Dynamic Analysis of a One-cylinder Engine Crankshaft

P. Horváth, J. Égert

Department of Applied Mechanics, Faculty of Mechanical Engineering
Information Technology and Electrical Engineering
Széchenyi István University
9026 Győr 1 Egyetem tér, Hungary
horvpe@sze.hu, egert@sze.hu

Abstract: The paper deals with the vibration analysis of a one-cylinder engine crankshaft. For the analysis, a sequential torsional multi-mass vibration model was created of the crank-mechanism. The natural frequencies and natural modes of the crankshaft were determined from this model. The results of the torsional model were verified with a 3D finite element model, which consists of not only the torsional, but every other natural frequency and mode of the crankshaft. From the torsional natural frequencies and the harmonic expansion of the excitation forces, the critical speeds were determined, where resonance phenomena can occur. The relative angular deflection of the two ends of the crankshaft was also determined which characterizes the torsional stiffness of the crankshaft. From the angular velocity function of the clutch, the measure of speed irregularity of the clutch was defined.

Keywords: crankshaft, multi-mass model, vibration, natural frequency, natural mode

1. Introduction

As result of periodic loading in time, every mechanism that contains elastic elements – e.g. the crankshaft of an internal combustion engine – vibrations may occur. The investigated crank-mechanism can be regarded as a vibration system from a mechanical point of view. In this vibration system under specific conditions resonance phenomena may occur. In the analysis, the goals were:
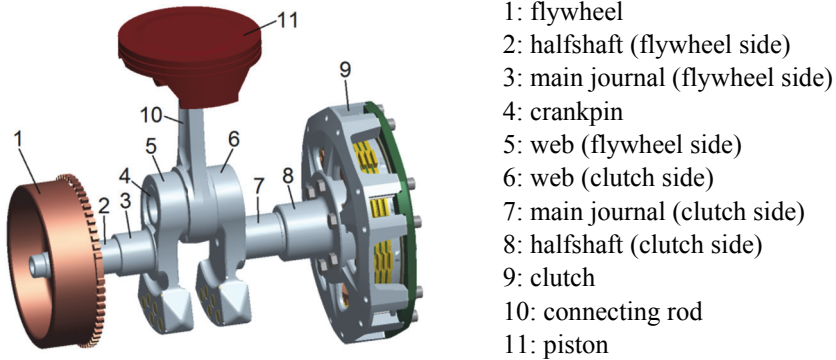
- to perform a free vibration analysis of the crank-mechanism in order to obtain the natural frequencies and natural modes of the crankshaft,
- to investigate the effect of the excitation loads,
- to determine critical speeds, where resonance phenomenon can occur,
- to define the relative angular deflection of the two ends of the crankshaft, which is characteristic for the stiffness of the crankshaft,
- to define the measure of speed irregularity of the clutch, which has a great impact on the fatigue life of the gearbox.

The investigation and clarification of the abovementioned points is crucial because of the constructional modifications [1]. Since the stiffness of the crankshaft has been

reduced [1], the structure has become more sensitive to excitation loads, to vibration and to resonance phenomena.
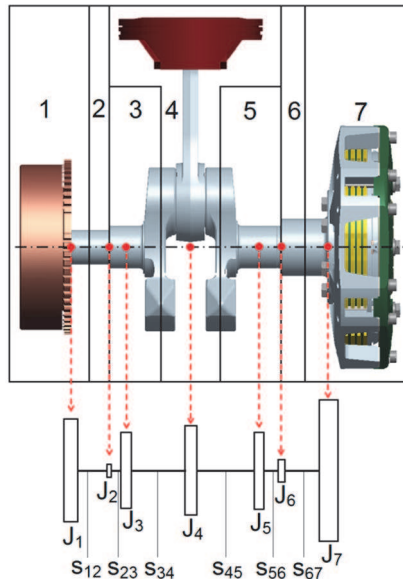
2. Creating the torsional multi-mass model of the crank mechanism

For the torsional analysis of the crankshaft (Fig. 1.) a simple multi-mass model was created from the original crank-mechanism which only describes the torsional vibrations of the crankshaft.

1: flywheel
2: halfshaft (flywheel side)
3: main journal (flywheel side)
4: crankpin
5: web (flywheel side)
6: web (clutch side)
7: main journal (clutch side)
8: halfshaft (clutch side)
9: clutch
10: connecting rod
11: piston

Figure 1. Parts of the analysed crank mechanism

A seven-degree of freedom sequential multi-mass torsional model was created (Fig. 2.) for modelling the crank mechanism (Fig. 1.) according to [2], [3] and [4].

Figure 2. The sequential torsional multi-mass model of the crank mechanism

In this model, the rigid mass elements are described by moments of inertia and elastic elements are characterized by torsional stiffnesses. The unknown parameters, which should be calculated, are the angular deflections of the rigid mass elements. As the

unknown parameters are only angular deflections, the mass properties can be given simply by moments of inertia. The main bearings of the crankshaft have no effect on the angular deflections, so the torsional multi-mass model has no kinematic boundary conditions, it is a so-called free vibration model. In order to create the torsional multi-mass model, the mass- and stiffness reduction of the crankshaft had to be performed.

2.1 Mass-reduction of the crankshaft

The moments of inertia for a simple shaft, journal or pin (parts 2, 3, 4, 7 and 8 in Fig. 1), calculated with respect to the rotational axis of the crankshaft, can be calculated as follows:

$$ J = mr^2 + \frac{1}{2}m\left(\frac{D^2}{4} + \frac{d^2}{4}\right), \tag{1} $$

where m is the mass of the given part, d is the inner diameter, D is the outer diameter of the part and r is the distance from the rotation-axis (in case of the crankpin, it is the throw of crank, in other cases it is zero).

The moments of inertia of the webs and counterweights were gained from the CAD model.

The moments of inertia of the connecting rod and piston reduced to the rotation-axis are function of the crank angle. In this case an equivalent moment of inertia was calculated according to [12]:

$$ J_e = \left[m_{hA} + \left(m_{hB} + m_d \right)\left(\frac{1}{2} + \frac{\lambda^2}{8} \right) \right] r^2, \tag{2} $$

where: - $m_{hA} = (1-u_h)m_h$ is the rotating mass part of the connecting rod, where u_h is the length ratio of the big end - center of gravity distance to the big end - small end distance and m_h is the mass of the connecting rod,

- $m_{hB} = u_h m_h$ is the alternating mass part of the connecting rod,
- m_d is the mass of the piston,
- r is the length of the throw of crank,
- $\lambda = r / l$ is the length ratio of throw of crank to the connecting rod length

In *Table 1.* the moments of inertia of each mass element of the crank mechanism, obtained from the formulas (1) and (2) and CAD calculations [2] [3] are summarized.

From the ten moments of inertia in Table 1. seven moments of inertia were calculated by the formulas in Table 2., which are assumed to be mass elements in the ends of the shafts and in the middle of the pins (Fig. 2.). In Table 2. these seven moments of inertia and their derivation from the former ten moments of inertia shown in Table 1. are summarized.

