Mathematical analysis of the geometric mapping in the finite element method

Cedric M. A. Ayala Bravo[a], Victor Ayala[b,*]

[a]Researcher at DPR Engenharia, Campinas, Brazil.
[b]Universidad de Tarapacá, Casilla 7D Arica, Chile, And, Universidad Católica del Norte, Casilla 1280, Antofagasta, Chile.

Abstract

The finite element method FEM is an important tool used in various areas of science, where partial differential equations need to be discretized. The problem's domain is approximated by means of a geometric element partition, polyhedrons or polygons, with well defined properties. Then, a standard or reference element is associated with each distinct geometric figure present in the partition. All the operations to be made in the deformed elements are loaded instead of this in the reference element by means of an affine transformation. Thus, for example, instead of defining a numerical integration rule for each deformed element, one defines a single integration rule in the reference element, and the calculation is performed employing the affine transformation. In the case of integration of the equations, using the Jacobian of the transformation, too.

In this paper we make a rigorous analysis of the formal mathematical aspects of mapping between the finite geometric elements of zero, one, two, and three dimensions, commonly employed in the finite element theory. We show that this kind of mapping preserves all the geometric properties present between the reference element and the deformed element, alike the same number of vertices, edges, faces, and its dimension. ©2016 All rights reserved.

Keywords: Finite element method, affine transformation, geometric mapping, reference element.
2010 MSC: 74S05, 51N10.

1. Introduction

Nowadays, the use of the finite element method has expanded into a large number of applications in different areas of knowledge. The use of finite element method has grown steeply over the past

---

*Corresponding author
Email addresses: cedric@dprengenharia.com (Cedric M. A. Ayala Bravo), vayala@uta.cl, vayala@ucn.cl (Victor Ayala)

decades covering most - if not all - areas of the Exact Sciences. The advantage to using FEM as a method of numerical interpolation against the use of analytical solutions lies in the ease of implementing in different programming languages, as shown in Becker et al. [3]. Another advantage is that, with the same program, it is possible to obtain approximate solutions for different geometries, including the more complicated ones, and different combinations of boundary conditions without having to reprogram the method.

A geometric element in the usual sense is defined by a simple geometric figure, such as a closed interval in 1D, one triangle or quadrilateral in 2D, a tetrahedron, hexahedron, pyramid, and pentahedron in 3D, as defined in Ayala et al. [1, 2, 6]. These elements are introduced with the objective of providing an interpolation for the solution in a small portion of the domain. Therefore, a polynomial space of certain degree is associated to each element. All of it is commonly constructed by the FEM defining certain points within the element called nodes and associating polynomials of certain degree to them, Zienkiewicz et al. [8]. The polynomials defined like this constitute a basis set. Thus, the interpolation space of the element is generated by the combination of basic polynomials. Basically, the finite element is composed by the geometric element plus the interpolation space associated.

From this small brick or interpolating unit the solution is estimated by juxtaposition in whole domain according to the systematics of the FEM method. There is an extensive and good literature that teaches the basics of FEM method. For a detailed description of the polynomial approximation of a family of hierarchical finite elements see Ayala et al. [1, 6].

In this work we make a rigorous mathematical study to define accurately which is the mapping used to obtain the deformed finite element that has all the good properties for the correct application of the method.

2. Mathematical description of the geometric mapping

It is possible to define the topological space of the geometric element as the convex wrapping of certain special points, the nodes. Let n be a positive integer and $A \subset \Re^n$.

Definition 2.1. A is said to be a convex set if for each $x, y \in A$ we have

$$t\, x + (1 - t)\, y \in A \ , \ 0 \le t \le 1.$$

That is, A contains the segment joining any two points in A.

Definition 2.2. The convex hull of A, denoted *conv* A, is defined as the smallest convex set of $\Re^n$ containing A.

Remark 2.3. From the above definition it follows that the convex hull of A is given by:

$$conv\, A = \cap F.$$

where, F is the family of all convex sets in $\Re^n$ containing the set A.

