_{u \in H_p^k(\omega_h^i)} \frac{|F(u)|}{\|u\|_{H_p^k(\omega_h^i)}} \le \|T\|_{(L_p(\omega_h^i))'} \sup_{u \in H_p^k(\omega_h^i)} \frac{\|D^\alpha(J_h u - u)\|_{L_p(\omega_h^i)}}{\|u\|_{H_p^k(\omega_h^i)}}$$

or, by the definition of the dual norm over the dual space:

$$\|F\|_{(H_p^k(\omega_h^i))'} \leq \|T\|_{(L_p(\omega_h^i))'} \cdot \sup_{u \in H_p^k(\omega_h^i)} \frac{\|D^\alpha (J_h u - u)\|_{L_p(\omega_h^i)}}{\|u\|_{H_p^k(\omega_h^i)}} \quad \ldots \quad (2\text{-}174)$$

By using now the inequality (2-163), we get:

$$\sup_{u \in H_p^k(\omega_h^i)} \frac{\|D^\alpha (J_h u - u)\|_{L_p(\omega_h^i)}}{\|u\|_{H_p^k(\omega_h^i)}} \leq \sup_{u \in H_p^k(\omega_h^i)} \frac{\|D^\alpha J_h u\|_{L_p(\omega_h^i)} + \|D^\alpha u\|_{L_p(\omega_h^i)}}{\|u\|_{H_p^k(\omega_h^i)}} \leq$$

$$\leq \sup_{u \in H_p^k(\omega_h^i)} \frac{Ch^{-|\alpha|}\|u\|_{L_p(\omega_h^i)} + \|D^\alpha u\|_{L_p(\omega_h^i)}}{\|u\|_{H_p^k(\omega_h^i)}} \leq \sup_{u \in H_p^k(\omega_h^i)} \left[Ch^{-|\alpha|} + \frac{\|D^\alpha u\|_{L_p(\omega_h^i)}}{\|u\|_{H_p^k(\omega_h^i)}}\right] \leq$$

$$\leq Ch^{-|\alpha|} + 1 \leq C h^{-|\alpha|}$$

for some numerical constant C, since by definition:

$$\|D^\alpha u\|_{L_p(\omega_h^i)} \leq \|u\|_{H_p^k(\omega_h^i)}$$

Therefore, from the inequality (2-174), we obtain:

$$\|F\|_{(H_p^k(\omega_h^i))'} \leq \|T\|_{(L_p(\omega_h^i))'} \cdot Ch^{-|\alpha|}, \text{ for all } 0 \leq |\alpha| \leq k. \quad \ldots \quad (2\text{-}175)$$

Next, we need the following result from Functional Analysis, which is nothing more than the formula (2-125) considered in the previous section:

$$\|D^\alpha (J_h u - u)\|_{L_p(\omega_h^i)} = \sup_{T \in (L_p(\omega_h^i))'} \frac{|T(D^\alpha (J_h u - u))|}{\|T\|_{(L_p(\omega_h^i))'}} \quad \ldots \quad (2\text{-}176)$$

Then, from the Bramble and Hilbert lemma combined with

the inequality (2-175), we get:

$$|F(v)| \leq \|F\|_{(W_p^k(\omega_h^i))'} \cdot \|v\|_{W_p^k(\omega_h^i)} \leq C_1 \|F\|_{(W_p^k(\omega_h^i))'} \cdot h^k |u|_{W_p^k(\omega_h^i)}$$

$$\leq C \|T\|_{(L_p(\omega_h^i))'} \cdot h^{k-|\alpha|} |u|_{W_p^k(\omega_h^i)}$$

or, from (2-171):

$$|T(D^\alpha(J_h u - v))| \leq C \|T\|_{(L_p(\omega_h^i))'} \cdot h^{k-|\alpha|} |u|_{W_p^k(\omega_h^i)}$$

Thus,

$$\frac{|T(D^\alpha(J_h u - v))|}{\|T\|_{(L_p(\omega_h^i))'}} \leq C h^{k-|\alpha|} |u|_{W_p^k(\omega_h^i)}$$

and, combined with (2-176):

$$\|D^\alpha(J_h u - u)\|_{L_p(\omega_h^i)} = \sup_{T \in (L_p(\omega_h^i))'} \frac{|T(D^\alpha(J_h u - u))|}{\|T\|_{L_p(\omega_h^i)}} \leq C h^{k-|\alpha|} |u|_{W_p^k(\omega_h^i)}$$

for all the integers α such that $0 \leq |\alpha| \leq k$. Therefore, for any point $x \in \Omega_i$ we have so far proved that:

$$\|D^\alpha(J_h u - u)\|_{L_p(\omega_h^i)} \leq C h^{k-|\alpha|} |u|_{W_p^k(\omega_h^i)} \quad \ldots \quad (2\text{-}177)$$

The estimate (2-170) now comes from the local estimate (2-177) upon summing up for all the hypercubes:

$$\Omega_i \subset R^n \text{ such that } \bigcup_i \Omega_i = R^n.$$

Therefore,

$$\sum_{\Omega_i} \|D^\alpha (J_h u - u)\|_{L_p(\omega_h^i)} \leq \sum_{\Omega_i} C h^{k-|\alpha|} |u|_{H_p^k(\omega_h^i)}, \quad \text{or}$$

$$\|D^\alpha (J_h u - u)\|_{L_p(R^n)} \leq C\, h^{k-|\alpha|}\, |u|_{H_p^k(R^n)}, \quad \text{for all } 0 \leq |\alpha| \leq k$$

which completes the proof for the inequality (2-170).

Finally, the following theorem gives an estimate for the error between an arbitrary function $u \in H_p^k(\Omega)$ and its smoothed version $J_h u$ constructed by convolution:

THEOREM III

Let u be any function from the space $H_p^k(\Omega)$, where Ω is any bounded and open subset of the Euclidean space R^n and $J_h u$ is its smoothed version which belongs to the space $C_0^\infty(R^n)$ and satisfies the conditions (i) and (ii) stated above. Then,

$$\|J_h u - u\|_{H_p^m(\Omega)} \leq C\, h^{k-m} |u|_{H_p^k(\Omega)} \qquad \ldots \quad (2\text{-}178)$$

for any integer $0 \leq m \leq k$, where C is some numerical constant which does not depend on the function u and the parameter h.

PROOF

Indeed, from (2-170) and any integer $0 \leq m \leq k$, we have:

$$\sum_{|\alpha|\le m} \| D^{\alpha}(J_h u - u) \|^p_{L_p(R^n)} \le \sum_{|\alpha|\le m} C h^{p(k-|\alpha|)} |u|^p_{H^k_p(R^n)}$$

and, by assuming that $0 < h \le 1$, we get:

$$\| J_h u - u \|^p_{H^m_p(R^n)} \le C\, h^{p(k-|\alpha|)} |u|^p_{H^k_p(R^n)}, \quad \text{or}$$

$$\| J_h u - u \|_{H^m_p(R^n)} \le C\, h^{k-m} |u|_{H^k_p(R^n)} \qquad \ldots \quad (2\text{-}179)$$

Next, since by definition:

$$\| J_h u - u \|_{H^m_p(\Omega)} \le \| J_h u - u \|_{H^m_p(R^n)}$$

for any bounded domain $\Omega \subset R^n$, from (2-179), we have:

$$\| J_h u - u \|_{H^m_p(\Omega)} \le C\, h^{k-m} |u|_{H^k_p(R^n)}, \quad \text{for all } 0 \le m \le k \qquad \ldots \quad (2\text{-}180)$$

It remains now to find an upper bound for the semi-norm $|u|_{H^k_p(R^n)}$ which is defined over the entire space R^n in terms of the same quantity, but defined this time only over the subdomain $\Omega \subset R^n$. According to Calderon's theorem (Agmon [1], p. 171) there exists a transformation ε of the space $H^k_p(\Omega)$ into the space $H^k_p(R^n)$ such that, for any function $u \in H^k_p(\Omega)$, the restriction of εu to Ω coincides with u, ie,

$$\varepsilon u = u, \text{ for all } u \in H^k_p(\Omega)$$

and, therefore, there exists a constant C such that:

$$\|\varepsilon u\|_{H_p^k(R^n)} \leq C\|u\|_{H_p^k(\Omega)}, \text{ for all } u \in H_p^k(\Omega) \quad \ldots \quad (2\text{-}181)$$

Nevertheless, the Calderon's inequality (2-181) involves the full norm of the function rather than the semi-norm which we actually need. However, following G. Strang [23] we can replace the extension operator ε above by the following:

$$\ldots = \varepsilon(I-\Pi_{k-1}) + E\Pi_{k-1} \quad \ldots \quad (2\text{-}182)$$

where E is the operator which extends a polynomial defined on Ω to its equivalent polynomial defined on R^n and Π_{k-1} is the projection operator of the space $H_p^k(\Omega)$ onto the space $P_{k-1}(\Omega)$ of all the polynomials of degree less than or equal to (k-1) defined over Ω.

Then, from (2-182), we get:

$$|\varepsilon_k u|_{H_p^k(R^n)} = |\varepsilon(I-\Pi_{k-1})u + E\Pi_{k-1}u|_{H_p^k(R^n)} = |\varepsilon(I-\Pi_{k-1})u|_{H_p^k(R^n)} \quad \ldots \quad (2\text{-}183)$$

since the semi-norm is always zero for polynomials of degree less than k. Thus, from (2-183) combined with (2-181), we obtain:

$$|\varepsilon_k u|_{H_p^k(R^n)} \leq \|\varepsilon(I-\Pi_{k-1})u\|_{H_p^k(R^n)} \leq C\|u-\Pi_{k-1}u\|_{H_p^k(\Omega)} \leq C|u|_{H_p^k(\Omega)}$$

where in the last inequality we have used the approximation result (2-143) of the theorem II, for m = k.

Finally, if we combine this result with the inequality (2-179), we obtain:

$$\|J_h u - u\|_{H_p^m(\Omega)} \leq C h^{k-m}|u|_{H_p^k(\Omega)}, \quad \text{for all } 0 \leq m \leq k$$

which completes the proof.

2.4.4 Some Useful Remarks

If we apply the two basic results (2-143) and (2-178) of Theorems II and III, respectively, over any simplex:

$$\bar{S}_n^\nu, \quad \nu \in I$$

of a partition Δ - satisfying the regularity condition (2-128) - of the domain Ω and by taking:

$$\hat{\Omega} = \hat{S}_n; \quad \Omega = \bar{S}_n^\nu, \quad \nu \in I$$

we have:

$$\|u-Pu\|_{W_p^m(\bar{S}_n^\nu)} \leq C\, h_\nu^{k-m} |u|_{W_p^k(\bar{S}_n^\nu)}, \text{ for all } 0 \leq m \leq k, \ \nu \in I \quad \ldots \quad (2\text{-}184)$$

and

$$\|u-J_h u\|_{W_p^m(\bar{S}_n^\nu)} \leq C\, h_\nu^{k-m} |u|_{W_p^k(\bar{S}_n^\nu)}, \text{ for all } 0 \leq m \leq k, \ \nu \in I \quad \ldots \quad (2\text{-}185)$$

where P is the following linear mapping:

$$P : W_p^k(\bar{S}_n^\nu) \to W_p^m(\bar{S}_n^\nu), \text{ for all } 0 \leq m \leq k, \ \nu \in I$$

ie,

$$P \in L(W_p^k(\bar{S}_n^\nu); W_p^m(\bar{S}_n^\nu)), \ \nu \in I$$

such that:

$$Pu = u, \text{ for all } u \in P_{k-1}(\bar{S}_n^\nu), \ \nu \in I$$

and J_h denotes the smoothing operator such that $J_h u$ is the smoothed version of any function:

$$u \in H_p^k(\bar{S}_n^\nu), \quad \nu \in I.$$

The numerical constant C, again, does not depend on u and h_ν, $\nu \in I$.

Next, what is naturally expected and indeed it turns out that this is the case, is the fact that the more general approximation schemes analyzed at the present and previous sections over the Sobolev spaces contain that of the Hermite approximation as a special case. For, if we consider any simplex:

$$\bar{S}_n^\nu, \quad \nu \in I$$

of the partition and any function:

$$u \in H_p^k(\bar{S}_n^\nu), \quad \nu \in I$$

such that the Hermite interpolating conditions (2-52) are satisfied and, furthermore, that:

$$\lambda < k-n/p \qquad \ldots \quad (2\text{-}186)$$

then, by vitue of the Sobolev imbedding theorem (2-99), the inequality (2-186) guarantees that:

$$H_p^k(\bar{S}_n^\nu) \subset C^\lambda(\bar{S}_n^\nu), \quad \nu \in I$$

and, therefore, the Hermite interpolating polynomial $u_n^{(k-1)}(x)$ is indeed well defined. However, since the derivatives of the function:

$$u \in H_p^k(\bar{S}_n^\nu), \quad \nu \in I$$

are unlikely to exist in the pointwise sense and the Taylor expansion for u is, therefore, in general meaningless,

we have got to make use of the results from the Functional Analysis as well as that of the Lemma of Bramble and Hilbert in order to obtain the following approximation estimate:

$$\| u(x) - u_n^{(k-1)}(x) \|_{H_p^m(\bar{S}_n^\nu)} \leq C\, h_\nu^{k-m} |u|_{H_p^k(\bar{S}_n^\nu)}, \quad \text{for all } 0 \leq m \leq k,\ \nu \in I \qquad \ldots \quad (2\text{-}187)$$

for any function $u \in H_p^k(\bar{S}_n^\nu)$, $\nu \in I$, where the numerical constant C does not depend on the particular function u and the geometrical parameter h_ν. Whenever the inequality (2-186) above, is not satisfied, instead, the interpolating polynomial $u_n^{(k-1)}(x)$ for any function $u \in H_p^k(\bar{S}_n^\nu)$, $\nu \in I$, does not exist and we are faced with the problem, first of smoothing the function u considerably and then applying to the smoothed function $J_h u$ the pointwise approximation scheme described in the section 2.2.