Table 1. Moments of inertia of the parts of the crank mechanism

Symbol of moment of inertia	Description of the part of the crank mechanism	Value of moment of inertia [kgm^2]
J_l	flywheel	$2{,}906 \cdot 10^{-3}$
J_{vt}	halfshaft (flywheel side)	$6{,}57157 \cdot 10^{-6}$
J_{vcs}	main journal (flywheel side)	$1{,}44886 \cdot 10^{-5}$
J_{vk}	web (flywheel side)	$1{,}135432 \cdot 10^{-3}$
J_{fcs}	crankpin	$1{,}23428 \cdot 10^{-4}$
J_{hk}	web (clutch side)	$1{,}135622 \cdot 10^{-3}$
J_{hcs}	main journal (clutch side)	$1{,}57778 \cdot 10^{-5}$
J_{ht}	halfshaft (clutch side)	$6{,}9956 \cdot 10^{-5}$
J_k	clutch	$6{,}206 \cdot 10^{-3}$
J_e	piston and connecting rod	$4{,}343 \cdot 10^{-4}$

Table 2. The moments of inertia of the mass elements used for the torsional analysis

Symbol of mass element	Description of mass element	Calculation of moment of inertia of the mass elements	Moment of inertia J_i [kgm^2]
$i=1$	flywheel	$J_1 = J_l + J_{vt} / 2$	$2{,}910 \cdot 10^{-3}$
$i=2$	halfshaft 1 - main journal 1	$J_2 = \dfrac{J_{vt}}{2}$	$3{,}286 \cdot 10^{-6}$
$i=3$	main journal 1	$J_3 = J_{vcs} + J_{vk} / 2$	$5{,}822 \cdot 10^{-4}$
$i=4$	crankpin	$J_4 = \dfrac{J_{vk}}{2} + \dfrac{J_{hk}}{2} + J_{fcs} + J_e$	$1{,}693 \cdot 10^{-3}$
$i=5$	main journal 2	$J_5 = J_{hcs} + J_{hk} / 2$	$5{,}836 \cdot 10^{-4}$
$i=6$	main journal 2 - halfshaft 2	$J_6 = \dfrac{J_{ht}}{2}$	$2{,}998 \cdot 10^{-5}$
$i=7$	clutch	$J_7 = J_k + J_{ht} / 2$	$6{,}206 \cdot 10^{-3}$

2.2 Stiffness-reduction of the crankshaft

For the sequential torsional vibration system, six elastic elements are needed between the above described seven reduced rigid mass elements. Fig. 3. shows these elastic elements of the crankshaft.

In case of shafts and pins, the stiffness can be calculated with the following formula:

$$s = \frac{M_{cs}}{\varphi} = \frac{I_p G}{l} \,, \tag{3}$$

where I_p is the polar moment of inertia of the cross section area, G is the shear modulus of the crankshaft's material and l is the length of the shaft / pin.
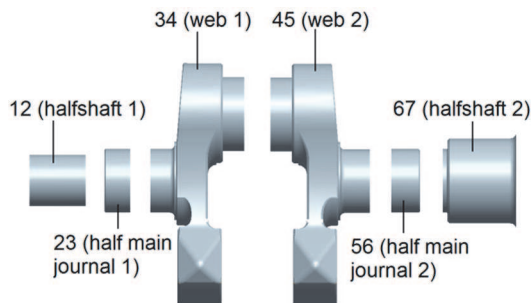
Figure 3. The six elastic elements connecting the rigid mass elements.

In case of webs, the stiffness cannot be calculated with a simple formula. In this paper, the web's stiffness was calculated with the aid of the AutoSHAFT modul of AVL Excite Designer software [2]. This modul automatically creates the FEM model of the crank web and then with static FEM analysis it calculates the torsional stiffness of the webs, applying a tentative torque on the webs. In Table 3. the stiffness of the elastic elements are summarized.

Table 3. Stiffnesses of the connecting elastic elements

Symbol of connecting element	Description of connecting element	Value of stiffness s_i [Nm/rad]
$i=12$	halfshaft 1	$9,628 \cdot 10^4$
$i=23$	half main journal 1	$5,399 \cdot 10^5$
$i=34$	web 1	$1,766 \cdot 10^5$
$i=45$	web 2	$1,643 \cdot 10^5$
$i=56$	half main journal 2	$4,293 \cdot 10^5$
$i=67$	halfshaft 2	$2,024 \cdot 10^5$

3. The natural frequencies and natural modes from the multi-mass model

A free sequentional torsional model consisting of *n* mass has (*n*-1) non zero natural frequencies and modes, so the above created multi-mass model consisting of seven mass elements has six non zero natural frequencies. The determination of the movement of this multi-mass model can be expressed using *Lagrange's* equations [8]:

$$\frac{d}{dt}\left(\frac{\partial E}{\partial \dot{\varphi}_i} \right) - \frac{\partial E}{\partial \varphi} = Q_{si} ,$$
(4)

where E is the kinetic energy of the whole torsional model, φ_i is the angular deflection of mass element i and Q_{Si} is the spring force. Executing the differentiations in (4) the following second order differential equation system is obtained:

$$\underline{\underline{M}}\, \underline{\ddot{\varphi}} + \underline{\underline{K}}\, \underline{\varphi} = \underline{0} ,$$
(5)

where $\underline{\varphi}$ is the vector of angular deflection (6) of the mass elements, $\underline{\underline{M}}$ is the mass matrix (7) and $\underline{\underline{K}}$ is the stiffness matrix (8).

$$\underline{\varphi}^T = \begin{bmatrix} \varphi_1 \varphi_2 & \dots & \varphi_7 \end{bmatrix} \tag{6}$$

$$\underline{\underline{M}} = \begin{bmatrix} J_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & J_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & J_3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & J_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & J_5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & J_6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & J_7 \end{bmatrix} \tag{7}$$

$$\underline{\underline{K}} = \begin{bmatrix} s_{12} & -s_{12} & 0 & 0 & 0 & 0 & 0 \\ -s_{12} & s_{12}+s_{23} & -s_{23} & 0 & 0 & 0 & 0 \\ 0 & -s_{23} & s_{23}+s_{34} & -s_{34} & 0 & 0 & 0 \\ 0 & 0 & -s_{34} & s_{34}+s_{45} & -s_{45} & 0 & 0 \\ 0 & 0 & 0 & -s_{45} & s_{45}+s_{56} & -s_{56} & 0 \\ 0 & 0 & 0 & 0 & -s_{56} & s_{56}+s_{67} & -s_{67} \\ 0 & 0 & 0 & 0 & 0 & -s_{67} & s_{67} \end{bmatrix} \tag{8}$$

The differential equation system (4) leads to the (9) eigenvalue problem, the eigenvalues of which provide the natural circular frequencies ω_i, while the eigenvectors provide the natural modes $\underline{\varphi_0}_i$ for the natural circular frequency ω_i
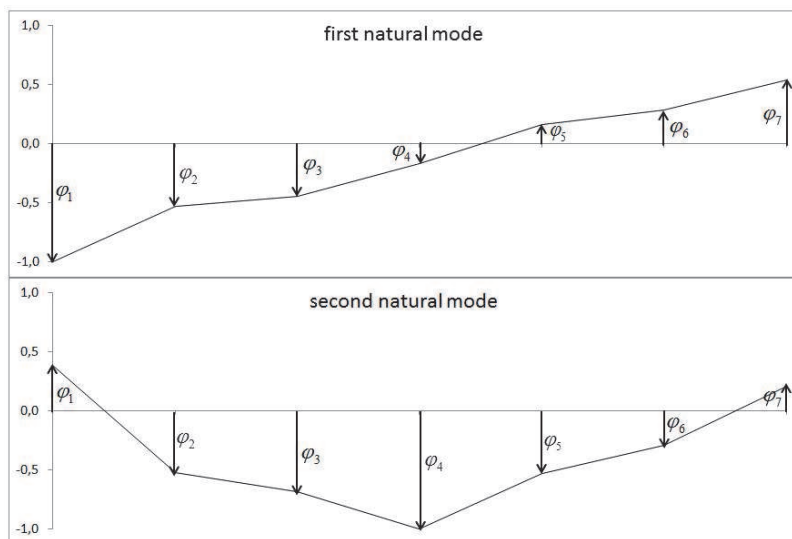
$$\left(\underline{\underline{K}} - \omega_i^2 \underline{\underline{M}} \right) \underline{\varphi_0}_i = \underline{0} . \tag{9}$$