Definition 2.4. Geometric elements of reference:

$$Pt = conv\, \{0\}.$$

$$L = conv\, \{-1, 1\}.$$
$$T = conv\, \{(0,0), (1,0), (0,1)\}.$$

$$Q = conv\{(-1,-1), (1,-1), (1,1), (-1,1)\}.$$

$$H = conv\{(-1,-1,-1), (1,-1,-1), (1,1,-1), (-1,1,-1),$$
$$(-1,-1,1), (1,-1,1), (1,1,1), (-1,1,1)\}.$$

$$Te = conv\{(0,0,0), (1,0,0), (0,1,0), (0,0,1)\}.$$

$$Pr = conv\{(0,0,-1), (1,0,-1), (0,1,-1), (0,0,1), (1,0,1), (0,1,1)\}.$$

$$Pi = conv\{(-1,-1,0), (1,-1,0), (1,1,0), (-1,1,0), (0,0,1)\}.$$

This nomenclature refers to a point, line, triangle, quadrilateral, hexahedron, tetrahedron, prism and pyramid, respectively.

In fact it is possible to demonstrate the following result.

Proposition 2.5. *The convex hull of $A \subset \Re^n$ is the set of all convex combinations of points in A, that is,*

$$conv\, A = \left\{ \sum_{j=1}^{k} a_j x_j \; : \; x_j \in A \text{ and such that } \sum_{j=1}^{k} a_j = 1 \right\}$$

Definition 2.6. Let X and Y be two topological spaces. A bijective and continuous application $f : X \to Y$ is said to be a **homeomorphism** if the function $f^{-1} : Y \to X$ is continuous.

Remark 2.7. Note that each geometric reference element is a compact set.

Remark 2.8. Consider the function $f : (L - \{1\}) \to S^1$ defined by $f(x) = e^{2\pi i x}$. Clearly f is bijective and continuous. Although the application f^{-1} is not continuous in $(1, 0)$. Indeed, there are points on the circle S^1 near to the point $(1,0)$ separated of f^{-1} by a distance greater than $\frac{1}{2}$. To see this consider the points $(1-\epsilon, +\sqrt{1-(1-\epsilon)^2})$ and $(1-\epsilon, -\sqrt{1-(1-\epsilon)^2})$ the circle with ϵ sufficiently small.

In this work we are interested in the study of the finite element as geometric entities that are constructed from deformations obtained in the geometric elements of reference declared in the Definition 2.4. In particular, it is desired that the topology of the deformed geometry elements be determined by the topology of its referential. So it is necessary and sufficient to consider the functions of deformation as homeomorphisms, Lima [7].

In particular the topological varieties obtained are compacts and the topological dimension of any deformed geometric element EG obtained by the homeomorphism f equals the dimension of the variety with boundary $f^{-1}(EG)$. In fact, let (U, h) be a **chart** of EG, that is, U is an opening of the topology of EG and h,

$$h : U \to h(U) \subset \Re^s$$

is a homeomorphism for any $s \in \{1, 2, 3\}$. Then,

$$(f^{-1}(U), h \circ f)$$

is a chart for $f^{-1}(EG)$. Indeed, $f^{-1}(U)$ is an opening of the referential and

$$h \circ f : f^{-1}(U) \to h(U)$$

is a homeomorphism, Figure 1. Thus, the dimensions coincide.
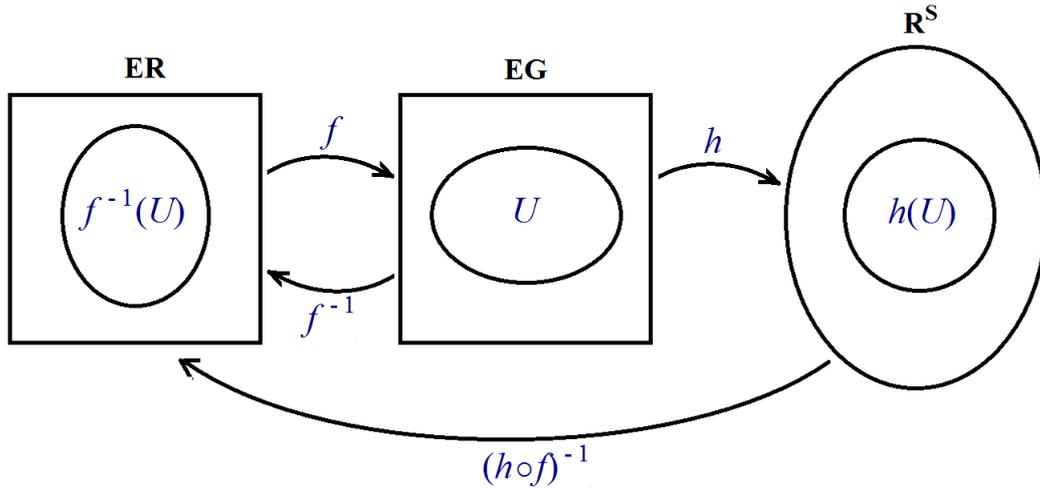
Figure 1: Chart of the deformed geometric element EG.