2.5 MULTIVARIATE APPROXIMATION OVER CURVED ELEMENTS

For sake of completeness in our analysis of the approximation problem, we shall simply outline in this section an approximation technique over curved finite elements of isoparametric type which is a generalization of the case considered earlier, in paragraph 2.1.4, over the straight quadrilateral finite element in the plane. An extensive analysis of this technique, from which this brief presentation is taken, is given by P.G. Ciarlet and P.A. Raviart (see [8] or [3] pp 409-474) where a number of applications

are also described to several types of elements which are extremely useful from the practical point of view.

Therefore, suppose that the domain Ω is a bounded and open subset of the space R^n whose boundary $\partial\Omega$ now is curved - ie, the closed domain $\bar{\Omega} = \Omega + \partial\Omega$ is no more a polyhedron as was assumed to be the case in our previous analysis - and an elliptic boundary-value problem is described over Ω. Then, again, its solution u can always be approximated by the Finite Element Method: the domain $\bar{\Omega}$ is replaced by a finite union of finite elements which for simplicity are assumed to be of simplicial type, the case for hypercubic being similar under only minor alterations being made on the several assumptions. Then, it is obvious that for those simplices which are situated in the interior of the domain and have no point in common with the boundary $\partial\Omega$ are assumed to have plane faces, whereas for those which share some points with the boundary $\partial\Omega$ are generally assumed to be curved.

Let us, once more, consider a partition:

$$\Delta : \{\bar{S}_n^\nu\}, \quad \nu \in I$$

of the domain Ω into n-simplices (being straight in the interior and curved near the boundary) and let:

$$\bar{S}_n^\nu, \quad \nu \in I$$

be any simplex of the partition. With the simplex $\bar{S}_n^\nu$, $\nu \in I$, we associate its discrete set of interpolating points:

$$S_n^{\nu}(k-1),\ \nu \in I,\ k \geq 1$$

which contains exactly, as before,

$$N = \binom{k+n-1}{k-1}$$

members. Then, once and for all, with every simplex $\bar{S}_n^{\nu}$, $\nu \in I$, of the partition Δ, we associate a straight *reference simplex* $\hat{\bar{S}}_n$ which is related to the simplices $\bar{S}_n^{\nu}$, $\nu \in I$, through the following mapping:

$$F_{\nu}:\ \hat{x} \in \hat{\bar{S}}_n \rightarrow F_{\nu}(\hat{x}) \in R^n \qquad \ldots\ (2\text{-}188)$$

with:

$$F_{\nu}(\hat{x}) = (F_{\nu,1}(\hat{x}),\ F_{\nu,2}(\hat{x}),\ldots,F_{\nu,n}(\hat{x}))$$

such that:

$$F_{\nu}(\hat{\bar{S}}_n) = \bar{S}_n^{\nu},\ \nu \in I$$

ie, the simplex $\bar{S}_n^{\nu}$, $\nu \in I$, is the image of the reference simplex $\hat{\bar{S}}_n$ under the mapping (2-188) and, in particular:

$$F_{\nu}(\hat{a}_i) = a_i^{\nu},\ 1 \leq i \leq N,\ \nu \in I$$

where $\hat{a}_i \in \hat{S}_n(k-1)$ and $\hat{S}_n(k-1)$ denotes the corresponding discrete set of interpolating points which is attached to the reference simplex $\hat{\bar{S}}_n$. There are two possibilities as far as the mapping (2-188) is concerned: in the first, the mapping is linear and the simplex $\bar{S}_n^{\nu}$, $\nu \in I$, is straight. This situation corresponds to the approximation problem considered in the whole of our previous discussion. In the second, the mapping F_{ν} is not linear and in that case the simplex $\bar{S}_n^{\nu}$, $\nu \in I$, is curved.

With every function:

$$\hat{u}: \hat{\bar{S}}_n \to R$$

we associate a function:

$$u: \bar{S}_n^\nu \to R, \ \nu \in I$$

defined by:

$$u(x) = \hat{u} o F^{-1}(x), \text{ for all } x \in \bar{S}_n^\nu, \ \nu \in I$$

and vice versa; ie, with every function:

$$u: \bar{S}_n^\nu \to R, \ \nu \in I$$

we associate a function:

$$\hat{u}: \hat{\bar{S}}_n \to R$$

defined by:

$$\hat{u}(\hat{x}) = u o F(\hat{x}), \text{ for all } \hat{x} \in \hat{\bar{S}}_n.$$

By using now the same notation as before together with the fact that in the isoparametric technique one always uses the same polynomials to express the coordinate transformation between the simplices $\hat{\bar{S}}_n$ and $\bar{S}_n^\nu$, $\nu \in I$, (which is carried out through the mapping (2-188)) as well as the interpolating function over each simplex $\bar{S}_n^\nu$, $\nu \in I$, the mapping F_ν for the special cases corresponding to the Lagrange and Hermite interpolation problems, respectively, are given as follows:

(i) For the Lagrange problem:

$$F_\nu = \sum_{i=1}^{N} \hat{\phi}_i(\hat{x}) a_i^\nu, \ 1 \le i \le N, \ \nu \in I \qquad \ldots (2\text{-}189)$$

where $\hat{\phi}_i(\hat{x})$, $1 \le i \le N$, are the basis functions over the reference (straight) n-simplex $\hat{S}_n$ and they may always be expressed in terms of the barycentric coordinates of an arbitrary point $\hat{x} \in \hat{S}_n$ with respect to the (n+1) vertices $\hat{a}_i$, $1 \le i \le n+1$, of $\hat{S}_n$. Thus, obviously:

$$\hat{\phi}_i(\hat{a}_j) = \delta_{ij},\ 1 \le i,j \le N$$

where $\hat{a}_j \in \hat{S}_n(k-1)$, $1 \le j \le N$.

(ii) For the Hermite problem:

$$F_\nu = \sum_{i=1}^{N_0} \hat{\phi}_i^0(\hat{x}) a_i^{\nu,0} + \sum_{i=1}^{N_1} \sum_{\ell=1}^{\nu_i} \hat{\phi}_{i\ell}^1(\hat{x})\, \xi_{i\ell}^{\nu,1} + \ldots + \sum_{i=1}^{N_\lambda} \sum_{\ell=1}^{\nu_i} \hat{\phi}_{i\ell}^\lambda(\hat{x})\, \xi_{i\ell}^{\nu,\lambda} \qquad \ldots \quad (2\text{-}190)$$

where the basis functions:

$$\hat{\phi}_i^0(\hat{x}),\ 1 \le i \le N_0$$

$$\hat{\phi}_{i\ell}^\mu(\hat{x}),\ 1 \le i \le N_\mu,\ 1 \le \mu \le \lambda,\ 1 \le \ell \le \nu_i$$

satisfy the conditions (2-54), or, equivalently, they are uniquely determined through the conditions (2-54), and:

$$a_i^{\nu,0} = F_\nu(\hat{a}_i^0),\ 1 \le i \le N_0,\ \nu \in I$$

and

$$\xi_{i\ell}^{\nu,\mu} = \mathcal{D}^\mu F_\nu(a_i^{\nu,\mu})\hat{\xi}_{i\ell}^\mu,\ 1 \le i \le N_\mu,\ 1 \le \mu \le \lambda,\ 1 \le \ell \le \nu_i,\ \nu \in I.$$

Next, the following bound serves as a first measure for the difference between an arbitrary function u defined over any simplex S_n^ν, $\nu \in I$, of the partition Δ and its

interpolating polynomial generally denoted by Pu:

$$|u-Pu|_{H_p^m(\tilde{S}_n^\nu)} \le C \max_{\hat{x}\in\hat{S}_n} |J_{F_\nu}(\hat{x})|^{1/p} \sum_{k=1}^{m} \sum_{i\in I(k,m)} \sup_{x\in\tilde{S}_n^\nu} [\|DF_\nu^{-1}(x)\|^{i_1} \cdot$$

$$\cdot \|D^2F_\nu^{-1}(x)\|^{i_2} \cdots \|D^mF_\nu^{-1}(x)\|^{i_m}] \cdot \frac{1}{\min_{\hat{x}\in\hat{S}_n} |J_{F_\nu}(\hat{x})|^{1/p}} \cdot \sum_{k=1}^{k} |u|_{H_p^k(\tilde{S}_n^\nu)} \cdot$$

$$\cdot \sum_{j\in I(k,k)} \sup_{\hat{x}\in\hat{S}_n} [\|DF_\nu(\hat{x})\|^{j_1} \|D^2F_\nu(\hat{x})\|^{j_2} \cdots \|D^kF_\nu(\hat{x})\|^{j_k}], \quad \nu \in I \quad \ldots (2\text{-}191)$$

where P denotes the following mapping between the Sobolev spaces $H_p^k(\tilde{S}_n^\nu)$, $\nu \in I$, and $H_p^m(\tilde{S}_n^\nu)$, $\nu \in I$:

$$P: u \in H_p^k(\tilde{S}_n^\nu) \to Pu \in H_p^m(\tilde{S}_n^\nu), \quad \nu \in I$$

for some integer m with $0 \le m \le k$ and is such that:

$$Pu = u, \text{ for all } u \in P_{k-1}(\tilde{S}_n^\nu), \quad \nu \in I$$

Moreover, $|J_{F_\nu}(\hat{x})|$ denotes the Jacobian matrix of the mapping F_ν defined as in (2-188) at the particular point $\hat{x} \in \hat{S}_n$, $I(\ell,m)$ and $I(\ell,k)$ are the following sets of points respectively:

$$I(\ell,m) = \{i = (i_1, i_2, \ldots, i_m) \in N^m : \sum_{\nu=1}^{m} i_\nu = \ell, \ \sum_{\nu=1}^{m} \nu i_\nu = m\}, \ 1 \le \ell \le m$$

and

$$I(\ell,k) = \{j = (j_1, j_2, \ldots, j_k) \in N^k : \sum_{\nu=1}^{k} i_\nu = \ell, \ \sum_{\nu=1}^{m} \nu j_\nu = k\}, \ 1 \le \ell \le k$$

and C is some numerical constant which does not depend

on the function u and the geometry of the simplex $\bar{S}_n^\nu$, $\nu \in I$. Again, for the special cases of the Lagrange and Hermite interpolation problems and in accordance with the formulae (2-189) and (2-190), the corresponding interpolating polynomials are given, respectively, by the following functions:

$$Pu \approx \sum_{i=1}^{N} u(a_i^\nu)\phi_i^\nu(x), \quad \nu \in I$$

and,

$$Pu \approx \sum_{i=1}^{N_0} u(a_i^{\nu,0})\phi_i^{\nu,0}(x) + \sum_{i=1}^{N_1}\sum_{\ell=1}^{\nu_i} [Du(a_i^{\nu,1})\cdot\xi_{i\ell}^{\nu,1}]\phi_{i\ell}^{\nu,1}(x) + \ldots$$

$$\ldots + \sum_{i=1}^{N_\lambda}\sum_{\ell=1}^{\nu_i} [D^\lambda u(a_i^{\nu,\lambda})\cdot\xi_{i\ell}^{\nu,\lambda}]\phi_{i\ell}^{\nu,\lambda}(x), \quad \nu \in I$$

where,

$$a_i^{\nu,\mu} = F_\nu(\hat{a}_i^\mu), \quad 1 \le i \le N_\mu, \quad 1 \le \mu \le \lambda, \quad \nu \in I.$$

All that is needed now, in order to determine the asymptotic order of convergence between the functions $u \in H_p^k(\bar{S}_n^\nu)$, $\nu \in I$, and $Pu \in H_p^m(\bar{S}_n^\nu)$, $\nu \in I$, for some integer m such that $0 \le m \le k$, is to find an upper bound for the various derivatives of the mapping F_ν - as well as its inverse F_ν^{-1} - in (2-191) under the assumption that F_ν is a C^k-diffeomorphism over the simplex $\bar{S}_n^\nu$, $\nu \in I$, (ie, the mappings F_ν and F_ν^{-1} are of class $C^k(\bar{S}_n^\nu)$, $\nu \in I$) as well as an upper bound for the quantity:

$$\max_{\hat{x} \in \hat{\bar{S}}_n} |J_{F_\nu}(\hat{x})| \Big/ \min_{\hat{x} \in \hat{\bar{S}}_n} |J_{F_\nu}(\hat{x})|$$