Table 4.
The natural circular frequencies and natural frequencies of the multi-mass model

	Natural circular frequencies (rad/s)	Natural frequencies (Hz)
ω_1	3944,6	627,8
ω_2	8857,4	1409,7
ω_3	21934,6	3491,0
ω_4	24975,0	3974,9
ω_5	146949,3	23387,7
ω_6	440913,7	70173,6

In Table 4. the natural circular frequencies and the natural frequencies are summarized, while Fig. 4. illustrates the first two natural modes, which belong to the lowest two natural frequencies. The natural modes illustrate the relative ratio and

direction of the amplitudes of the angular deflection of the mass elements of the multi-mass model vibrating on the given natural frequency. The actual vibration amplitudes will be influenced by the excitation forces and the damping of the crank mechanism. It can be observed that the first natural mode has one non-vibrating (non-rotating) point, while the second mode has two, so the number of the non-vibrating points increases according to the number of the natural modes. In case of engines, in general, it is important to know only the first two natural modes, as the detectable harmonics of the excitation torques – as it will be seen in the next chapter – are only in the range of the first two natural frequencies.

Figure 4. First two natural modes of the crankshaft

4. The natural frequencies and natural modes calculated by FEM modeling

In this point the natural frequencies and natural modes will also be calculated by FEM method. In order to verify the result of the multi-mass model, these results will be compared to the former values.

For the vibration analysis of the FEM model, the Modal modul of ANSYS software [5] was used. From the crankshaft–flywheel–clutch assembly, the FEM model was created only for the crankshaft, the flywheel and the clutch were added as mass elements, characterized by their mass and their moments of inertia.

As a kinematic boundary condition in static analysis, an elastic support was used on the contact surface of the main bearings [1]. At free vibration analysis kinematic boundary condition does not need to be defined. In an eigenvalue problem dynamic boundary conditions (loadings) are not needed either.

The FEM analysis has covered only the first two natural frequences which are the most important ones from the resonance phenomena point of view [6] – [8]. Fig. 5.

illustrates the first two natural modes belonging to the lowest two natural frequencies of the crankshaft.
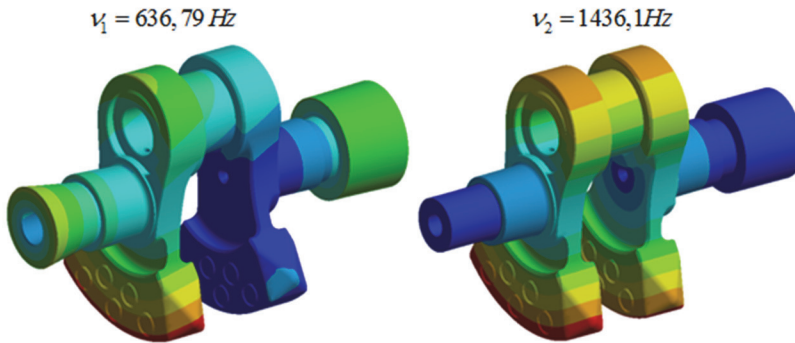
$$\nu_1 = 636,79\,Hz \qquad\qquad \nu_2 = 1436,1\,Hz$$

Figure 5. The first two natural modes belonging to the lowest two natural frequencies

In *Table 5*. the first two natural frequencies of the FEM model are compared with the first two natural frequencies of the torsional multi-mass model.

Table 5. Comparison of the natural frequencies of the torsional multi-mass model with those of the FEM model

Natural frequencies of the multi-mass model (Hz)	Natural frequencies of the FEM model (Hz)	Relative error [%]
627,8	636,79	1,43
1409,7	1436,1	1,87

As can be seen from Table 5. the first two natural frequencies calculated from the torsional multi-mass model and those calculated from the FEM model show a very good match, the relative error is below 2 % in both cases. Furthermore, the natural modes obtained from the FEM model come close to those obtained from the multi-mass model. On the basis of this, it can be stated that from the torsional vibrations aspect, in the frequency range of operation, the multi–mass model comes close to the real system in respect to natural frequencies and natural modes.

5. Forced torsional vibrations

The periodic tangential force generates not only a static torque on the crankshaft, but it causes additional torsional vibrations as well. Both the tangential force – crank angle function and the torque – crank angle function, due to the periodical characteristic of these functions, they can be expanded into harmonic series. In course of the harmonic series development, the torque – crank angle function is expanded into an infinite series, where the terms of the series are harmonic (sine and cosine) functions. The harmonic expansion can be performed mathematically with the Fourier-Transformation. According to this, the torque – crank angle function $M(\alpha)$ can be written with the use of harmonic functions as the following series [9]:

$$M(\alpha) = M_k + 2\sum_{n=1}^{\infty} a_n \cos\left(\frac{2\pi}{T} n\alpha\right) + 2\sum_{n=1}^{\infty} b_n \sin\left(\frac{2\pi}{T} n\alpha\right), \tag{10}$$

where M_k is the mean torque, while a_n and b_n coefficients can be calculated with the following formulas:

$$a_n = \frac{1}{T}\int_0^T M(\alpha)\cos\left(\frac{2\pi}{T} n\alpha\right) d\alpha, \qquad b_n = \frac{1}{T}\int_0^T M(\alpha)\sin\left(\frac{2\pi}{T} n\alpha\right) d\alpha. \tag{11}$$

In case of a four stroke engine the whole engine cycle is $T = 4\pi$. Substituting this value into the equations (11), the argumentums of the sine and cosine functions will be $\frac{n}{2}\alpha$, where $x = \frac{n}{2}$ is the n-th order of the series development, which is the number of harmonic oscillations during a whole engine cycle. So in case of a four stroke engine, the harmonics will have half orders, as in case of $n = 1, 2, 3$ we get $x = 0,5, 1, 1,5$ oscillation in the harmonic functions.

With the amplitude of each harmonic, the spectrum of the exciting harmonics can be obtained. The spectrum represents the amplitude of the harmonics in function of the order. Fig. 6–8. show the spectrum of the exciting gas-, mass- and resultant rotation-forces which are proportional to the torques caused by the tangential gas-, mass- and resultant forces in case of crankshaft speed of 6500 rpm.
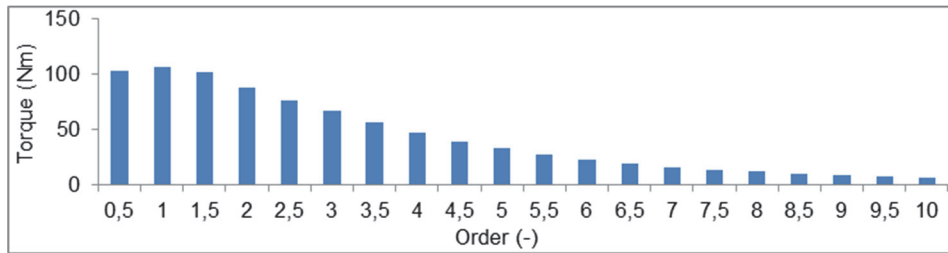