Observe that this is a consequence of the definition of the possible deformations that have been considered, the geometric transformations, which will be explicated below in details. Note that, for example, the circle S^1 can be constructed via one homeomorphism, from the boundary of a triangle or a rectangle or in general from a planar polygonal of $k - sides$, for each $k \geq 3$. Since we don't want that type of anomaly, we consider only certain subspaces of homeomorphism: *those that preserve the essential characteristics of the reference elements.* More precisely, the idea is to consider the classes of differentiable and bijective functions with inverse differentiable, that is, global **diffeomorphisms** that deform the reference geometric elements in surfaces which border is contained in $\Re^3$, keeping in essence the vertices number.

Denote by O any reference geometric element and $f(O)$ the deformed element in $\Re^3$, constructed from the diffeomorphism f

$$f : O \subset \Re^s \to f(O) \subset \Re^3 , \ s = 1, 2, 3,$$

where s depends on the dimension of O. Hereafter, an appropriated mathematical object is used to define vertices, edges, faces, and the inside of a deformed element $f(O)$. In the first place, for $x \in O$ is introduced the concept of the **Tangent Cone** generated by x, and denoted by C_x as follows:

$$C_x = \big\{ \dot{\gamma}(0) | \gamma : (-\epsilon, \epsilon) \to O, \ continuous, \ differentiable \ in \ t = 0, \ and \ \gamma(0) = x \big\}.$$

In words, the C_x cone generated by x, is the set which elements are the derivatives in $t = 0$ from the curves γ (vectors) differentiable in 0, continuous in its domain, that go through x, that is, $\gamma(0) = x$, is such that exists a number ϵ non negative, sufficiently small in such a way that the image $\gamma(-\epsilon, \epsilon)$ is contained in the reference element O. It is important to observe the following:

$$v \in C_x \Leftrightarrow \lambda v \in C_x, \forall \lambda \gneq 0.$$

A question that justified the cone name. A tangent cone is said **Proper Cone** if it does not contain a nontrivial subspace[1], that is, different from $\{0\}$.

---

[1] *Proper subspaces of R^2 can be $\{0\}$ and lines through the origin. Proper subspaces of R^3 can be $\{0\}$, lines and planes through the origin.*
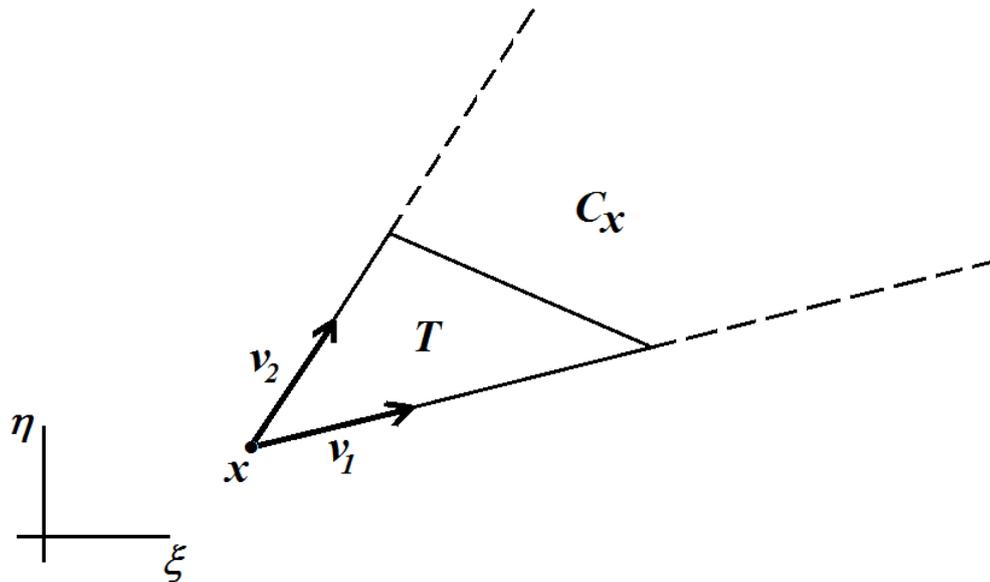
Figure 2: Tangent cone to vertex x.