This problem, however, is considerably simplified by introducing the following auxiliary linear mapping:

$$F_\nu^* : \hat{x} \in \hat{\bar{S}}_n \to F_\nu^*(\hat{x}) \in R^n \qquad \ldots \quad (2\text{-}192)$$

where:

$$F_\nu^*(\hat{x}) = (F_{\nu,1}^*(\hat{x}),\ F_{\nu,2}^*(\hat{x}),\ \ldots,\ F_{\nu,n}^*(\hat{x}))$$

such that, if we assume that the first (n+1) points α_i, $1 \le i \le n+1$, of the set $S_n^\nu(k-1)$, $\nu \in I$, constitute the (n+1) vertices of the simplex $\bar{S}_n^\nu$, $\nu \in I$, we always have:

$$F_\nu^*(\hat{\alpha}_i) = \alpha_i^\nu, \quad 1 \le i \le n+1, \quad \nu \in I$$

and for all the other points:

$$F_\nu^*(\hat{\alpha}_i) = {}^*\alpha_i^\nu, \quad n+2 \le i \le N, \quad \nu \in I$$

Then, the image of the reference simplex $\hat{\bar{S}}_n$ under the linear mapping (2-192) defines the straight simplex:

$$\bar{\bar{S}}_n^\nu = F_\nu^*(\hat{\bar{S}}_n), \quad \nu \in I$$

which shares the same vertices with the simplex $\bar{S}_n^\nu$, $\nu \in I$, and the non-linear (in general) mappings (2-189) and (2-190) are respectively given as follows, in terms of the linear mapping (2-192)

$$F_\nu = F_\nu^* + \sum_{i=1}^{N} \hat{\phi}_i(\hat{x})\,(\alpha_i - {}^*\alpha_i^\nu), \quad 1 \le i \le N, \quad \nu \in I \qquad \ldots \quad (2\text{-}193)$$

where:

$${}^*\alpha_i^\nu = F_\nu^*(\hat{\alpha}_i), \quad 1 \le i \le N, \quad \nu \in I$$

and

$$F_\nu = F_\nu^* + \sum_{i=1}^{N_0} \hat{\phi}_i^0(\hat{x})\,(a_i^{\nu,0} - {}^*a_i^{\nu,0}) + \sum_{i=1}^{N_1}\sum_{\ell=1}^{\nu_i} \hat{\phi}_{i\ell}^1(\hat{x})\,(\xi_{i\ell}^{\nu,1} - {}^*\xi_{i\ell}^{\nu,1}) + \ldots$$

$$\ldots + \sum_{i=1}^{N_\lambda}\sum_{\ell=1}^{\nu_i} \hat{\phi}_{i\ell}^\lambda(\hat{x})\,(\xi_{i\ell}^{\nu,\lambda} - {}^*\xi_{i\ell}^{\nu,\lambda}) \qquad \ldots \text{(2-194)}$$

where:

$${}^*a_i^{\nu,0} = F_\nu^*(\hat{a}_i^0), \quad 1 \le i \le N_0, \quad \nu \in I$$

and:

$${}^*\xi_{i\ell}^{\nu,\mu} = D^\mu F_\nu^*(\hat{a}_i^\mu)\hat{\xi}_{i\ell}^\mu, \; 1 \le i \le N_\mu, \; 1 \le \mu \le \lambda, \; 1 \le \ell \le \nu_i, \; \nu \in I$$

We note, however, that whenever the simplex $\bar{S}_n^\nu$, $\nu \in I$, is not curved, ie, the mapping F_ν defined by (2-188) is linear, then, we obviously have:

$$\bar{S}_n^\nu \equiv {}^*\bar{S}_n^\nu, \quad \nu \in I$$

Next, with every such simplex ${}^*\bar{S}_n^\nu$, $\nu \in I$, we associate the following two geometrical parameters:

$$h_\nu = \text{diameter of } {}^*\bar{S}_n^\nu, \quad \nu \in I$$
$$\rho_\nu = \text{diameter of the inscribed sphere in } {}^*\bar{S}_n^\nu, \quad \nu \in I \qquad \ldots \text{(2-195)}$$

Then, for a regular partition Δ such that the regularity condition (2-128) is satisfied, we can obtain the asymptotic order of convergence over the simplices $\bar{S}_n^\nu$, $\nu \in I$, provided that they do not differ too much from the simplices ${}^*\bar{S}_n^\nu$, $\nu \in I$. Since, by definition of the linear mapping F_ν^*, we always have for the first (n+1) points that:

$$a_i^\nu - {}^*a_i^\nu = 0, \quad 1 \le i \le n+1, \quad \nu \in I$$

a natural way of measuring this deviation between the simplices $\hat{S}_n^\nu$, $\nu \in I$ and $\overset{*}{S}{}_n^\nu$, $\nu \in I$, is to consider the difference:

$$\alpha_i^\nu - \overset{*}{\alpha}{}_i^\nu, \quad n+2 \leq i \leq N, \quad \nu \in I$$

where again:

$$\overset{*}{\alpha}{}_i^\nu = F_\nu^*(\hat{a}_i), \quad n+2 \leq i \leq N, \quad \nu \in I$$

However, for the Jacobian of the mapping F_ν, the following result holds:

$$C_1 |J_{F_\nu^*}(\hat{x})| \leq |J_{F_\nu}(\hat{x})| \leq C_2 |J_{F_\nu^*}(\hat{x})| \qquad \ldots \quad (2\text{-}196)$$

with the constants C_1 and C_2 explicitly given in [8], and for the derivatives of the mapping F_ν - as well as those of its inverse F_ν^{-1} - under the assumption that the linear mapping F_ν^* is C^k-diffeomorphism:

$$\sup_{\hat{x} \in \hat{S}_n} \| \mathcal{D}^j F_\nu(\hat{x}) \| \leq K \sup_{\hat{x} \in \hat{S}_n} \| \mathcal{D} F_\nu^*(\hat{x}) \|^j, \quad 1 \leq j \leq k, \quad \nu \in I \qquad \ldots \quad (2\text{-}197)$$

and,

$$\sup_{x \in \hat{S}_n^\nu} \| \mathcal{D}^j F_\nu^{-1}(x) \| \leq L \sup_{\overset{*}{x} \in \overset{*}{S}{}_n^\nu} \| \mathcal{D} F_\nu^{*-1}(\overset{*}{x}) \|^j, \quad 1 \leq j \leq k, \quad \nu \in I \qquad \ldots \quad (2\text{-}198)$$

with, again, the constants K and L explicitly calculated and given in [8]. Then, for a regular partition Δ of the domain such that the geometrical parameters defined by (2-195) satisfy the inequality (2-128), we have that:

$$\sup_{\hat{x} \in \hat{S}_n} \| \mathcal{D} F_\nu^*(\hat{x}) \| = O(h_\nu), \quad \nu \in I \qquad \ldots \quad (2\text{-}199)$$

a natural way of measuring this deviation between the simplices $\hat{S}_n^\nu$, $\nu \in I$ and $\overset{*}{S}{}_n^\nu$, $\nu \in I$, is to consider the difference:

$$\alpha_i^\nu - \overset{*}{\alpha}{}_i^\nu, \quad n+2 \leq i \leq N, \quad \nu \in I$$

where again:

$$\overset{*}{\alpha}{}_i^\nu = F_\nu^*(\hat{a}_i), \quad n+2 \leq i \leq N, \quad \nu \in I$$

However, for the Jacobian of the mapping F_ν, the following result holds:

$$C_1 |J_{F_\nu^*}(\hat{x})| \leq |J_{F_\nu}(\hat{x})| \leq C_2 |J_{F_\nu^*}(\hat{x})| \qquad \ldots \quad (2\text{-}196)$$

with the constants C_1 and C_2 explicitly given in [8], and for the derivatives of the mapping F_ν - as well as those of its inverse F_ν^{-1} - under the assumption that the linear mapping F_ν^* is C^k-diffeomorphism:

$$\sup_{\hat{x} \in \hat{S}_n} \| D^j F_\nu(\hat{x}) \| \leq K \sup_{\hat{x} \in \hat{S}_n} \| D^j F_\nu^*(\hat{x}) \|^j, \quad 1 \leq j \leq k, \quad \nu \in I \qquad \ldots \quad (2\text{-}197)$$

and,

$$\sup_{x \in \overset{*}{S}{}_n^\nu} \| D^j F_\nu^{-1}(x) \| \leq L \sup_{\hat{x} \in \overset{*}{S}{}_n^\nu} \| D F_\nu^{*-1}(\hat{x}) \|^j, \quad 1 \leq j \leq k, \quad \nu \in I \qquad \ldots \quad (2\text{-}198)$$

with, again, the constants K and L explicitly calculated and given in [8]. Then, for a regular partition Δ of the domain such that the geometrical parameters defined by (2-195) satisfy the inequality (2-128), we have that:

$$\sup_{\hat{x} \in \hat{S}_n} \| D F_\nu^*(\hat{x}) \| = O(h_\nu), \quad \nu \in I \qquad \ldots \quad (2\text{-}199)$$

and

$$\sup_{\overset{*}{x}\in\overset{*}{S}_n^\nu} \| DF_\nu^{*-1}(\overset{*}{x}) \| = O(h_\nu^{-1}), \quad \nu \in I \qquad \ldots \quad (2\text{-}200)$$

Finally, by combining the results (2-196), (2-197), (2-198), (2-199) and (2-200) with the estimate (2-191), we get the following assymtotic estimate:

$$|u-Pu|_{H_p^m(\bar{S}_n^\nu)} \leq C\, h_\nu^{k-m} |u|_{H_p^k(\bar{S}_n^\nu)}, \quad \nu \in I \qquad \ldots \quad (2\text{-}201)$$

for any integer m with $0 \leq m \leq k$, which gives the exact order of convergence over any simplex $\bar{S}_n^\nu$, $\nu \in I$, of the partition Δ. The numerical constant C, moreover, which does not depend on the function u and the parameter h_ν, $\nu \in I$, is, of course, very difficult to compute explicitly, but nevertheless this fact does not lower the importance of the result (2-201) which simply expresses that the same order can be obtained in the approximation, either over straight elements or over curved elements. Of course, one can easily notice that in the second part of (2-201) the full norm for the function u appears instead of the k-th order derivatives only and this is due to the fact that one has to introduce, first the reference simplex $\hat{S}_n$ and then go back to the original simplex $\hat{S}_n^\nu$, $\nu \in I$.

PART III

ERROR ESTIMATES FOR THE FINITE ELEMENT APPROXIMATION

3.1 ERROR ANALYSIS FOR CONFORMING FINITE ELEMENTS

3.1.1 A Model Problem

In this third and final part of our analysis we apply the previously obtained results on approximation in order to determine the correct order of convergence of the finite element approximation u_h to the exact solution u of the particular problem under consideration. Therefore, let us consider the following n-variable elliptic boundary-value problem of order 2m as a general model problem defined on a bounded and open subset Ω of R^n:

$$A.u(x) = \sum_{|\alpha|,|\beta|\leq m} (-1)^{|\beta|} D^{\beta}[q_{\alpha\beta}(x) D^{\alpha}u(x)] = f(x), \; x \in \Omega \quad \ldots (3\text{-}1)$$

subject to the Dirichlet conditions:

$$D^{\beta}u(x) = 0, \text{ for all } x \in \partial\Omega, \; |\beta| \leq m-1 \quad \ldots (3\text{-}2)$$

where the functions $q_{\alpha\beta}(x)$ are required to satisfy certain continuity conditions over the closed domain $\bar{\Omega}$; viz:

$$q_{\alpha\beta}(x) \in C^{\beta}(\bar{\Omega})$$

and are such that, there exists a positive constant C_1, with:

$$\sum_{|\alpha|,|\beta|\leq m} \int_{\Omega} q_{\alpha\beta}(x) D^{\alpha}u(x) D^{\beta}u(x)\,dx \geq C_1 \|u\|^2_{H^m(\Omega)} \qquad \ldots \quad (3\text{-}3)$$

Then, by integrating the inner-product term $(Au,v)_{L_2(\Omega)}$ by parts (see eg. P.M. Prenter [19] p.266), we get the following result:

$$a(u,v) = \sum_{|\alpha|,|\beta|\leq m} \int_{\Omega} q_{\alpha\beta}(x) D^{\alpha}u(x) D^{\beta}v(x)\,dx + \gamma(u,v), \quad u,v \in V \subset H^m(\Omega)$$

where the expression $\gamma(u,v)$, generally a surface integral, vanishes because of the boundary conditions (3-2) which have to be satisfied by the functions $u,v \in V$. Therefore, the following formula:

$$a(u,v) = \sum_{|\alpha|,|\beta|\leq m} \int_{\Omega} q_{\alpha\beta}(x) D^{\alpha}u(x) . D^{\beta}v(x)\,dx \qquad \ldots \quad (3\text{-}4)$$

defines a bilinear form on the product-space $V \times V$ and, because of the inequality (3-3), it satisfies:

$$|a(u,u)| \geq C_1 \|u\|^2_{H^m(\Omega)}, \quad u \in V \qquad \ldots \quad (3\text{-}5)$$

for some positive constant C_1, which expresses the ellipticity condition of the problem. On the other hand, one can easily see that:

$$|a(u,u)| = \Big| \int_{\Omega} \sum_{|\alpha|,|\beta|\leq m} q_{\alpha\beta}(x) D^{\alpha}u(x) D^{\beta}u(x)\,dx \Big| \leq$$

$$\leq \max_{|\alpha|,|\beta|\leq m} \max_{x\in\Omega} |q_{\alpha\beta}(x)| \sum_{|\alpha|,|\beta|\leq m} \int_{\Omega} D^{\alpha}u(x) . D^{\beta}u(x)\,dx = C_2 \|u\|^2_{H^m(\Omega)} . \text{ Thus,}$$

$$|a(u,u)| \leq C_2 \|u\|^2_{H^m(\Omega)}, \quad u \in V \qquad \ldots (3\text{-}6)$$

where the numerical constant $C_2 = \max_{|\alpha|,|\beta| \leq m} \max_{x \in \Omega} |q_{\alpha\beta}(x)|$.