Figure 6. Spectrum of the gas rotational excitation harmonics at crankshaft speed of 6500 rpm
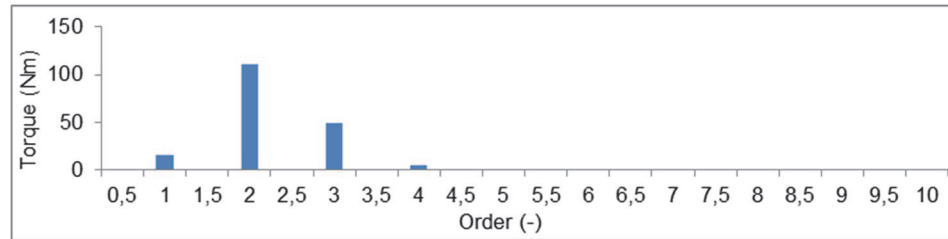

Figure 7. Spectrum of the mass rotational excitation harmonics at crankshaft speed of 6500 rpm

It can be seen from the spectrum of the excitation harmonics (Fig. 6. – 8.) that the rotational mass forces appear only in case of integer orders. Furthermore, it can be observed from the column diagrams that as an effect of the rotational mass forces, the

amplitude of the first order rotational gas force increases, while the amplitude of the second, third and slightly the fourth order rotational gas force decreases.
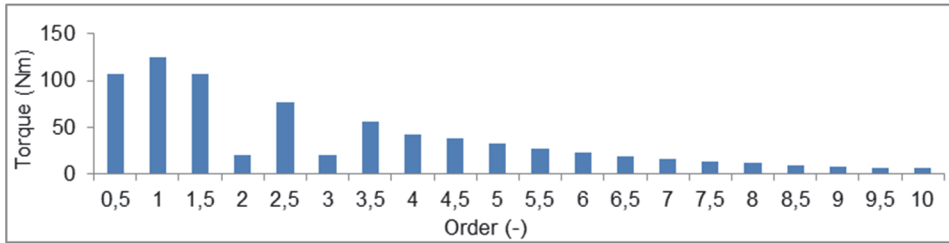

Figure 8. Spectrum of resultant rotational excitation harmonics at crankshaft speed of 6500 rpm

From the fifth order on, the effect of the rotational mass forces is practically negligible. The amplitudes of the rotational resultant forces decrease also significantly in function of the orders, and as a result the analysis does not have to be extended for higher orders (in this paper the last order is the tenth).
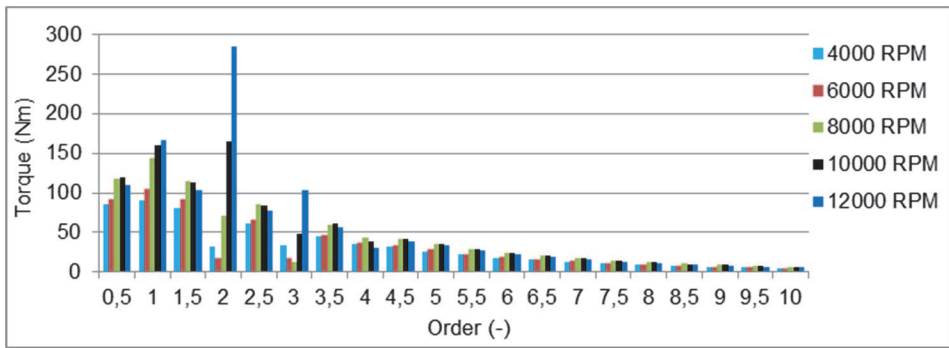
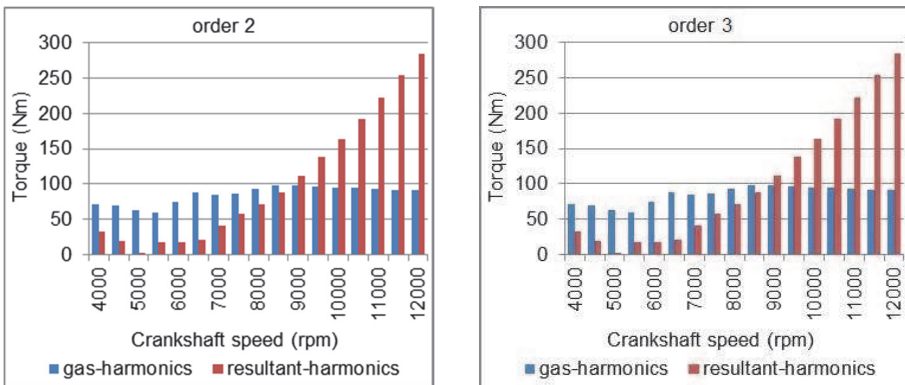*Figure 9. Resultant excitation harmonics at different crankshaft speeds*

Figure 10. The second- and third order rotational gas- and resultant harmonics in function of crankshaft speed

Studying the excitation harmonics in function of the crankshaft speed, it can be observed that – in case of any value of the crankshaft speed – the second and the third order rotational mass forces (Fig. 8.) have a major effect on the amplitude of the rotational gas force (Fig. 9.). Furthermore – as can be seen from Fig. 10. – in case of lower engine speeds, this effect means the decrease of the amplitude of the rotational gas force, while in case of higher engine speeds this effect means the increase of the same.

6. Defining critical speeds

Unlike the natural frequencies, the frequencies of the excitation harmonics (v_e) are not constant, but they are increasing linearly in function of the crankshaft speed:

$$v_e = x \frac{n}{60} , \tag{12}$$

where *n* has to be substituted in rpm.

Plotting the excitation and natural frequencies in function of crankshaft speed, the excitation frequencies will be lines starting from the origin, while the natural frequencies will be horizontal lines. The critical speeds can be derived from the intersections of the excitation and natural frequency curves [11], [12].

According to Fig. 11. the first six excitation harmonics have no intersections even with the first natural frequency. Since the first six harmonics have the highest amplitudes – as can be seen from Fig. 9. – they are the most dangerous ones so avoiding them has an advantageous effect on the dynamic loading of the crankshaft. The first excitation harmonic that resonates with the first natural frequency is the 3,5th excitation harmonic. From Fig. 8. it can be seen that from this excitation harmonic on, the amplitudes of the harmonics decrease, so they generate smaller and smaller angular deflection. The examined first 20 excitation harmonics have 19 intersection points with the first and second natural frequencies, so that means that there are 19 critical speed values. Not every critical speed causes dangerous amplitudes, however.
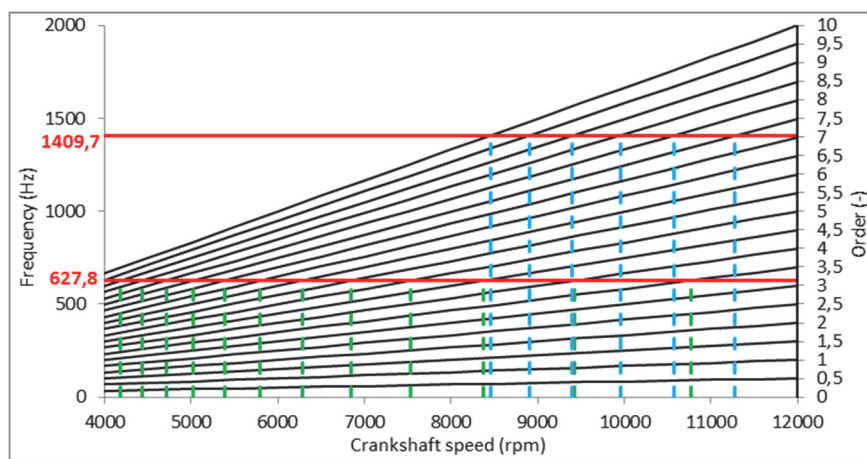

Figure 11. Defining critical engine speeds

On the one hand, the magnitude of the angular deflections depend on the amplitude of the excitation harmonic, on the other hand, the different damping effects in the system have a major decreasing effect on the magnitudes of the angular deflection. The next chapter deals with these magnitudes.