Example 2.9. To elucidate some ideas take as an example the triangle T.

- Case 1: In case $x \in \overset{\circ}{T}$ (interior of T) then $C_x = T_x \Re^2$, that is, the tangent cone coincides with the tangent plane in the point $x \in \Re^2$, the proper $\Re^2$.

- Case 2: If x belongs to an edge of T but is not a vertex, then, C_x is a semiplane of $T_x \Re^2$.

- Case 3: If x is a vertex then C_x is a proper cone, that is, C_x does not contain any nontrivial subspace of $T_x \Re^2$. Indeed, this cone is generated from the linear combinations with non negatives coefficients of the vectors constructed with the triangle sides whose intersection determines the vertex itself, Figure 2.

Example 2.10. In case of tetrahedron:

- Case 1: If $x \in \overset{\circ}{Te}$ then $C_x = Te_x \Re^2$, that is, the tangent cone coincides with all the space at $x \in \Re^3$ (the $\Re^3$ itself).

- Case 2: If x belongs to a face of Te but it is not within an edge or vertex, then C_x is a semispace of the $Te_x \Re^3$, a plane.

- Case 3: If x belongs to an edge of Te but it is not a vertex of Te, then C_x is a cone that contains a single subspace of dimension 1 of $Te_x \Re^3$, a straight line. Indeed, this cone is constructed with non negative linear combinations of vetoers starting from x and that belongs to faces whose intersection determines the edge in question and arbitrary coefficients of any non null vectors that belongs to this edge.

- Case 4: If x is a vertex of Te, then C_x is a proper cone, that is, it only contains the subspace $\{0\}$. This cone is generated from the linear combinations with non negative coefficients of three
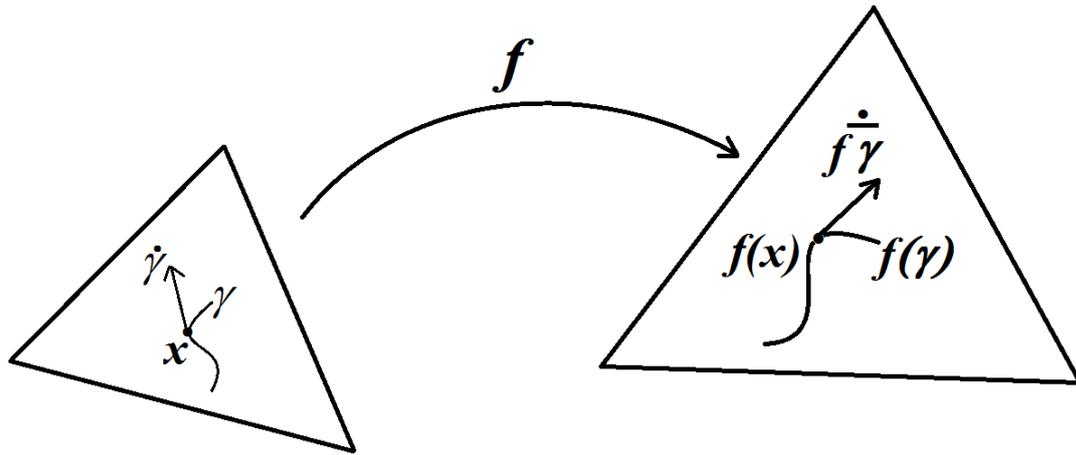
Figure 3: Cone elements in x and in $f(x)$.

vectors constructed with the three edges whose intersection determines the vertex in question.

Since $f(O)$ is a compact surface with boundary, it is possible to define in analogous way to C_x the tangent cone to $f(O)$ in point $f(x)$ in a natural manner, as it follows.

$$C_{f(x)} = \left\{ (f \circ \gamma)^{\cdot}(0) \,|\, \gamma : (-\epsilon, \epsilon) \to O, \ continuous, \ differentiable \ in \ t = 0, \ and \ \gamma(0) = x \right\}.$$

Next, $C_{f(x)}$ is a set of elements that are initiated in $f(x)$ and that are obtained from tangent vectors at continuous curves the type $f \circ \gamma$ over an interval $(-\epsilon, \epsilon)$ and whose derivative exist in $t = 0$, Figure 3.