Inequality (3-6) expresses the boundedness condition and, by combining the two inequalities (3-5) and (3-6), we get the following basic result:

$$C_1 \|u\|^2_{H^m(\Omega)} \leq |a(u,u)| \leq C_2 \|u\|^2_{H^m(\Omega)} \qquad \ldots (3\text{-}7)$$

for some numerical constants C_1 and C_2, with $C_1 > 0$. The double inequality (3-7) simply states, by the standard definition of the equivalence norms, that the two norms $\|u\|_{H^m(\Omega)}$ and $a(u,u)$ are equivalent. The importance of that result rests upon the fact that, once an estimate for the error in $u - u_h$ has at first been derived in terms of the energy norm $a(u,u)$ - this always is the case since the Ritz method is minimizing the appropriate functional with respect to that kind of norm - then, we shall use that inequality in order to compute an equivalent estimate for the error, this time, involving the H^m-norm.

3.1.2 Admissible Triangulation of $\bar{\Omega}$

In order to outline once more the essential points of our previous analysis as well as to generalize them to include any kind of finite element instead of the n-simplex which was considered in the approximation problem of the previous part, suppose that the closed domain $\bar{\Omega}$ has been

replaced by a finite union of finite elements e_i such that:

$$\bar{\Omega} = \bigcup_{e_i \in T_h} e_i$$

We say that this union of elements consists of an admissible partition or an admissible generalized triangulation, usually denoted by T_h, of Ω if and only if it is such that: if e_i and $e_j \in T_h$ are two elements of the partition T_h, then, $e_i = e_j$ or $e_i \cap e_j = \emptyset$ or $e_i \cap e_j =$ a common vertex or a common side.

With every such element e_i of the partition T_h, we associate the following two geometrical parameters:

h_{e_i} = diameter of the element e_i . . . (3-8)

ρ_{e_i} = sup {diameter of the inscribed sphere in e_i}

and we let:

$$h = \max_{e_i \in T_h} h_{e_i} \quad . . . \quad (3\text{-}9)$$

Then, we say that a sequence $\{T_h\}$, $h \in H$, where H denotes some collection of positive parameters h, of generalized triangulations of the domain Ω constitute a regular family if and only if the following two conditions are satisfied:

(i) the parameter h, defined by (3-9), approaches 0, and

(ii) $$0 \leq a \leq \min_{e_i \in T_h} \frac{\rho_{e_i}}{h_{e_i}} \quad . . . \quad (3\text{-}10)$$

for all the generalized triangulations T_h of the sequence

$\{T_h\}$, $h \in H$, where the constant a does not depend on the parameter h. The condition (3-10) can be regarded as a generalized form of the more particular angle condition assumed by Zlamal [28] or that of the uniformity condition of G. Strang [23], which for the case where $n = 2$ simply states that: there exists an angle $\theta_0 > 0$ such that the interior smallest angle θ found in all the triangles of the given triangulation T_h, satisfies:

$$\theta \geq \theta_0 \qquad \ldots \quad (3\text{-}11)$$

Next, for any such admissible triangulation of the domain Ω, the method is applied, at first, over each element e_i at a time and, then, a result concerning the entire domain Ω can be derived by computing the several contributions from within all the elements.

3.1.3 The Error in the Energy Norm

For the error analysis the starting point is always the same: If the function u_h denotes the finite element approximation and u the exact solution of the problem, it is well-known that:

$$a(u-u_h, u-u_h) = \min_{v_h \in S_h} a(u-v_h, u-v_h) \qquad \ldots \quad (3\text{-}12)$$

ie, the approximate solution u_h *obtained over the subspace* $S_h \in V$ is closer, in terms of the energy norm, to the exact solution u than any other *function* v_h *of the space* S_h.

Then, since the problem of convergence of the Finite Element Method is a question in pure approximation theory in estimating the distance between the function u and the subspace S_h, or particularly between the function u and its interpolating polynomial, which is a member of the space S_h, the essence of the entire analysis on the approximation problem of the previous Part, is that: *given a function* $u \in H^k(e_i)$, *restricted over an element* $e_i \in T_h$ at a time, there exists a unique interpolating polynomial Pu such that:

$$\|u-Pu\|_{H^m(e_i)} \leq C\, h_{e_i}^{k-m} |u|_{H^k(e_i)}, \quad \text{for any } 0 \leq m \leq k \qquad \ldots (3\text{-}13)$$

for some numerical constant C which depends neither on the function u nor on the geometrical parameter h_{e_i}, defined by (3-8). By adding the inequalities (3-13) over all the elements e_i into which the domain Ω has been partitioned, together with (3-9), we obtain:

$$\|u-Pu\|_{H^m(\Omega)} \leq C\, h^{k-m} |u|_{H^k(\Omega)}, \quad \text{for any } 0 \leq m \leq k \qquad \ldots (3\text{-}14)$$

Here, again, the constant C does not depend on the function u and the parameter h. Then, the approximation error estimate (3-14) together with the minimum principle described by (3-12), will give us the starting point in order to give several error bounds for the difference $u-u_h$. Indeed, from (3-12), we easily get that:

$$a(u-u_h, u-u_h) \leq a(u-v_h, u-v_h), \quad \text{for all } v_h \in S_h$$

and, by taking $v_h \equiv Pu \in S_h$, that:

$$\alpha(u-u_h, u-u_h) \leq \alpha(u-Pu, u-Pu) \qquad \ldots (3\text{-}15)$$

and we have now to compute an upper bound for the energy in the difference between the function u and its interpolating polynomial Pu. For that, we make use of the definition of the bilinear form $\alpha(u,v)$ which is given by the formula (3-4); ie,

$$\alpha(u-Pu, u-Pu) = \sum_{|\alpha|,|\beta| \leq m} \int_\Omega q_{\alpha\beta}(x) D^\alpha(u-Pu) D^\beta(u-Pu)\, dx \leq$$

$$\leq \max_{|\alpha|,|\beta| \leq m} \max_{x \in \Omega} |q_{\alpha\beta}(x)| \sum_{|\alpha|,|\beta| \leq m} \int_\Omega D^\alpha(u-Pu)\, D^\beta(u-Pu)\, dx = C' \| u-Pu \|^2_{H^m(\Omega)}$$

where the constant:

$$C' = \max_{|\alpha|,|\beta| \leq m} \max_{x \in \Omega} |q_{\alpha\beta}(x)|$$

Thus,

$$\alpha(u-Pu, u-Pu) \leq C' \| u-Pu \|^2_{H^m(\Omega)} \qquad \ldots (3\text{-}16$$

and we make use now of the approximating result (3-14). Therefore, by combining the inequalities (3-16) and (3-14), we have:

$$\alpha(u-Pu, u-Pu) \leq C\, h^{2(k-m)} |u|^2_{H^k(\Omega)} \qquad \ldots (3\text{-}17)$$

which is exactly an upper bound for the difference u-Pu in energy. Next, from (3-15) together with (3-17), we

can obtain a first bound for the error in $u-u_h$ in terms of the energy norm:

$$a(u-u_h,u-u_h) \leq C\ h^{2(k-m)}\ |u|_{H^k(\Omega)} \qquad \ldots \quad (3\text{-}18)$$

for any function $u \in H^k(\Omega)$, where the numerical constant C, as usual, does not depend on u and the parameter h and the integer m here denotes the order of the highest derivative which is involved in the bilinear form $a(u,v)$ associated with the problem (3-1) - (3-2).

3.1.4 The Constant Strain Condition

There is an important remark which has to be emphasized as far as the convergence of the method is concerned: from the error estimate (3-18) one can easily see that convergence does occur when and only when:

$$k > m \qquad \ldots \quad (3\text{-}19)$$

Inequality (3-19), in other words, means that any solution u which is a polynomial of degree m should exactly be reproduced by the finite element method. This condition is known, among the engineers, as the *constant strain condition* and its validity has gradually been established from the numerical failures which were experienced whenever it was violated.

3.1.5 The Error in the H^m-norm

Once the error bound (3-18) has been established, it is an easy task to give an equivalent estimate in terms, this time, of the H^m-norm. Indeed, from the ellipticity condition (3-5), we get:

$$C_1 \| u-u_h \|^2_{H^m(\Omega)} \leq a(u-u_h, u-u_h) \quad \ldots \quad (3\text{-}20)$$

or, from (3-20) combined with (3-18), we obtain:

$$\| u-u_h \|^2_{H^m(\Omega)} \leq C\, h^{2(k-m)} |u|^2_{H^k(\Omega)} \quad \ldots \quad (3\text{-}21)$$

for any function $u \in H^k(\Omega)$ and some numerical constant C which does not depend on the function u and the space parameter h. Notice, however, that since convergence in the energy norm essentially means convergence in the m-th derivative of the finite element approximation u_h to that of the exact solution u, this derivative is something special.

3.1.6 The Nitsche Trick

A different and somewhat more difficult problem arises from the possibility of estimating the rate of convergence in the s-th derivative, or in terms of the H^s-norm, where s may be smaller or larger than m. This approach requires the application of an elegant variational argument which was derived simultaneously by Aubin and Nitsche

and is simply known as the *Nitsche trick*. Following G. Strang and G. Fix [24], we shall prove that the correct order of convergence is not of h^{k-s}, but it must also depend on the integer m as well as on k and s. The whole strategy in proving that consists of introducing an auxiliary differential equation of the same order 2m:

$$L\,\omega = g \qquad \ldots \quad (3\text{-}22)$$

which, through the equation of the vanishing of the first variation, can be written as:

$$a(\omega,v) = (g,v)_{L_2(\Omega)}, \text{ for all } v \in V \subset H^m(\Omega) \qquad \ldots \quad (3\text{-}23)$$

Then, for $v = u-u_h$ in (3-23), we get:

$$a(\omega,u-u_h) = (g,u-u_h)_{L_2(\Omega)} \qquad \ldots \quad (3\text{-}24)$$

or since,

$$a(\omega,u-u_h) = a(\omega-v_h,u-u_h)$$

since:

$$a(u-u_h,v_h) = 0, \text{ for all } v_h \in S_h$$

from (3-24), we have that:

$$a(\omega-v_h,u-u_h) = (g,u-u_h)_{L_2(\Omega)}$$

or,

$$|a(\omega-v_h,u-u_h)| = |(g,u-u_h)_{L_2(\Omega)}| \qquad \ldots \quad (3\text{-}25)$$

By applying the Schwarz inequality, in terms of the energy norm, on the left side of (3-25), we get:

$$|a(\omega-v_h, u-u_h)| \leq (a(\omega-v_h, \omega-v_h))^{\frac{1}{2}} (a(u-u_h, u-u_h))^{\frac{1}{2}} \quad \ldots (3\text{-}26)$$

and suppose that the function v_h is the best approximation to the exact solution ω of the auxiliary differential equation (3-22). We distinguish now between the following two possibilities as far as the order s of the derivatives is concerned, with respect to which an error bound for the difference $u-u_h$ is sought, as well as the other two parameters k and m. Thus:

1. For $s \geq 2m-k$: $\ldots$ (3-27)

Then, for any function $\omega \in H^k(\Omega)$, since:

$$k \geq 2m-s$$

or, by definition:

$$H^k(\Omega) \subset H^{2m-s}(\Omega)$$

it follows that:

$$\omega \in H^{2m-s}(\Omega)$$

Therefore, if v_h denotes the best approximation to the exact solution ω, from the approximation estimate (3-17), we have that:

$$(a(\omega-v_h, \omega-v_h))^{\frac{1}{2}} \leq c_1 h^{(2m-s)-m} |\omega|_{H^{2m-s}(\Omega)}$$

or

$$(a(\omega-v_h, \omega-v_h))^{\frac{1}{2}} \leq c_1 h^{m-s} |\omega|_{H^{2m-s}(\Omega)} \quad \ldots (3\text{-}28)$$

for any function $\omega \in H^{2m-s}(\Omega)$.