7. Damping of the crank mechanism

Since the system has a non-negligible damping, finite revolutions will occur at the resonance locations. Basically, there are two damping effects acting on the crank mechanism:

- damping of the piston.
- damping of the valve control and different auxiliary equipment (like the oil pump or the water pump).

The damping effect of the piston - characterized by the damping coefficient k_d - acting on the crankpin (mass element J_4) can be calculated with the following formula according to [3]:

$$k_d \left[\frac{Nms}{rad} \right] = 1,5 \cdot 10^{-7} r^2[m] \frac{d_d^{\,2}[m]\pi}{4} \,, \tag{13}$$

where r is the throw of crank and d_d is the diameter of the piston. The damping from the effects of the valve control and auxiliary equipment was defined on the flywheel (mass element 1). The value of this can only be obtained through experiments. Without such experience one can only rely on the literature. Based on experiments measured by AVL [12] on similar one-cylinder engines, the value of this damping was considered as $1,2 \; Nms / rad$

8. Solving the equations of motion in case of forced vibrations

Taking into consideration excitation and damping, the equation system of motion looks like the following formula [9]:

$$\underline{\underline{M}} \, \underline{\ddot{\varphi}} + \underline{\underline{C}} \, \underline{\dot{\varphi}} + \underline{\underline{K}} \, \underline{\varphi} = \underline{M}_g(t) \,, \tag{14}$$

where: $\underline{\underline{M}}$ is the mass matrix, $\underline{\underline{C}}$ is the damping matrix, $\underline{\underline{K}}$ is the stiffness matrix, $\underline{M}_g(t)$ is the vector of the excitation forces / torques and $\underline{\varphi}(t)$ is the vector of angles of revolution (unknown parameters).

The solution of this second order, ordinary, inhomogeneous differential equation system provides the forced angular deflection, the forced angular velocity and the forced angular acceleration of each mass element. Fig. 12. illustrates the angular deflection between the two ends of the crankshaft, between the flywheel (mass element 1) and the clutch (mass element 7). The angular deflection between the two ends of the crankshaft (15) is characteristic for the torsional stiffness of the crankshaft.
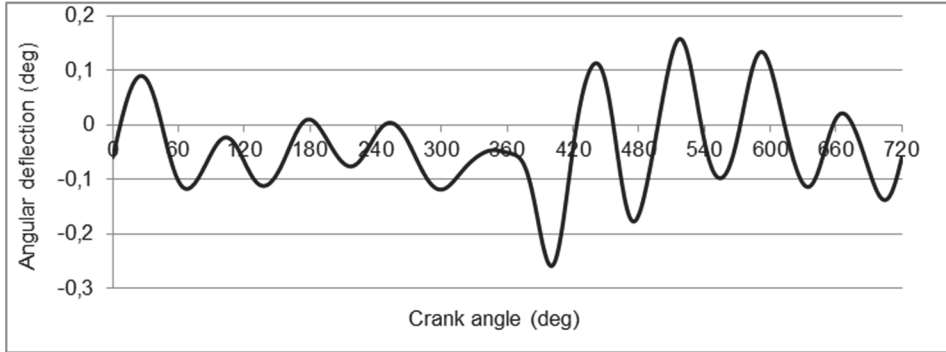
$$\varphi_{17} = \varphi_7 - \varphi_1 \,. \tag{15}$$

Figure 12. The angular deflection between the flywheel and the clutch in function of crank angle at 8000 rpm engine speed

In order to be able to plot the relative angular deflection of the crankshaft in function of the crankshaft speed, the mean relative angular deflection is calculated as follows:

$$\bar{\varphi}_{17} = \frac{|\varphi_{17\,max}| + |\varphi_{17\,min}|}{2} . \tag{16}$$
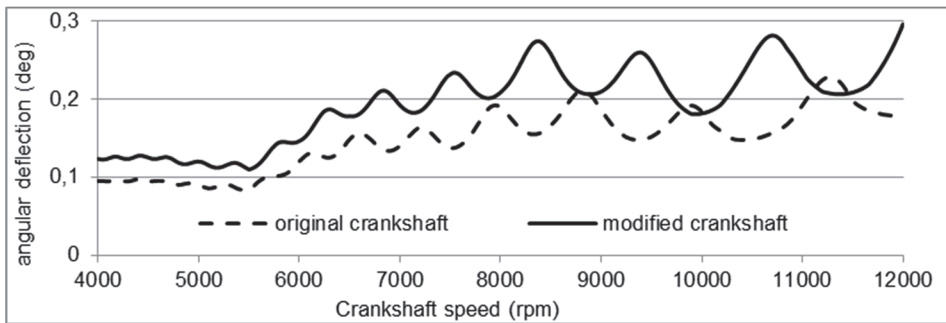

Figure 13. Mean relative angular deflection of the two ends of the original crankshaft and that of the two ends of the modified crankshaft in function of engine speed

Fig. 13. shows the mean relative angular deflection of the crankshaft (16) in function of the crankshaft speed, while Fig. 14. shows the same quantity, but caused by the separate orders of the excitation forces. In order to be able to compare the results, the mean relative angular deflection of the original crankshaft [1] is also plotted in Fig. 13. According to Fig. 13. the maximal mean relative angular deflection of the modified crankshaft is approximately 30% higher compared to the original crankshaft. According to AVL design references [10], this value should be below 0,6°. As can be seen from Fig. 13. this criteria is fulfilled by both crankshafts.

As to the mean relative angular deflections caused by the separate orders of the excitation loads, AVL indicates a maximal value of 0,1° in case of serial engines [10]. Above this limit value the engine noise increases. According to Fig. 14. this criteria is not fulfilled in case of orders below five, but since the examined crankshaft is from a race engine the engine noise is not an important issue.
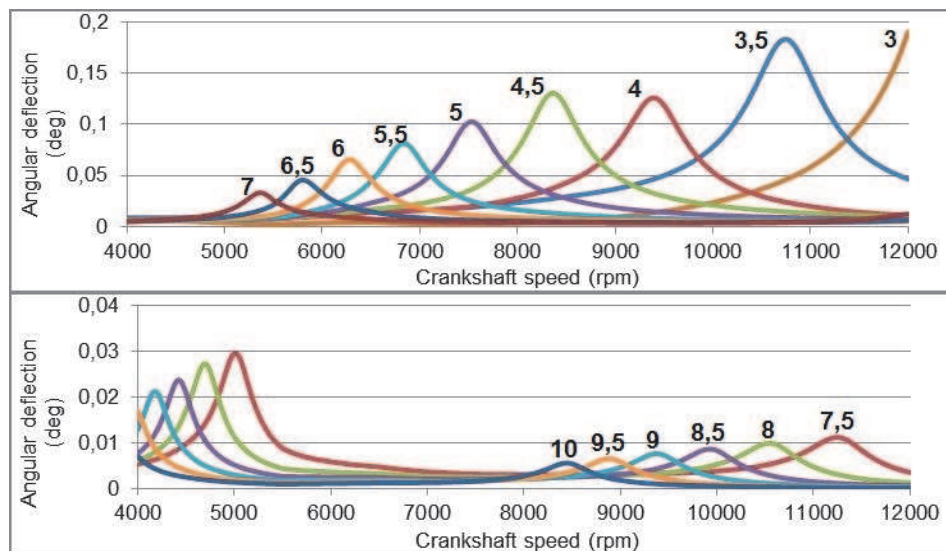
Figure 14. Mean relative angular deflection of the two ends of the modified crankshaft caused by the separate orders of the excitation force in function of engine speed

From Fig. 14. it can be observed, that the maximum values of angular deflection of the separate orders (the resonance locations) occur at critical speeds determined from Fig. 9. According to Fig. 9. in case of orders 3,5 – 7, the curve has one, while in case of orders 7,5 – 10, the curve has two maximum values. Furthermore, as excpected, the highest angular deflection occurs in case of order 3,5, while the maximum deflections decrease as the orders increase.