In others words, the cone in $f(x)$ is an image from the cone in x by the Jacobian transformation, or

$$C_{f(x)} = Jac(f)(x)(C_x),$$

where

$$Jac(f)(x) = [Df(x)]_{can}^{can}$$

and the linear transformation

$$Df(x) : T_x \Re^3 \to T_x \Re^3$$

is the derivative of f in point x, that is, the linear approximation of f in x. From the definition of derivative it follows:

$$v \in T_x \Re^s \Rightarrow Df(x)(v) = \frac{d}{dx} f(\gamma(t))|_{t=0},$$

when $\gamma(t)$ is justly a continuous curve defined on the interval of type $(-\epsilon, \epsilon)$ differentiable in $t = 0$ and such that $\gamma(t) = v$. That proves that the Jacobian of f carry any tangent cone in a tangent cone. Once that $f : O \to f(O)$ admits to being a diffeomorphism, in particular we conclude that for each $x \in O$,

$$Df(x) : T_x \Re^s \to T_{f(x)} f(O)$$

is an isomorphism between tangent subspaces. In particular, if S is a subspace of k dimension of $T_x\Re^s$, $0 \le k \le s$, then $Df(x)(S)$ is a subspace of k dimension. Consequently, if f is a diffeomorphism, for each $x \in O$ it follows:

- Case 1: If $x \in \overset{\circ}{O}$

  Consider $f(O)$ a topological space with the induced topology from $\Re^3$, then

  $$C_x = T_x\Re^s \Rightarrow Df(x)(T_x\Re^s) = T_{f(x)}f(O) \Rightarrow f(x) \in \overset{\circ}{f(O)}.$$

- Case 2: If $x \in \partial O$ then the following situations are given:

  ○ for $s = 1$

  C_{-1} and C_1 are semi-straight and the same happens with $C_{f(-1)}$ and $C_{f(1)}$.

  ○ for $s = 2$

  1. if x belongs to interior of an edge A of O. Then C_x is a semi-plane of $T_x\Re^2$ with support the straight line L_A generated by the edge A. Next, $C_{f(x)}$ is a semi-plane of $T_{f(x)}f(O)$ with support to the subspace $Df(x)(L_A)$;

  2. if x is a vertex, in this case the proper cone C_x is taken by $Df(x)$ in the proper cone $C_{f(x)}$. Indeed, if $C_{f(x)}$ contains one nontrivial subspace then $Df^{-1}(f(x))(C_{f(x)}) = C_x$ would contain one nontrivial subspace;

  ○ for $s = 3$

  1. if x belongs to the inside of the face F of O. Then C_x is one semi-space of $T_x\Re^3$ and contains a single bidimensional subspace S. Then, $C_{f(x)}$ is one semi-space that is contained in one of the connected components of $T_{f(x)}f(O)$ that determines the subspace $Df(x)(S)$;

  2. if x belongs to the interior of one edge A of O. Then, C_x is one cone that contains a single unidimensional subspace L_A, the one generated by the edge A. Soon, the same happens with $C_{f(x)}$, that is, it is one cone in $T_{f(x)}f(O)$ that contains a single unidimensional subspace $Df(x)(L_A)$;

  3. if x is one vertex. In this case, both C_x and $C_{f(x)}$ are proper cones. Indeed, neither one of two cones contains nontrivial subspaces.

Based on previous deductions, it follows that the admissible classes of deformations of the reference geometric element are exactly the diffeomorphism classes.

Definition 2.11. Let $f : O \subset \Re^s \to f(O) \subset \Re^3$ be a geometric element.

(i) One point $x \in f(O)$ is said to be a vertex if, $C_x \subset T_x f(O)$ is a proper cone.

(ii) A curve $a : [0, 1] \to O$ is said to be an edge of $f(O)$,

  (a) $a(0)$ and $a(1)$ are vertices;

  (b) for each point x in the inside of an edge A, the tangent cone C_x contains a single subspace of $T_x f(O)$ of the dimension 1.