On the other hand, from (3-18), we have:

$$(a(u-u_h,u-u_h))^{\frac{1}{2}} \le C_2 h^{k-m} |u|_{H^k(\Omega)} \quad \ldots \quad (3\text{-}29)$$

for any function $u \in H^k(\Omega)$. Then, from the inequality (3-26) combined with (3-28) and (3-29), we get:

$$|a(\omega-u_h,u-u_h)| \le C\, h^{m-s} h^{k-m} |\omega|_{H^{2m-s}(\Omega)} \; |u|_{H^k(\Omega)}$$

or,

$$|a(\omega-u_h,u-u_h)| \le C\, h^{k-s} |\omega|_{H^{2m-s}(\Omega)} \; |u|_{H^k(\Omega)} \quad \ldots \quad (3\text{-}30)$$

Next, since the solution ω of the differential equation (3-22) is always 2m derivatives smoother than the data g; ie,

$$|\omega|_{H^{2m-s}(\Omega)} \le \|\omega\|_{H^{2m-s}(\Omega)} \le K \|g\|_{H^{-s}(\Omega)} \quad \ldots \quad (3\text{-}31)$$

from (3-30) together with (3-25) and 3-31), we have:

$$|(g,u-u_h)_{L_2(\Omega)}| \le C\, h^{k-s} \|g\|_{H^{-s}(\Omega)} \; |u|_{H^k(\Omega)}$$

or

$$\frac{|(g,u-u_h)_{L_2(\Omega)}|}{\|g\|_{H^{-s}(\Omega)}} \le C\, h^{k-s} |u|_{H^k(\Omega)} \quad \ldots \quad (3\text{-}32)$$

or, by taking the maximum in (3-32) over the space $H^s(\Omega)$, we obtain:

$$\|u-u_h\|_{H^s(\Omega)} \equiv \max_{g \in H^s(\Omega)} \frac{|(g,u-u_h)_{L_2(\Omega)}|}{\|g\|_{H^{-s}(\Omega)}} \le C\, h^{k-s} |u|_{H^k(\Omega)}$$

Thus, for the first case where $s \ge 2m-k$, we have:

$$\|u-u_h\|_{H^s(\Omega)} \leq C\, h^{k-s} |u|_{H^k(\Omega)}, \text{ for any function } u \in H^k(\Omega) \quad . \ . \ . \quad (3\text{-}33)$$

2. For $s \leq 2m-k$ . . . (3-34)

Then, again, if the function v_h is the best approximation to ω, for any function $\omega \in H^k(\Omega)$, we have that:

$$(a(\omega-v_h, \omega-v_h))^{\frac{1}{2}} \leq C_1 h^{k-m} |\omega|_{H^k(\Omega)} \quad . \ . \ . \quad (3\text{-}35)$$

and, on the other hand, for any function $u \in H^k(\Omega)$ we have:

$$(a(u-u_h, u-u_h))^{\frac{1}{2}} \leq C_2 h^{k-m} |u|_{H^k(\Omega)} \quad . \ . \ . \quad (3\text{-}36)$$

Thus, from (3-26) combined together with (3-35) and (3-36), we get:

$$|a(\omega-v_h, u-u_h)| \leq C\, h^{2(k-m)} |\omega|_{H^k(\Omega)} |u|_{H^k(\Omega)} \quad . \ . \ . \quad (3\text{-}37)$$

Next, for any function $\omega \in H^k(\Omega)$, since:

$$k \leq 2m-s$$

by definition, we have that:

$$|\omega|_{H^k(\Omega)} \leq \|\omega\|_{H^k(\Omega)} \leq \|\omega\|_{H^{2m-s}(\Omega)}$$

or, by considering the basic result (3-31), that:

$$|\omega|_{H^k(\Omega)} \leq K \|g\|_{H^{-s}(\Omega)} \quad . \ . \ . \quad (3\text{-}38)$$

Thus, from (3-25) combined with (3-37) and (3-38), we get:

$$\left| (g, u-u_h)_{L_2(\Omega)} \right| \leq C\, h^{2(k-m)} \|g\|_{H^{-s}(\Omega)} \left| u \right|_{H^k(\Omega)}$$

or,

$$\frac{\left| (g, u-u_h)_{L_2(\Omega)} \right|}{\|g\|_{H^{-s}(\Omega)}} \leq C\, h^{2(k-m)} \left| u \right|_{H^k(\Omega)} \qquad \ldots \quad (3\text{-}39)$$

and, by taking the maximum in (3-39) over the space $H^s(\Omega)$, we obtain:

$$\|u-u_h\|_{H^s(\Omega)} \equiv \max_{g \in H^s(\Omega)} \frac{\left| (g, u-u_h)_{L_2(\Omega)} \right|}{\|g\|_{H^{-s}(\Omega)}} \leq C\, h^{2(k-m)} \left| u \right|_{H^k(\Omega)}$$

Therefore, for the second case where $s \leq 2m-k$, we have:

$$\|u-u_h\|_{H^s(\Omega)} \leq C\, h^{2(k-m)} \left| u \right|_{H^k(\Omega)} \qquad \ldots \quad (3\text{-}40)$$

for any function $u \in H^k(\Omega)$. finally, by combining the two results (3-33) and (3-40), as far as the *rate of convergence is concerned* we can write that:

$$\|u-u_h\|_{H^s(\Omega)} = O(h^{k-s} + h^{2(k-m)}) \qquad \ldots \quad (3\text{-}41)$$

for the H^s-norm. Nevertheless, it happens that - with the majority of the cases in practice - the first term in (3-41) governs the rate of convergence and this agrees with what naturally is expected from the approximation theory.

3.1.7 The Error in Displacement

For the particular case where $s = 0$, the error in displacement - or the error in the L_2-norm - can be derived by following almost the same steps as those of the above analysis. Therefore, suppose that the auxiliary problem (3-22) is again given and also suppose that the data g now is taken to be equal to the error $u-u_h$. Then, by the vanishing of the first variation:

$$a(\omega,v) = (u-u_h,v)_{L_2(\Omega)}, \text{ for all } v \in V$$

and, by taking $v = u-u_h$:

$$a(\omega,u-u_h) = (u-u_h,u-u_h)_{L_2(\Omega)} = \|u-u_h\|^2_{L_2(\Omega)} \quad \ldots (3-42)$$

Then, since always:

$$a(\omega-v_h,u-u_h) = a(\omega,u-u_h)$$

from (3-42), we have:

$$|a(\omega-v_h,u-u_h)| = \|u-u_h\|^2_{L_2(\Omega)} \quad \ldots (3-43)$$

By applying the Schwarz inequality to the left side of (3-43), we get:

$$|a(\omega-v_h,u-u_h)| \leq (a(\omega-v_h,\omega-v_h))^{\frac{1}{2}} (a(u-u_h,u-u_h))^{\frac{1}{2}} \quad \ldots (3-44)$$

Thus, for any function $u \in H^k(\Omega)$ and for the usual case in practice where $k > 2m$ since, by definition:

$$H^k(\Omega) \subset H^{2m}(\Omega)$$

it follows that:

$$u \in H^{2m}(\Omega)$$

and, if v_h is again considered to be the best approximation to the exact solution ω of the auxiliary problem, we have:

$$(a(\omega-v_h,\omega-v_h))^{\frac{1}{2}} \leq C_1 h^{2m-m} |\omega|_{H^{2m}(\Omega)}$$

or

$$(a(\omega-v_h,\omega-v_h))^{\frac{1}{2}} \leq C_1 h^m |\omega|_{H^{2m}(\Omega)} \quad \ldots (3\text{-}45)$$

On the other hand:

$$(a(u-u_h,u-u_h))^{\frac{1}{2}} \leq C_2 h^{k-m} |u|_{H^k(\Omega)}, \text{ for any } u \in H^k(\Omega) \quad \ldots (3\text{-}46)$$

Therefore, from (3-43) combined together with (3-44), (3-45) and (3-46), we obtain:

$$\|u-u_h\|^2_{L_2(\Omega)} \leq C\, h^k |\omega|_{H^{2m}(\Omega)} \, |u|_{H^k(\Omega)} \quad \ldots (3\text{-}47)$$

and since:

$$|\omega|_{H^{2m}(\Omega)} \leq \|\omega\|_{H^{2m}(\Omega)} \leq K \|g\|_{L_2(\Omega)} = K \|u-u_h\|_{L_2(\Omega)}$$

from (3-47), we get:

$$\|u-u_h\|^2_{L_2(\Omega)} \leq C\, h^k \|u-u_h\|_{L_2(\Omega)} \, |u|_{H^k(\Omega)}$$

or

$$\|u-u_h\|_{L_2(\Omega)} \leq C\, h^k\, |u|_{H^k(\Omega)} \qquad \ldots \quad (3\text{-}48)$$

for any function $u \in H^k(\Omega)$, which gives the correct order of convergence in displacement.

3.1.8 The Phenomenon of Superconvergence

Let us consider now the case where $s < 0$ and, in particular, the case where $s = -1$. Then, by definition of the dual norm:

$$\|u-u_h\|_{H^{-1}(\Omega)} = \max_{v \in H^1(\Omega)} \frac{\left|\int_\Omega (u-u_h)\, v\, dx\right|}{\|v\|_{H^1(\Omega)}}$$

from which, by taking the function $v(x) = 1$, we get:

$$\|u-u_h\|_{H^{-1}(\Omega)} \geq \frac{\left|\int_\Omega (u-u_h)\, dx\right|}{(\mathrm{vol}\Omega)^{\frac{1}{2}}} \qquad \ldots \quad (3\text{-}49)$$

or, the inequality (3-49), can be written as:

$$\left|\int_\Omega (u-u_h)\, dx\right| \leq C\, \|u-u_h\|_{H^{-1}(\Omega)}$$

for some numerical constant C, and by combining this result with the error estimate (3-33) for the usual case in practice where $k > 2m$, ie, with the error estimate:

$$\|u-u_h\|_{H^{-1}(\Omega)} \leq C\, h^{k+1} |u|_{H^k(\Omega)}$$

we obtain:

$$\left|\int_{\Omega}(u-u_h)\,dx\right| \leq C\,h^{k+1}\,|u|_{H^k(\Omega)} \qquad \ldots (3\text{-}50)$$

which expresses the average error over the entire domain Ω. If we compare now the two estimates (3-48) and (3-50), ie, the error in *displacement at a point* and *the average* error over the entire domain Ω, we see that the second is smaller than the first and, therefore, it is expected that somewhere within the domain the error alternates rapidly in sign. Hence, the problem of finding the so-called *special points* within each element where such changes do occur is indeed of great practical importance and near these points the approximate solution u_h will naturally be of exceptional accuracy. This phenomenon of super convergence at the special points has been analysed by Dupont and Douglas (see eg. [11]) and the following estimate can easily be established.

$$|u(x_o)-u_h(x_o)| \leq \left(a(u-u_h,u-u_h)\right)^{\frac{1}{2}}\left(a(G_o-v_h,G_o-v_h)\right)^{\frac{1}{2}} \qquad \ldots (3\text{-}51)$$

for any function v_h, where x_o denotes one of these *special points* and G_o denotes the fundamental solution corresponding to the particular point x_o, ie, G_o satisfies the following equation:

$$a(G_o,v) = (f_o,v)_{L_2(\Omega)}, \text{ for all } v \in V$$

where:

$$f_o = \delta(x-x_o)$$

corresponds to a point load.

Indeed, since:

$$\alpha(G_o, v) = (f_o, v)_{L_2(\Omega)} = (\delta(x-x_o), v)_{L_2(\Omega)} =$$

$$= \int_\Omega \delta(x-x_o)\, v(x)\, dx = v(x_o), \text{ for any } v \in V$$

and a similar result holds for the approximating function u_h, we have that:

$$u(x_o) - u_h(x_o) = \alpha(G_o, u) - \alpha(G_o, u_h) = \alpha(G_o, u-u_h) =$$

$$= \alpha(G_o - v_h, u-u_h)$$

since,

$$\alpha(u-u_h, v_h) = 0, \text{ for all } v_h \in S_h \subset V.$$

Thus,

$$u(x_o) - u_h(x_o) = \alpha(G_o - v_h, u-u_h)$$

or

$$|u(x_o) - u_h(x_o)| = |\alpha(G_o - v_h, u-u_h)| \qquad \ldots (3\text{-}52)$$

By applying the Schwarz inequality on the right side of (3-52), we get:

$$|u(x_o) - u_h(x_o)| \leq (\alpha(G_o - v_h, G_o - v_h))^{\frac{1}{2}} (\alpha(u-u_h, u-u_h))^{\frac{1}{2}}$$

and since the term $(\alpha(G_o - v_h, G_o - v_h))^{\frac{1}{2}}$ will certainly add some finite power of the parameter h to the h^{k-m} obtained by the other term, the convergence of the method at the point x_o will indeed be exceptional.

3.2 ERROR ANALYSIS FOR NON-CONFORMING FINITE ELEMENTS

We extend the above convergence analysis of the Finite Element Method in order to include an extremely important

case arising from the following situation: The finite elements are non-conforming - it was stated at the very beginning and consequently assumed that the variational problem requires the trial functions to have a certain smoothness in order that they *conform* to the theory - and the space S_h does not now constitute *a subspace of* the energy space $V \subset H^m(\Omega)$. When such a situation arises, the application of a particular test is all that is needed in order to determine whether or not convergence may take place. However, since the inclusion $S_h \subset V$ does not hold in this case, the rules of the Ritz procedure are violated and since:

$$a(u_h, v_h) = \infty, \text{ for } v_h \in S_h$$

the minimization process of the quadratic functional (1-18) can no longer proceed. To compensate, the technique consists of computing the energies within each element e_i separately, by ignoring the interelement discontinuities of the trial functions and, then, to add together the results. In so doing, however, we have replaced the functional (1-18) by the following functional:

$$F^*(v_h) = a^*(v_h, v_h) - 2(f, v_h)^*_{L_2(\Omega)} \qquad \ldots (3\text{-}53)$$

where,

$$a^*(u_h, v_h) = \sum_{e_i \in T_h} [a(u_h, v_h)]_{e_i}; \quad (f, v_h)^*_{L_2(\Omega)} = \sum_{e_i \in T_h} (f, v_h)_{L_2(e_i)}$$

and the difference between (1-18) and (3-53) is that

$F(v_h) = \infty$ for non-conforming elements. In this case, the *convergence* of the method is no longer an automatic consequence of the approximation properties of the subspace S_h and it becomes an exception rather than the rule. However, since non-conforming elements have been successfully used to date - most commonly by the engineers who believe that the conformity requirement leads to complicated finite elements - there exist a test, which was first devised by Irons and is known as *the patch test*, which when it is passed convergence does occur. The patch test says (see G. Strang and G. Fix [24] p.174): Suppose that within a patch of elements the exact solution u is equal to a polynomial $P_m(x)$ of degree $m \leq k-1$. Then, since the constant strain condition is still satisfied by non-conforming elements, this polynomial is present at the subspace S_h and, therefore, the finite element approximation u_h coincides identically with the polynomial $P_m(x)$. After that, the test consists of checking out whether the non-conforming finite element approximation u_h^*, which minimizes the new functional (3-53) over S_h, is still identical to the polynomial $P_m(x)$ in spite of shifting from the functional (1-18) to the new functional (3-53). This is the test devised by Irons. G. Strang in [23], however, has shown that if, for any polynomial $P_m(x)$ of degree $m \leq k-1$ and each non-conforming basis $\phi_j(x)$, the following equation holds:

$$a^*(P_m(x), \phi_j(x)) = a(P_m(x), \phi_j(x)) \quad . \ . \ . \ (3\text{-}54)$$

then the patch test is passed and vice versa. Thus, what

we actually do in practice is to prove whether or not equality in (3-54) holds and, correspondingly, we determine whether or not convergence does occur.