Besides the relative angular deflection, another characteristic quantity that is usually analysed is the speed irregularity of the clutch. In order to be able to calculate this, first the angular velocity of the clutch must be defined. This can be achieved by differentiating the angular deflection of the clutch (Fig. 15.).
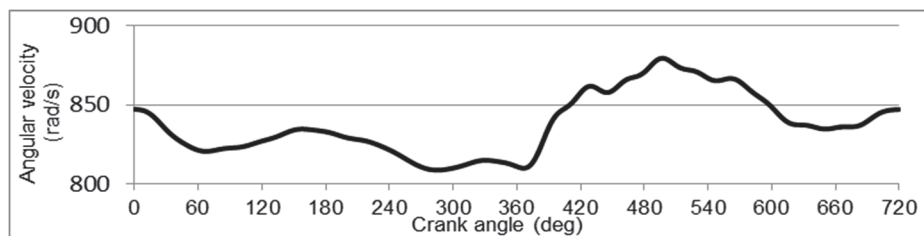

Figure 15. Angular velocity of the clutch in function of crank-angle at 8000 rpm engine speed

The angular speed irregularity δ of the clutch is the difference between the maximum and the minimum angular speeds divided by the mean angular speed ω_k:

$$\delta = \frac{\omega_{max} - \omega_{min}}{\omega_k}, \quad \omega_k = \frac{\omega_{max} + \omega_{min}}{2}. \tag{17}$$

The speed irregularity of the clutch of the original, as well as of the modified crankshaft [1] can be seen in Fig. 16.
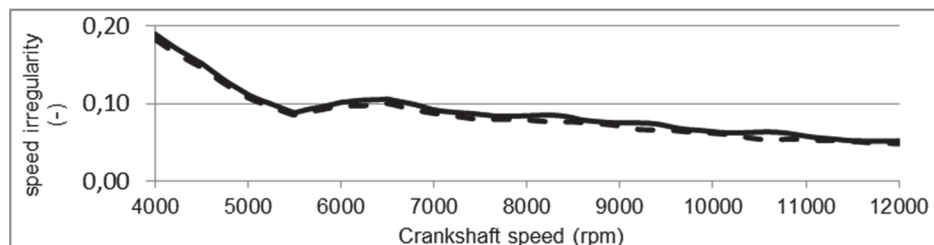

Figure 16. The speed irregularity of the clutch of the original crankshaft (dashed line) and that of the modified crankshaft (continuous line) in function of engine speed

A too high value of speed irregularity has a negative effect on the fatigue life of the gearbox. In case of serial engines, the maximum acceptable value is approximately 0,16 –0,2. According to Fig. 16. the maximum value of the speed irregularity is approximately 0,18 in case of the original crankshaft, while in case of the modified crankshaft this value is slightly higher, approximately 0,19 at 4000 rpm engine speed. But this value rapidly decreases as the engine speed increases, at 5500 rpm engine speed it is already below 0,1 and it remains below this value in the largest part of the operational range.

9. Summary of the results

For the torsional dynamic analysis, we have created a sequential torsional multi-mass vibration model of the crankshaft mechanism. In order to be able to check the vibration characteristics of the torsional model, the first two natural frequencies and natural modes obtained from the multi-mass model were compared with those obtained from the 3D finite element model of the crankshaft. The error was below 2 % at the lowest two natural frequencies of the two different models. The natural modes have also shown a proper match and therefore the forced vibration analysis was performed only on the multi-mass model.

During the forced vibration analysis, the following statements could be made:

- Up to the first ten excitation harmonics, two natural frequencies appear in the operational range of the engine, which result nineteen resonance locations altogether.

- The first six excitation harmonics have no intersection with the natural frequencies. Since these harmonics have the highest amplitudes they are the most dangerous ones, avoiding them has an advantageous effect on the dynamic loading of the crankshaft.

- Also taking into account the dampings of the system, the mean relative angular deflection of the two ends of the modified crankshaft has increased by approximately 30 % compared to that of the two ends of the original crankshaft, so the result was a maximum value of 0,3°, which is still half of the limit-value of 0,6° specified by AVL [10].

- The speed irregularity of the clutch has slightly increased compared to the original crankshaft, the maximum value of 0,18 occuring in case of the original crankshaft has increased to approximately 0,19 at an engine speed of 4000 rpm. However this value rapidly decreases as the engine speed increases. At 5500 rpm engine speed this is already below 0,1 and it remains below this value in the largest part of the operational range.

References

[1] Horváth P, Égert J: Stress analysis and weight reduction of a one-cylinder engine crankshaft. Acta Technica Jaurinensis, Vol. 8, No. 3, pp. 201-217, 2015. DOI: 10.14513/actatechjaur.v8.n3.377

[2] AVL Excite Designer Primer (User's Manual). AVL List Gmbh, Graz, 2009.

[3] AVL Excite Designer Theory. AVL List Gmbh, Graz, 2009.

[4] Parikyan T, Resch T, Priebsch HH: Structured model of crankshaft in the simulation of engine dynamics with AVL/Excite. AVL List Gmbh, Graz, 2001.

[5] ANSYS 14.5 Mechanical APDL Theory Reference. Ansys Inc., 2012.

[6] Mitianiec W, Buczek K: Torsional vibration analysis of a crankshaft in heavy duty six cylinder inline engine. Politechnika Krakowska Publisher, 2008.

[7] Meirelles PS, Zampieri DE, Mendes AS: Mathematical Model for Torsional Vibration Analysis in Internal Combustion Engines. 12th IFToMM World Congress, Besançon (France), June 18-21, 2007.

[8] Reynolds DD: Engineering principles of mechanical vibration (3rd edition). Trafford publishing, 2013.

[9] Osgood B: The Fourier Transform and its Applications. Stanford University, 2009.

[10] AVL Design References. AVL List Gmbh, Graz, Oral information.

[11] Corbo MA, Malanoski SB: Practical design against torsional vibration. In Proceedings of the Twenty-Fifth Turbomachinery Symposium, September, pp. 189-222, 1996.

[12] Huang Y, Yang S, Zhang F, Zhao C, Ling Q, Wang H: Nonlinear Torsional Vibration Characteristics of an Internal Combustion Engine Crankshaft Assembly. Chinese Journal of Mechanical Engineering, Vol. 25, Issue 4, pp. 797-808, July 2012.