(iii) A set $F \subset f(O)$ is said to be a face of $f(O)$ if:

  (a) for each point $x \in \overset{\circ}{F}$, the tangent cone C_x contains a single subspace of dimension 2;

  (b) ∂F consists of exactly $n - vertices$ and $n - edges$, in the case of a $n - polygonal$ face[1].

With these premises in mind, the following result has been proven.

---

[1]A triangle face have 3 vertices and 3 edges. One quadrilateral face have 4 vertices and 4 edges, etc.

Theorem 2.12. *Let O be a reference element and $f : O \subset \Re^s \to f(O) \subset \Re^3$. Then, the diffeomorphism f takes vertex by vertex, edge by edge, face by face and volume by volume. In particular, the cardinality of vertices, edges, and faces of the elements O and $f(O)$ coincide.*

Remark 2.13. Despite that both Q and Te have the same cardinality of vertices, there is no ambiguity with the previous definition. This is due to the fact that the reference elements belong to the environment with spaces of different dimensions, dimension that is kept by the diffeomorphism f, as explained previously. Obviously, f is a homeomorphism. In other words, $f(Q)$ cannot be diffeomorphic to $g(Te)$ to any diffeomorphisms f and g, because of the theorem of invariance of Brouwer [5].

Example 2.14. We finished illustrating different types of finite elements where such geometric transformations are present.

(1) **Standard elements**. Deformed elements like the image of a linear transformation. The inverse is also a linear transformation. For elements with straight edges and flat faces this is a common transformation in standard finite element programs. They are known as Affine Transformation.

(2) **Isoparametric elements**. The T transformation is defined using the shape functions associated with the vertices of the elements. Then, the deformed element is defined as follows:

$$T(x) = \sum_{i=1}^{n} v_i \psi_i(x),$$

where v_i represents the vertices of the deformed element, $i = 1, 2, ..., n$. The shape functions of the reference elements are obtained satisfying $\psi_i(x_j) = \delta_{ij}$, where the last symbol is the Cronecker's delta. Thus the functions are linear over the element's edge. A less trivial example is shown in Ayala et al. [2] for the pyramid element of high $p - order$ with rational terms.

(3) **Curved elements**. In general the transformation is non-linear and can be given by a polynomial of higher degree. Sometimes this kind of element can be obtained from standard reference elements, as given in Bernadou [4]. This reference presented specific mathematically well prepared examples to demonstrate all the necessary properties. Curved elements may also be isoparametric elements (or not).

References

[1] C. M. A. Ayala Bravo, *Um sistema de refinamento H-P adaptativo utilizando elementos finitos hierárquicos multidimensionais*, Tese de doutorado. Faculdade de Engenharia Mecânica, Universidade Estadual de Campinas, Campinas, SP, (2000).1

[2] C. M. A. Ayala Bravo, R. Pavanello, P. R. B. Devloo, J. L. D. Calle, *Definition of a P-Interpolating Space of Hierarchical Bases of Finite Elements on the Pyramid*, Linear Algebra Appl., **460** (2014), 174–204. 1, 2

[3] E. B. Becker, G. F. Carey, J. T. Oden, *Finite elements*, Englewood Cliffs, NJ: Prentice-Hall, (1981). 1

[4] M. Bernadou, C^1 *curved finite elements with numerical integration for thin plate and thin shell problems, Part 1: Construction and interpolation properties of curved C^1 finite elements*, Computer Methods Appl. Mech. Eng., **102** (1993), 255–289. 3

[5] L. E. J. Brouwer, *Beweis der Invarianz des n-dimensionalen Gebiets*, Math. Ann., **71** (1912), 305–315, see also **72** (1912), 55–56. 2.13

[6] P. R. B. Devloo, C. M. A. Ayala Bravo, E. C. Rylo, *Systematic and generic construction of shape functions for p-adaptive meshes of multidimensional finite elements*, Comput. Methods Appl. Mech. Engrg., **198** (2009), 1716–1725. 1

[7] E. L. Lima, *Elementos de Topologia Geral. IMPA - Instituto de Matemtica Pura e Aplicada*, Editora da Universidade de São Paulo. Rio de Janeiro, (1970). 2

[8] O. C. Zienkiewicz, R. L. Taylor, J. Z. Zhu., *The finite element method: Its Basis and Fundamentals (Sixth ed.)*, Butterworth-Heinemann, (2005). 1