By making the assumption hereafter that the patch test is passed, we are again faced with the problem of giving an estimate for the error between the exact solution u *and its non-conforming finite element approximation* u_h^*. Indeed, if we make the additional assumption that the function f is sufficiently smooth, by the vanishing of the first variation, we have:

$$a^*(u_h^*, v_h) = (f, v_h)^*_{L_2(\Omega)}, \text{ for all } v_h \in S_h \qquad \ldots (3\text{-}55)$$

Then, since:

$$a(f, v_h)^*_{L_2(\Omega)} = (f, v_h)_{L_2(\Omega)}, \quad v_h \in S_h \qquad \ldots (3\text{-}56)$$

and:

$$a(u, v_h) = (f, v_h)_{L_2(\Omega)}, \quad v_h \in S_h \qquad \ldots (3\text{-}57)$$

we can, as a first step towards our final goal, easily prove that:

$$a^*(u-u_h^*, v_h) = a^*(u, v_h) - a(u, v_h) \qquad \ldots (3\text{-}58)$$

Indeed, by using the equations (3-55), (3-56) and (3-57), we have:

$$a^*(u-u_h^*, v_h) = a^*(u_h, v_h) - a^*(u_h^*, v_h) = a^*(u, v_h) - (f, v_h)^*_{L_2(\Omega)} =$$

$$= a^*(u, v_h) - (f, v_h)_{L_2(\Omega)} = a^*(u, v_h) - a(u, v_h).$$

Suppose now that the function $\omega \in S_h$ denotes the orthogonal projection of the exact solution u onto the subspace S_h, and consider the following triangle inequality:

$$|\alpha^*(u-u_h^*, u-u_h^*)| \leq |\alpha^*(u-\omega, u-\omega)| + |\alpha^*(\omega-u_h^*, \omega-u_h^*)| \quad \ldots (3\text{-}59)$$

The first part of this inequality is exactly an expression of the error $u-u_h^*$ in terms of the energy norm and, therefore, we shall make use of this inequality in order to estimate an upper bound for the error by giving special attention to its second part. Then, by keeping in mind that the function ω is the closest element in S_h to u, the term:

$$|\alpha^*(u-\omega, u-\omega)|$$

precisely reflects the error in approximation and, thus:

$$|\alpha^*(u-\omega, u-\omega)| = \min_{\omega_h \in S_h} |\alpha^*(u-\omega_h, u-\omega_h)|$$

For the other term; viz:

$$|\alpha^*(\omega-u_h^*, \omega-u_h^*)|$$

we note that:

$$\max_{v_h \in S_h} \frac{|\alpha^*(\omega-u_h^*, v_h)|^2}{|\alpha^*(v_h, v_h)|} = |\alpha^*(\omega-u_h^*, \omega-u_h^*)| \quad \ldots (3\text{-}60)$$

Then, since:

$$\alpha^*(u-\omega, v_h) = 0, \text{ for all } v_h \cdot S_h$$

or

$$a^*(u-v_h) = a^*(\omega,v_h) \text{ for all } v_h \in S_h \qquad \ldots (3\text{-}61)$$

from (3-60) combined together with (3-61) and (3-58), we have:

$$|a^*(\omega-u_h^*,\omega-u_h^*)| = \max_{v_h\in S_h} \frac{|a^*(\omega-u_h^*,v_h)|^2}{|a^*(v_h,v_h)|} = \max_{v_h\in S_h} \frac{|a^*(\omega,v_h)-a^*(u_h^*,v_h)|^2}{|a^*(v_h,v_h)|} =$$

$$= \max_{v_h\in S_h} \frac{|a^*(u,v_h)-a^*(u_h^*,v_h)|^2}{|a^*(v_h,v_h)|} = \max_{v_h\in S_h} \frac{|a^*(u,v_h)-a(u,v_h)|^2}{|a^*(v_h,v_h)|}.$$

Therefore,

$$|a^*(\omega-u_h^*,\omega-u_h^*)| = \max_{v_h\in S_h} \frac{|a^*(u,v_h)-a(u,v_h)|^2}{|a^*(v_h,v_h)|} \qquad \ldots (3\text{-}62)$$

and the inequality (3-59) can be written as:

$$|a^*(u-u_h^*,u-u_h^*)| \le \min_{w_h\in S_h} |a^*(u-w_h,u-w_h)| + \max_{v_h\in S_h} \frac{|a^*(u,v_h)-a(u,v_h)|}{|a^*(v_h,v_h)|} \qquad \ldots (3\text{-}63)$$

Then, since from the approximation-theory results, we always have that:

$$\min_{\omega_h\in S_h} |a^*(u-\omega_h,u-\omega_h)| \le C_1 h^{2(k-m)} |u|^2_{H^k(\Omega)} \qquad \ldots (3\text{-}64)$$

for any function $u \in H^k(\Omega)$ and some numerical constant C_1, the only problem that still remains to be solved is to

find an upper bound for the other term in (3-63), ie, for the term:

$$\max_{v_h \in S_h} \frac{|a^*(u,v_h) - a(u,v_h)|^2}{|a^*(v_h,v_h)|}$$

for this, and in order to generalize the result given by G. Strang and G. Fix in [24] p.180, we assume the following important convention: The new trial space of non-conforming elements is regarded as a conforming one to which a number of non-conforming trial functions have been added. Thus, any function $v_h \in S_h$ can be written as:

$$v_h(x) = \sum_{e_i \in \Gamma_h} \left(\sum_{e_i} a_i \phi_i(x) + \sum_{e_i} b_i \psi_i(x) \right) \quad \ldots \quad (3\text{-}65)$$

where,

$$\sum_{e_i} a_i \phi_i(x) + \sum_{e_i} b_i \psi_i(x)$$

expresses the contribution within each element e_i of the given partition Γ_h of the domain Ω and the $\psi_i(x)$'s denote the non-conforming basis functions which have been added to the set of the conforming basis functions. Note, however, that since the basis functions $\phi_i(x)$ in (3-65) are conforming, their contribution to the present kind of error analysis will be equal to zero. Then, within each element $e_i \in \Gamma_h$, we have:

$$|a^*(u,v_h) - a(u,v_h)|_{e_i} = |a^*(u, \sum_{e_i} b_i\psi_i) - a(u, \sum_{e_i} b_i\psi_i)|_{e_i} =$$

$$= |a^*(u-P_m, \sum_{e_i} b_i\psi_i) - a(u-P_m, \sum_{e_i} b_i\psi_i)|_{e_i}$$

for some polynomial $P_m(x)$ of degree $m \le k-1$. It is obvious from equation (3-54) that the inclusion of the polynomial into the above expression has not any effect whatsoever. Then:

$$|a^*(u,v_h)-a(u,v_h)|_{e_i} = |a^*(u-P_m, \Sigma b_i\psi_i) - a(u-P_m, \Sigma b_i\psi_i)|_{e_i} \le$$

$$\le |a^*(u-P_m, \Sigma b_i\psi_i)|_{e_i} + |a(u-P_m, \Sigma b_i\psi_i)|_{e_i} \le$$

$$\le C'\|u-P_m\|_{H^m(e_i)} \|\Sigma b_i\psi_i\|_{H^m(e_i)} + C''\|u-P_m\|_{H^m(e_i)} \|\Sigma b_i\psi_i\|_{H^m(e_i)} \le$$

$$\le C'''\|u-P_m\|_{H^m(e_i)} \|\Sigma b_i\psi_i\|_{H^m(e_i)} \le C\, h_{e_i}^{k-m}|u|_{H^k(e_i)} \|\Sigma b_i\psi_i\|_{H^m(e_i)}$$

Thus, over each element $e_i \in T_h$, *we have:*

$$|a^*(u,v_h)-a(u,v_h)|_{e_i} \le C\, h_{e_i}^{k-m}|u|_{H^k(e_i)} \|\Sigma b_i\psi_i\|_{H^m(e_i)} \quad \ldots \quad (3\text{-}66)$$

By adding the inequalities (3-66) for all the elements e_i of the partition T_h, we get:

$$\sum_{e_i\in T_h} |a^*(u,v_h)-a(u,v_h)|_{e_i} = |a^*(u,v_h)-a(u,v_h)| \le$$

$$\le C\, h^{k-m} \sum_{e_i\in T_h} |u|_{H^k(e_i)} \|\Sigma b_i\psi_i\|_{H^m(e_i)} \quad \ldots \quad (3\text{-}67)$$

Next, from (3-67) and by using the Schwarz inequality, we obtain:

$$|a^*(u,v_h)-a(u,v_h)| \le C\, h^{k-m}\Big(\sum_{e_i\in T_h}|u|^2_{H^k(e_i)}\Big)^{\frac{1}{2}}\cdot\Big(\sum_{e_i\in T_h}\|\sum_i b_i\varphi_i\|^2_{H^m(e_i)}\Big)^{\frac{1}{2}} \le$$

$$\le C\, h^{k-m}|u|_{H^k(\Omega)}\|v_h\|_{H^m(\Omega)} \le C\, h^{k-m}|u|_{H^k(\Omega)}[a^*(v_h,v_h)]^{\frac{1}{2}}$$

or, by dividing by $[a^*(v_h,v_h)]^{\frac{1}{2}}$ we get:

$$\frac{|a^*(u,v_h)-a(u,v_h)|}{[a^*(v_h,v_h)]^{\frac{1}{2}}} \le C\, h^{k-m}|u|_{H^k(\Omega)},$$

or,

$$\max_{v_h\in S_h}\frac{|a^*(u,v_h)-a(u,v_h)|^2}{|a^*(v_h,v_h)|} \le C_2\, h^{2(k-m)}|u|^2_{H^k(\Omega)} \quad \ldots \quad (3\text{-}68)$$

Finally, from (3-63) together with (3-64) and (3-68), we obtain the following estimate for the error in $u-u_h^*$ in terms of the energy norm:

$$|a^*(u-u_h^*,u-u_h^*)| \le C\, h^{2(k-m)}|u|^2_{H^k(\Omega)} \quad \ldots \quad (3\text{-}69)$$

for any function $u\in H^k(\Omega)$ and some numerical constant C which does not depend on the function u and the geometrical parameter h. From (3-69) we can see that the error of convergence of non-conforming elements which hav passed the test is the same as for the conforming ones considered in the previous section of this error analysis.