FEM Modeling and Weight Reduction of a Solar Energy Driven Racing Car Chassis

L. Fehér, J. Égert

Department of Applied Mechanics, Faculty of Mechanical Engineering
Information Technology and Electrical Engineering
Széchenyi István University
9026 Győr 1 Egyetem tér, Hungary
feherlb@sze.hu, egert@sze.hu

Abstract: The paper demonstrates the strength analysis of a solar energy driven racing car chassis with two different mechanical modeling, beam-bar and shell modeling. Weight reduction is also a key point in the stages of car design. Reducing the weight of the car results in higher fuel efficiency and usually better dynamical properties. The weight reduction was performed with three different materials, steel, aluminium-alloy and laminated fibre reinforced composites. The paper shows the complete mechanical modeling process from construction of the mechanical model to the analysis of the laminated composite chassis structure and the results of the weight reduction.

Keywords: car chassis, steel, aluminium-alloy, fibre reinforced composite, strength analysis, FEM, weight reduction

1. Introduction

The investigated car was designed by the SZEnergy team (SZEnergy is a students' team for solar car development at Széchenyi István University) for the Shell Eco Marathon race series. The aim of the Shell Eco Marathon competition is to achieve the lowest fuel consumption. Fuel efficiency is one of the most important points in car design not only at this competition but also in general in car manufacturing. There are many international regulations in the vehicle industry to limit the emission of carbon dioxide. A higher efficiency of fuel consumption is not only advantageous for the drivers' budget but also beneficial for the environment. This parameter depends on a lot of things (car weight, air-drag, type of combustion engine, etc.). The paper deals only with one aspect, the weight reduction.

There is a famous saying by Colin Chapman who was a car designer and builder. He said "Adding power makes you faster on the straights. Subtracting weight makes you faster everywhere." [14] This principle also works in the Eco Marathon race series but in this case it sounds like this: Subtracting weight makes you save fuel everywhere.

The main goal of the analysis was to determine the stress distribution in the chassis and therefore to reduce the weight of the racing car [1]. The investigation concentrated on the chassis because it is one of main parts of the car and it represents a considerable part of its mass. Three different materials and two different mechanical models were used to analyse the structure and to realize the weight reduction. The paper also examines in detail the determination of the critical load cases and it describes the results of stress analysis and weight reduction.

2. Geometry and load cases

2.1. Layout of the solar racing car

The investigated solar car was designed for the Urban Concept category in the Shell Eco Marathon race. The racing car has a traditional car layout, a car body and a chassis separately [2]. The solar cells were fixed on the car body.

The geometry of the chassis under the car body was available from the SZEnergy students' design team (Fig. 1.). The chassis carries all the loads and holds all other parts together. It is a welded beam-bar structure with closed thin-walled square and circle profiles (Fig. 1.). The weight of the original chassis geometry (Fig. 1.) was 55.6 kg for steel, 19.2 kg for aluminium alloy and 7.17 kg for laminated fibre reinforced composite material.
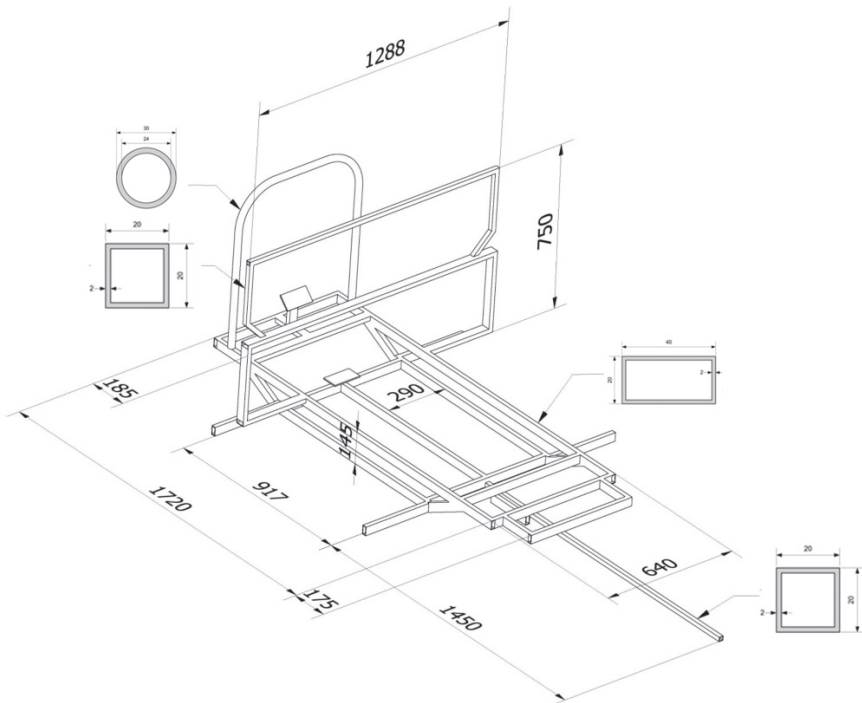

Figure 1. The geometry of the investigated chassis

2.2. Basic and critical load cases

For the stress analysis two different mechanical models were built and during the mechanical modelling only the chassis was considered. First the critical loads were determined which are acting on the structure during different driving manoeuvres. Two groups of loads were defined. The basic load cases (BL) is the first group which consist of elementary load parts and the other is the critical load cases (CL) which give the starting point for the dimensioning of the car chassis. The mechanical problem is linear therefore the critical loads can be built up as a superposition of basic loads.

The basic load cases group consist of seven basic load sets. These are the following:

BL1 dead weight of the chassis and other parts,
BL2 body forces from longitudinal acceleration at emergency braking,
BL3 body forces from lateral acceleration at turning with full speed,
BL4 moments from the front wheels at braking,
BL5 forces from the driver at braking,
BL6 loads from the acceleration of non-braked rear wheels,
BL7 moments from the wheels at turning.

From the loading point of view three critical situations can occur during the race. These critical situations can be characterised by the superposition of basic load sets:

CL1 emergency braking: *BL1 + BL2 + BL4 + BL5 + BL6*,
CL2 turning with full speed: *BL1 + BL3 + BL7*,
CL3 emergency braking at turning with full speed:
BL1 + BL2 + BL3 + BL4 + BL5 + BL6 + BL7.

BL1 - Loads from the dead weight of the chassis and other parts

This load case originates from the weight of the car with every coupled part and from the body of the driver. It is determined by the acceleration of gravity g and the mass density ρ of each component. The mass density values are for steel $\rho_s = 7\,820 \text{ kg/m}^3$, for aluminium alloy $\rho_a = 2\,700 \text{ kg/m}^3$ and for the composite material $\rho_c = 1200 \text{ kg/m}^3$. This body force loading occurs in every critical load cases.