3.3 TWO ILLUSTRATIVE EXAMPLES

To conclude, let us consider the following two simple examples in one and two variables, respectively, which

illustrate to some extent the process which is usually followed in practice in order to compute the approximate solution $u_h(x)$, as well as to estimate an upper bound for the error in $u(x) - u_h(x)$, where $u(x)$ denotes the exact solution of the differential equation under consideration:

1. Example 1

Consider the following ordinary differential equation of second order:

$$-\frac{d^2u}{dx^2} + u(x) = x, \quad 0 < x < 1 \qquad \ldots (3\text{-}70)$$

subject to the boundary conditions:

$$u(0) = u(1) = 0. \qquad \ldots (3\text{-}71)$$

Its exact solution is:

$$u(x) = (e^{-1}-e)^{-1}e^{x} + (e-e^{-1})^{-1}e^{-x} + x \qquad \ldots (3\text{-}72)$$

and the associated variational problem is to minimize the following integral:

$$F(u) = \int_0^1 [(\frac{du}{dx})^2 + (u(x))^2 - 2u(x).x]dx \qquad \ldots (3\text{-}73)$$

over the infinite-dimensional space $V \subset H^1[0,1]$ defined by:

$$V = \{u(x) \in H^1[0,1] : u(0) = u(1) = 0\}$$

where:

$$H^1[0,1] = \{u(x) : u(x) \in L_2[0,1];\ \frac{du}{dx} \in L_2[0,1]\}$$

However, since it can easily be seen that the differential operator defined by:

$$A.u(x) = -\frac{d^2u}{dx^2} + u(x)$$

is a linear positive definite and symmetric operator, the Rayleigh-Ritz version of the Finite Element Method can be employed in computing an approximate solution of the problem (3-70) - (3-71). The infinite-dimensional space V is then replaced by a finite-dimensional subspace S_h of V and for such a subspace in this example we consider the space of all the piecewise quadratic functions which are continuous at the nodes belonging to the following uniform partition:

$$\Delta: 0 < h < 2h < 3h < 4h = 1 \qquad \ldots (3\text{-}74)$$

of [0,1] and satisfy the boundary conditions (3-71). A basis for the subspace S_h can be constructed by introducing the following mid-points:

$$x = (2j-1)h/2,\ 1 \leq j \leq 4 \qquad \ldots (3\text{-}75)$$

of the above partition Δ, in addition to the end-points:

$$x = jh,\ 1 \leq j \leq 4 \qquad \ldots (3\text{-}76)$$

Then, over the subspace S_h, the quadratic functional (3-73) is replaced by the following functional:

$$F(u_h) = \int_0^1 [(\frac{du_h}{dx})^2 + (u_h(x))^2 - 2u_h(x).x]dx$$

or, subject to the partition (3-74) together with (3-75) and (3-76), by:

$$F(v_h) = \sum_{j=1}^{4} \int_{(j-1)h}^{jh} [(\frac{dv_h}{dx})^2 + (v_h(x))^2 - 2v_h(x).x]dx \quad \ldots (3\text{-}77)$$

Over each sub-interval:

$$(j-1)h \le x \le jh, \quad 1 \le j \le 4$$

the quadratic function $v_h(x) \in S_h$ which equals q_{j-1} at the node $x = (j-1)h$, $q_{(2j-1)/2}$ at $x = (2j-1)h/2$ and q_j at the node $x = jh$ is given by the following general formula:

$$v_h(x) = \frac{1}{\rho}[-\frac{1}{2}\frac{x^2}{h^2} + (j^2 - \frac{(2j-1)^2}{4})\frac{x}{h} - \frac{j(2j-1)}{4}]q_{j-1} +$$

$$+ \frac{1}{\rho}[\frac{x^2}{h^2} + ((j-1)^2 - j^2)\frac{x}{h} + j(j-1)]q_{(2j-1)/2} +$$

$$+ \frac{1}{\rho}[-\frac{1}{2}\frac{x^2}{h^2} + (\frac{(2j-1)^2}{4} - (j-1)^2)\frac{x}{h} - \frac{1}{4}(j-1)(2j-1)]q_j, \quad 1 \le j \le 4$$

$$\ldots (3\text{-}78)$$

where,

$$\rho = -\frac{1}{4}j(2j-1) + j(j-1) - \frac{1}{4}(j-1)(2j-1), \quad 1 \le j \le 4 \quad \ldots (3\text{-}79)$$

By performing a rigorous but simple computation on (3-78) and (3-79), we can get the following results:

$$\int_{(j-1)h}^{jh} (\frac{dv_h}{dx})^2 dx = \frac{1}{3h}[7q_{j-1}^2 + 16q_{(2j-1)/2}^2 + 7q_j^2 - 16q_{j-1}q_{(2j-1)/2} +$$

$$+ 2q_{j-1}q_j - 16q_{(2j-1)/2}q_j] =$$

$$= (q_{j-1}, q_{(2j-1)/2}, q_j)\frac{1}{3h}\begin{bmatrix} 7 & -8 & 1 \\ -8 & 16 & -8 \\ 1 & -8 & 7 \end{bmatrix} \cdot \begin{bmatrix} q_{j-1} \\ q_{(2j-1)/2} \\ q_j \end{bmatrix}, \quad 1 \le j \le 4$$

$$\int_{(j-1)h}^{jh} (v_h(x))^2 dx = \frac{h}{15}[2q_{j-1}^2 + 8q_{(2j-1)/2}^2 + 2q_j^2 + 2q_{j-1}q_{(2j-1)/2} -$$

$$- q_{j-1}q_j + 2q_{(2j-1)/2}q_j] =$$

$$= (q_{j-1}, q_{(2j-1)/2}, q_j)\frac{h}{15}\begin{bmatrix} 2 & 1 & -1/2 \\ 1 & 8 & 1 \\ -1/2 & 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} q_{j-1} \\ q_{(2j-1)/2} \\ q_j \end{bmatrix}, \quad 1 \le j \le 4$$

and,

$$\int_{(j-1)h}^{jh} v_h(x) \cdot x \, dx = (q_{j-1}, q_{(2j-1)/2}, q_j)h^2 \begin{bmatrix} 0 \\ 1/3 \\ 1/6 \end{bmatrix}, \quad 1 \le j \le 4 \quad \ldots \quad (3\text{-}80)$$

The matrix:

$$k_1 = \frac{1}{3h}\begin{bmatrix} 7 & -8 & 1 \\ -8 & 16 & -8 \\ 1 & -8 & 7 \end{bmatrix} \quad \ldots \quad (3\text{-}81)$$

known as the element stiffness matrix is replaced, over the entire domain [0,1], by a global stiffness matrix K_1 by adding together the resulting element matrices (3-81). Then, since the two extreme points q_0 and q_4 have to be discarded because of the boundary conditions (3-71), we get the following matrix:

$$K_1 = \frac{1}{3h} \cdot \begin{bmatrix} 16 & -8 & 0 & 0 & 0 & 0 & 0 \\ -8 & 14 & -8 & 1 & 0 & 0 & 0 \\ 0 & -8 & 16 & -8 & 0 & 0 & 0 \\ 0 & 1 & -8 & 14 & -8 & 1 & 0 \\ 0 & 0 & 0 & -8 & 16 & -8 & 0 \\ 0 & 0 & 0 & 1 & -8 & 14 & -8 \\ 0 & 0 & 0 & 0 & 0 & -8 & 16 \end{bmatrix}$$

Therefore,

$$\int_0^1 \left(\frac{du_h}{dx}\right)^2 dx = (q_{1/2}, q_1, q_{3/2}, \ldots, q_{7/2}) K_1 \cdot \begin{bmatrix} q_{1/2} \\ q_1 \\ q_{3/2} \\ \vdots \\ q_{7/2} \end{bmatrix} \quad \ldots \quad (3\text{-}82)$$

In exactly the same way, the matrix:

$$k_0 = \frac{h}{15} \cdot \begin{bmatrix} 2 & 1 & -1/2 \\ 1 & 8 & 1 \\ -1/2 & 1 & 2 \end{bmatrix}$$

known as the element mass matrix is replaced, over the entire domain [0,1], by a global mass matrix:

$$K_0 = \frac{h}{15} \cdot \begin{bmatrix} 8 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 4 & 1 & -1/2 & 0 & 0 & 0 \\ 0 & 1 & 8 & 1 & 0 & 0 & 0 \\ 0 & -1/2 & 1 & 4 & 1 & -1/2 & 0 \\ 0 & 0 & 0 & 1 & 8 & 1 & 0 \\ 0 & 0 & 0 & -1/2 & 1 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 8 \end{bmatrix}$$

and,

$$\int_0^1 (u_h(x))^2 dx = (q_{1/2}, q_1, q_{3/2}, \ldots, q_{7/2}) K_0 \cdot \begin{bmatrix} q_{1/2} \\ q_1 \\ q_{3/2} \\ \vdots \\ q_{7/2} \end{bmatrix} \quad \ldots \quad (3\text{-}83)$$

Finally, from (3-80), we get:

$$\int_0^1 v_h(x).x\,dx = (q_{1/2}, q_1, q_{3/2}, \ldots, q_{7/2})F$$

where the vector F is given by:

$$F = h^2(1/3, 1/6, 1/3, 1/6, 1/3, 1/6, 1/3)^T \qquad \ldots (3\text{-}84)$$

Thus, from (3-77) combined together with (3-82), (3-83) and (3-84), we obtain:

$$F(v_h) = (q_{1/2}, q_1, q_{3/2}, \ldots, q_{7/2})(K_1+K_0)\begin{bmatrix} q_{1/2} \\ q_1 \\ q_{3/2} \\ \cdot \\ \cdot \\ q_{7/2} \end{bmatrix} - 2(q_{1/2}, q_1, q_{3/2}, \ldots, q_{7/2})F \qquad \ldots (3\text{-}85)$$

The optimal vector:

$$Q = (Q_{1/2}, Q_1, Q_{3/2}, \ldots, Q_{7/2})$$

for which the above expression (3-85) attains its minimum value can be computed if we differentiate that expression with respect to the parameters $q_{1/2}, q_1, q_{3/2}, \ldots q_{7/2}$ and set the result equal to zero. This gives rise to the following linear system:

$$KQ = F \qquad \ldots (3\text{-}86)$$

where the stiffness matrix:

$$K = K_0 + K_1$$

and the vector F is defined as in (3-84). By employing

now the Gauss elimination procedure on the linear system (3-86), the optimal vector Q can easily be determined and, subsequently, the finite element approximation $u_h(x)$ to the exact solution (3-72).

Nevertheless, in order to give an upper bound for the error between the exact solution u and its approximation u_h, it is not necessary to compute the approximate solution u_h explicity. Indeed, by recalling the minimum principle:

$$a(u-u_h, u-u_h) \leq a(u-v_h, u-v_h) \quad \text{for all } v_h \in S_h \qquad \ldots \quad (3\text{-}87)$$

and by taking the function $v_h \in S_h$, in the second part of (3-87), to be the unique piecewise quadratic Lagrange interpolating polynomial denoted by $u_1^{(2)}(x)$ - in accordance with its general definition given in (2-7) - which interpolates the values of the function (3-72) at the knots of the partition (3-74), we have that:

$$a(u-u_h, u-u_h) \leq a(u-u_1^{(2)}, u-u_1^{(2)}) \qquad \ldots \quad (3\text{-}88)$$

Then, since:

$$a(u-u_1^{(2)}, u-u_1^{(2)}) = \int_0^1 [(Du-Du_1^{(2)})^2 + (u-u_1^{(2)})^2]\,dx =$$

$$= \int_0^1 (Du-Du_1^{(2)})^2\,dx + \int_0^1 (u-u_1^{(2)})^2\,dx = \|Du-Du_1^{(2)}\|^2_{L_2[0,1]} + \|u-u_1^{(2)}\|^2_{L_2[0,1]} =$$

$$= \|u-u_1^{(2)}\|^2_{H^1[0,1]}$$

from (3-88), we get the result:

$$a(u-u_h, u-u_h) \leq \| u-u_1^{(2)} \|^2_{H^1[0,1]} \qquad \ldots \quad (3\text{-}89)$$

and the accuracy of the approximation $u_h(x)$ depends on the smoothness of the function $u(x)$. However, since the data $f(x) = x$ is a continuous function and the solution $u(x)$ is always 2m derivatives smoother than the data $f(x)$ - where 2m is the order of the differential equation - from the approximation theory, we have that:

$$\| u-u_1^{(2)} \|_{H^1[0,1]} \leq C\, h^2 |u|_{H^3[0,1]}$$

for some numerical constant C, or, since by definition:

$$|u|_{H^3[0,1]} = \| D^3 u \|_{L_2[0,1]}$$

that:

$$\| u-u_1^{(2)} \|_{H^1[0,1]} \leq C\, h^2 \| D^3 u \|_{L_2[0,1]} \qquad \ldots \quad (3\text{-}90)$$

and the norm $\| D^3 u \|_{L_2[0,1]}$ can be computed exactly since the exact solution $u(x)$ is known. Therefore, from (3-89) combined with (3-90), we get the following result:

$$a(u-u_h, u-u_h) \leq C\, h^4 \| D^3 u \|^2_{L_2[0,1]} \qquad \ldots \quad (3\text{-}91)$$

which expresses a first upper bound between the functions u and u_h in terms of the energy norm. Then, since:

$$a(u-u_h, u-u_h) = \int_0^1 [(Du-Du_h)^2 + (u-u_h)^2]\,dx =$$

$$= \| Du-Du_h \|^2_{L_2[0,1]} + \| u-u_h \|^2_{L_2[0,1]} = \| u-u_h \|^2_{H^1[0,1]}, \quad \text{ie,}$$

$$a(u-u_h, u-u_h) = \| u-u_h \|^2_{H^1[0,1]} \qquad \ldots \quad (3\text{-}92)$$

from (3-92) and (3-91) we get the following result in terms, this time, of the H^1-norm:

$$\|u-u_h\|_{H^1[0,1]} \leq C\, h^2 \|D^3 u\|_{L_2[0,1]} \qquad \ldots \quad (3\text{-}93)$$

which is an equivalent result to those given by (3-91). This, however, is quite naturally expected since convergence in the energy norm essentially means convergence of the first derivatives of $u(x)$ to those of $u_h(x)$.

2. Example 2

As a second example, consider the following partial differential equation of second order:

$$-\Delta u = f, \quad (x,y) \in \Omega = [0,1] \times [0,1] \qquad \ldots \quad (3\text{-}94)$$

subject to the boundary conditions:

$$u(x,y) = 0, \quad (x,y) \in \partial\Omega \qquad \ldots \quad (3\text{-}95)$$

Suppose that the function f is a continuous function and there exists a unique solution $u(x,y)$ to the problem (3-94) - (3-95). One of the most frequent applications of this problem found in practice is the torsion of prismatic bars (see eg [13] p.138) where the domain Ω represents the normal cross-section of the bar and $\partial\Omega$ the boundary of the section.

Then, the problem of computing the exact solution $u(x,y)$

of the differential equation (3-94) subject to the boundary conditions (3-95) is again equivalent to that of finding that function u(x,y) which minimizes the following functional:

$$F(u) = \iint_{\Omega} [(\frac{\partial u}{\partial x})^2 + (\frac{\partial u}{\partial y})^2 - 2fu]dxdy \quad \ldots (3\text{-}96)$$

over the infinite-dimensional space $V \subset H^1(\Omega)$ defined by:

$$V = \{u(x,y) \in H^1(\Omega) : u(x,y) = 0,\ x,y \in \partial\Omega\}$$

where,

$$H^1(\Omega) = \{u(x,y) : u(x,y) \in L_2(\Omega) :\ Du(x,y) \in L_2(\Omega)\}.$$

However, since the differential operator defined by:

$$-\Delta.u(x,y) = -\frac{\partial^2 u(x,y)}{\partial x^2} - \frac{\partial^2 u(x,y)}{\partial y^2} \quad \ldots (3\text{-}97)$$

is a linear positive definite and symmetric operator, the Rayleigh-Ritz version of the Finite Element Method can be employed and an approximate solution $u_h(x,y)$ to the exact solution u(x,y) can be computed in a process quite analogous to that outlined in the first example. For a finite-dimensional subspace $S_h \subset V$ we consider the space of all the piecewise linear functions which are piecewise continuous at the nodes belonging to the triangulation of the domain $\Omega = [0,1] \times [0,1]$ which arises when it is divided into, say, 16 squares of side 0.25 and every such square in two triangles by the diagonal parallel to the axe of symmetry of the second quadrant.

Moreover, functions of the subspace S_h have also to satisfy the boundary conditions (3-95).

Thus, if by $u_2^{(1)}(x,y)$ we denote the unique piecewise linear Lagrange interpolating polynomial which interpolates the values of the function $u(x,y)$ at the nodes of the described partition, we have:

$$a(u-u_h,u-u_h) \leq a(u-u_2^{(1)},u-u_2^{(1)}) \qquad \ldots (3\text{-}98)$$

or, since:

$$a(u-u_2^{(1)},u-u_2^{(1)}) = \int_0^1\int_0^1 \left[\left(\frac{\partial(u-u_2^{(1)})}{\partial x}\right)^2 + \left(\frac{\partial(u-u_2^{(1)})}{\partial y}\right)^2\right]dxdy =$$

$$= \int_0^1\int_0^1 (D(u-u_2^{(1)}))^2 dxdy = \|D(u-u_2^{(1)})\|^2_{L_2(\Omega)} \leq \|u-u_2^{(1)}\|^2_{H^1(\Omega)}$$

from (3-98) we get the following result:

$$a(u-u_h,u-u_h) \leq \|u-u_2^{(1)}\|^2_{H^1(\Omega)} \qquad \ldots (3\text{-}99)$$

Then, since:

$$\|u-u_2^{(1)}\|_{H^1(\Omega)} \leq C\,h\,|u|_{H^2(\Omega)} \qquad \ldots (3\text{-}100)$$

for some numerical constant C, which is given through the general formula (2-88), from (3-99) combined with (3-100), we obtain:

$$a(u-u_h,u-u_h) \leq C\,h^2\,|u|^2_{H^2(\Omega)} \qquad \ldots (3\text{-}101)$$

where the geometrical parameter h denotes the greatest side of the triangles in the described triangulation.

Moreover, since by definition:

$$|u|^2_{H^2(\Omega)} = \sum_{|\alpha|=2} \|D^\alpha u\|^2_{L_2(\Omega)}$$

where $\alpha = (\alpha_1, \alpha_2)$, $|\alpha| = \alpha_1 + \alpha_2$ and $D^\alpha u = \frac{\partial^\alpha u}{\partial x^{\alpha_1} \partial y^{\alpha_2}}$, the estimate (3-101) can be written as:

$$a(u-u_h, u-u_h) \leq C\, h^2 \sum_{|\alpha|=2} \|D^\alpha u\|^2_{L_2(\Omega)} \quad \ldots \quad (3\text{-}102)$$

and, in order to give an a priori error estimate in (3-102), the problem depends upon the fact of whether or not we are able to estimate the unknown quantity $\|D^\alpha u\|^2_{L_2(\Omega)}$ in terms of the data f of the problem. For our particular example, however, it can easily be seen (see Birman and Skvortsov Theorem [19] p.234) that the differential operator defined by (3- strongly coercive, ie, there exists a constant ; nat:

$$\|D^\alpha u\|_{L_2(\Omega)} \leq \Gamma \|f\|_{L_2(\Omega)}, \text{ for all } |\alpha| \leq 2$$

and therefore, from (3-102), we can get:

$$a(u-u_h, u-u_h) \leq C\, h^2 \|f\|^2_{L_2(\Omega)} \quad \ldots \quad (3\text{-}103)$$

for some numerical constant C, which gives an upper error bound for the energy in $u-u_h$ depending on the *data of the problem only*.

On the other hand, in order to find an estimate for the error in terms of the H^1-norm, from Friedrichs's inequality (see [13] p.147), we have:

$$\int_0^1\int_0^1 [(\frac{\partial u}{\partial x})^2 + (\frac{\partial u}{\partial y})^2]dxdy \geq k\int_0^1\int_0^1 (u(x,y))^2 dxdy \qquad \ldots (3\text{-}104)$$

for some constant k, which for our particular example here is equal to 1. Then, since:

$$a(u,u) = \int_0^1\int_0^1 [(\frac{\partial u}{\partial x})^2 + (\frac{\partial u}{\partial y})^2]dxdy = \int_0^1\int_0^1 (Du)^2 dxdy \qquad \ldots (3\text{-}105)$$

by adding together the results (3-104) and (3-105) we get:

$$\int_0^1\int_0^1 [u^2 + (Du)^2]dxdy \leq 2\int_0^1\int_0^1 [(\frac{\partial u}{\partial x})^2 + (\frac{\partial u}{\partial y})^2]dxdy$$

or,

$$\|u\|^2_{H^1(\Omega)} \leq 2a(u,u) \qquad \ldots (3\text{-}106)$$

Therefore, from (3-106) combined together with the estimate (3-103), we obtain the following result:

$$\|u-u_h\|_{H^1(\Omega)} \leq C\, h \|f\|_{L_2(\Omega)}$$

which again depends only on the data of the problem.

CONCLUSIONS

The conclusion which we draw throughout the preceding discussion is that the error bounds for finite element approximations to elliptic boundary-value problems are of the following form:

$$\|u-u_h\| \leq C\, h^s |u|$$

where $\|\cdot\|$ and $|\cdot|$ represent the norm and semi-norm, respectively, of certain Sobolev spaces, h is a geometrical parameter which is closely related to the size of the elements that are used, C is some numerical constant which depends neither on the parameter h nor on the particular function u and the positive exponent s is the greatest possible exponent such that the above approximation estimate holds. However, although it is very difficult to give an explicit numerical estimate for the constant C, this is true even for the very simple cases, and the value of the quantity $|\cdot|$ can not be computed since the exact solution u is not known, the importance of this result is by no means insignificant and it provides us with the exact order of convergence of the approximate solution u_h to the exact solution u. Whenever we are able to estimate the semi-norm $|u|$ in terms of the data of the differential equation, we obtain an a priori error bound which depends on the data of the problem only.

It has been stressed throughout the analysis the wide basis of the finite element process as a general approximation tool and, although a great many papers have been published in the past few years, this will obviously remain an area for active research in the years to come. The possibility, for example, of obtaining uniform convergence - ie, convergenge in the L_∞-norm - of the approximate solution u_h to the exact solution u has by no means been completely analysed in the past and it remains an open question for the future researcher. Also, a few theoretical papers dealing with non-linear boundary-value problems as well as time dependent problems have been published and they present a challenge to the numerical analyst who is particularly concerned with error estimates.

REFERENCES

1. AGMON, S (1965): *Lectures on elliptic boundary-value problems*, Van Nostrand, Reinhold, New York
2. AUBIN, J.P. (1972): *Approximation of elliptic boundary-value problems*, Wiley, New York
3. AZIZ, A.K. (1972): *The mathematical foundations of the Finite Element Method with applications to partial differential equations*, Academic Press
4. BRAMBLE, J.H., and HILBERT, S.R. (1970): 'Estimation of linear functionals on Sobolev spaces with application to Fourier transforms and spline interpolation,' *SIAM J. Numer. Anal. 7*, pp. 112-124
5. BRAMBLE, J.H., and HILBERT, S.R. (1971): 'Bounds for a class of linear functionals with applications to Hermite interpolation,' *Numer. Math. 16*, pp. 362-369
6. CIARLET, P.G., and WAGSCHAL, C. (1971): 'Multipoint Taylor formulas and applications to the Finite Element Method,' *Numer. Math. 17*, pp. 84-100
7. CIARLET, P.G., and RAVIART, P.A. (1972): 'General Lagrange and Hermite interpolation in R^n with applications to Finite Element Methods,' *Arc. Rat. Mech. Anal. 46*, pp. 177-199
8. CIARLET, P.G., and RAVIART, P.A. (1972): 'Interplation theory over curved elements with applications to Finite Element Methods,' *Computer Methods in Appl. Mech and Eng. 1*, pp. 217-249

9. COURANT, R (1943): 'Variational methods for the solution of problems of equilibrium and vibrations,' *Bull. Amer. Math. Soc. 49*, pp. 1-23

10. DIEUDONNÉ, J. (1960): *Foundations of modern analysis*, Academic Press, New York

11. DOUGLAS, J., and DUPONT, T. (1973): 'Superconvergence for Galerkin Methods for the two point boundary problem via local projections,' *Numer. Math. 21*, pp. 270-278

12. LEAF, G.K., and KAPER, H.G. (1974): 'L^{∞}-error bounds for multivariate Lagrange approximation,' *SIAM J. Numer. Anal. 11*, pp. 363-381

13. MIKHLIN, S.G. (1964): *Variational methods in mathematical physics*, Pergammon Press, Elmsford, New York

14. MITCHELL, A.R. (1973): 'Numerical solution of partial differential equations, *NATO Advanced Study Institute*, Kjeller, Norway, 20th-24th Aug. 1973

15. MITCHELL, A.R. (1974): 'The Finite Element Method,' *J. Inst. Math. Applics. 10*, pp. 76-78

16. MORREY, C. (1966): *Multiple integrals in the calculus of variations*, Springer-Verlag, New York

17. NICOLAIDES, R.A. (1972): 'On a class of finite elements generated by Lagrange interpolation,' *SIAM J. Numer. Anal. 9*, pp. 435-445

18. NITSCHE, J. (1968): 'Ein kriterium für die Quasi-optimalität des Ritzschen Verfahrens,' *Numer. Math. 11*, pp. 346-348

19. PRENTER, P.M. (1975): *Splines and Variational Methods*, Wiley, New York.

20. SCHULTZ, M.H. (1969): 'L^{∞}-multivariate approximation theory,' *SIAM J. Numer. Anal. 6*, pp. 161-183

21. SCHULTZ, M.H. (1971): 'L^2-error bounds for the Rayleigh-Ritz-Galerkin method,' *SIAM J. Numer. Anal. 8*, pp. 737-748

22. STRANG, G. (1971): 'The Finite Element Method and approximation theory,' *Numerical solution of partial differential equations, (SYNSPADE) - II*, pp. 547-584

23. STRANG, G. (1971): 'Approximation in the Finite Element Method,' *Numer. Math. 19*, pp. 81-98

24. STRANG, G, and FIX, G. (1972): *An analysis of the Finite Element Method*, Prentice-Hall, Series in automatic computation

25. STRANG, G. (1973): 'Piecewise polynomials and the Finite Element Method, *Bull. Amer. Math. Soc. 79*, pp. 1128-1137

26. TAYLOR, A.E. (1958): *Introduction to Functional Analysis*, Wiley, New York.

27. ZIENKIEWICZ, O.C. (1971): *The Finite Element Method in Engineering Science*, 2nd ed. McGraw-Hill, New York

28. ZLAMAL, M. (1968): 'On the Finite Element Method,' *Num. Math. 12*, pp. 394-409

29. ZLAMAL, M. (1969): 'On some finite element procedures for solving second order boundary value problems,' *Numer. Math. 14*, pp. 42-48

30. ZLAMAL, M. (1970): 'A finite element procedure of the second order of accuracy,' *Numer. Math. 14*, pp. 394-402

31. ZLAMAL, M. (1973): 'The Finite Element Method in domains with curved boundaries,' *Int. J. for Num. Meth. in Eng. 5*, pp. 367-373

32. ZLAMAL, M. (1973): 'Curved elements in the Finite Element Method,' *SIAM. J. Numer. Anal. 10*, pp. 229-240.

Author Hassoulas Vasilios
Name of thesis General Multivariate Approximation Techniques Applied To The Finite Element Method. 1977

PUBLISHER:
University of the Witwatersrand, Johannesburg
©2013

LEGAL NOTICES:

Copyright Notice: All materials on the University of the Witwatersrand, Johannesburg Library website are protected by South African copyright law and may not be distributed, transmitted, displayed, or otherwise published in any format, without the prior written permission of the copyright owner.

Disclaimer and Terms of Use: Provided that you maintain all copyright and other notices contained therein, you may download material (one machine readable copy and one print copy per page) for your personal and/or educational non-commercial use only.

The University of the Witwatersrand, Johannesburg, is not responsible for any errors or omissions and excludes any and all liability for any errors in or omissions from the information on the Library website.