BL2 - Body forces from longitudinal acceleration at emergency braking

The maximum acceleration a_{Sm} (negative value) appears during the race at emergency braking. Only the front wheels of the car are supposed to be braked so the a_{Sm} maximum deceleration is determined by the slip point of the front wheels. This a_{Sm} maximum deceleration is caused by the tangential coordinate of $\vec{F}_E$ reaction force (Fig. 2.). Newton's second law and the principle of conservation of angular momentum are used to determine this a_{Sm} maximum deceleration:

$$F_{En} = \frac{b}{\left(a + b - c\mu_0\right)} m_c\, g \tag{1}$$

where F_{En} is normal reaction force on the front wheels, a, b, c are geometric data of the car, m_c is the mass of the car (Fig. 2.), μ_0 is the coefficient of static friction and g is the acceleration of gravity. Using equation (1) and Newton's second law the a_{Sm} maximum deceleration is

$$a_{Sm} = \frac{F_{Et}}{m_c} = -\mu_0 \frac{F_{En}}{m_c} = -\frac{\mu_0 b}{\left(a + b - c\mu_0 \right)} g \ . \tag{2}$$

In formulas (1) and (2) a static friction $\mu_0 = 0.9$ was considered as the upper limit of the real kinetic friction. The maximum deceleration values are for steel chassis $a_{Sm} = -3.75 \ \text{m/s}^2$, for aluminium chassis $a_{Sm} = -3.53 \ \text{m/s}^2$ and for composite chassis $a_{Sm} = -3.47 \ \text{m/s}^2$.
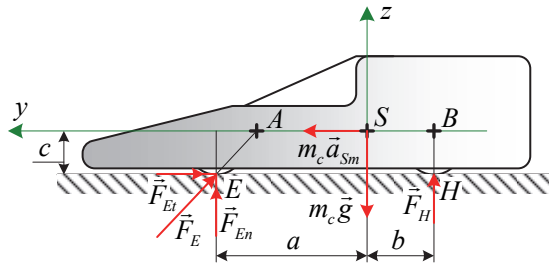

Figure 2. Geometric data and reaction forces in the slide point of the front wheels at braking

BL3 - Body forces from lateral acceleration at turning with full speed

At turning with full speed considerable body forces work on the car and all other coupled parts. They come from the lateral acceleration at turning with full speed. In this manoeuvre two possible critical events can occur. These are the turning-over and the sliding of the car. These two critical events cannot occur at the same time. The lateral accelerations were determined for both situations and two different values were obtained. From these the lower critical lateral acceleration value causes the real situation, which occurs first.
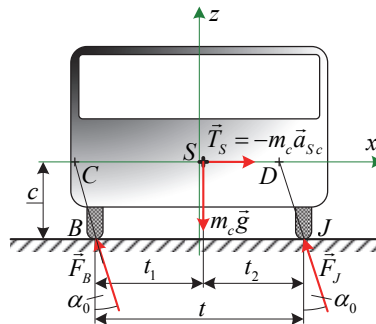

Figure 3. Geometric data and reaction forces in the slide point of the wheels at turning

Only the sliding event is shown because it happens first. Newton's second law was used to determine the a_{Sc} maximum side-slip lateral acceleration. It was assumed that both wheels slide at the same time (Fig. 3.)

$$F_{Bn} + F_{Jn} = m_c g \tag{3}$$

$$m_c a_{Sc} = F_{Bt} + F_{Jt} = -\mu_0 (F_{Bn} + F_{Jn}) = -\mu_0 m_c g \tag{4}$$

where a_{Sc} is the maximum lateral acceleration, F_{Bt}, F_{Jt} are the tangential reaction forces in lateral direction on the left and right side, F_{Bn}, F_{Jn} are the normal reaction forces. In formulas (3) and (4) a static friction $\mu_0 = 0.9$ was also considered as the upper limit of the real kinetic friction.

From equations (3), (4) the maximum lateral acceleration is

$$a_{Sc} = -\mu_0 g = -8.83 \text{ m/s}^2 . \tag{5}$$

BL4 - Moments from the front wheels at braking

At braking the tangential coordinate of the $\vec{F}_E$ reaction force (Fig. 2.) causes a torque on the front axle. This torque can be determined by the following formula:

$$M_{ct} = \frac{d_w}{2} F_{Et} = -\frac{d_w}{2} \mu_0 F_{En} - \frac{\mu_0 \, b \, d_w}{2(a + b - c \mu_0)} m_c g , \tag{6}$$

where $d_w = 558$ mm is the outside diameter of wheel, a, b, c are the dimensions of the centre of gravity S of the car, $\mu_0 = 0.9$ is coefficient of static friction, g is the gravitational acceleration. The torque on the front axle are $M_{ct} = 141$ Nm for steel chassis, $M_{ct} = 115$ Nm for aluminium alloy chassis and $M_{ct} = 110$ Nm for composite chassis.

BL5 - Forces from the driver at braking

At braking the driver pushes the brake pedal and this activity causes a reaction force in the driver's seat as well. This load was modelled with two concentrated forces. One force acts on the driver's seat console and the other acts on the mounting point of the brake pedal. This force was determined by measurement and its magnitude is 500 N.

BL6 - Loads from the acceleration of non-braked rear wheels

At emergency braking tangential forces originate from the rear wheels. The total tangential coordinate of $\vec{F}_a$ force (Fig. 4.) can be determined from the principle of conservation of angular momentum:

$$F_{at}^{**} = \frac{3}{2} m_w a_{Sm} \tag{7}$$

where m_w is the mass of the wheel, a_{Sm} is the maximum deceleration.
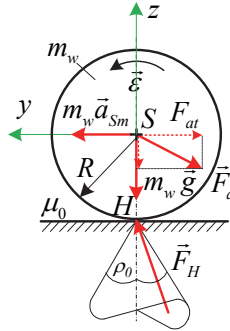

Figure 4. Forces on the rear non-braked wheel at emergency braking

The translational motion part of tangential coordinate F_{at}^* of $\vec{F}_a$ force was considered in *BL2* load case. Thus the load from non-braked rear wheels deceleration at emergency braking is

$$F_{at} = F_{at}^{**} - F_{at}^* = \frac{3}{2} m_w a_{Sm} - m_w a_{Sm} = \frac{1}{2} m_w a_{Sm} , \qquad (8)$$

where F_{at}^{**} is the total part of tangential coordinate and F_{at}^* is the translational motion part of tangential coordinate of the $\vec{F}_a$ force. In the analysis only aluminium alloy wheel-discs were considered with the mass of $m_w = 8.41 \, \text{kg}$. Therefore the additional force is $F_{at} = -15.77 \, \text{N}$ for steel chassis, $F_{at} = 14.84 \, \text{N}$ for aluminium alloy chassis and $F_{at} = 14.59 \, \text{N}$ for composite chassis.

BL7 - Moments come from turning loads

Bending moments in the front and rear axles come from the lateral forces when the car is turning. These bending moments were determined from the critical lateral forces F_{Bt} and F_{Bt} (Fig. 3.).

The four wheels were loaded with different bending moments, because the force values are different on the four wheels. These bending moments could be calculated with the following formulas:

$$M_{LFy} = \mu_0 F_{LFn} \frac{d_w}{2} = \mu_0 \frac{d_w}{2} \frac{b}{a+b} \frac{t_2 - \mu_0 c}{t} m_c g , \qquad (9)$$

$$M_{LSy} = \mu_0 F_{LSn} \frac{d_w}{2} = \mu_0 \frac{d_w}{2} \frac{a}{a+b} \frac{t_2 - \mu_0 c}{t} m_c g , \qquad (10)$$

$$M_{RFy} = \mu_0 F_{RFn} \frac{d_w}{2} = \mu_0 \frac{d_w}{2} \frac{b}{a+b} \frac{\mu_0 c + t_1}{t} m_c g , \qquad (11)$$

$$M_{RSy} = \mu_0 F_{RSn} \frac{d_w}{2} = \mu_0 \frac{d_w}{2} \frac{a}{a+b} \frac{\mu_0 c + t_1}{t} m_c g , \tag{12}$$

where L index means the left side of the car, R index means the right side of the car, F index is for the front wheels of the car, S index is for the rear wheel of the car, $d_w = 558$ mm is the outside diameter of the tyre, a, b are the coordinates of the centre of gravity of the car, $\mu_0 = 0.9$ is coefficient of static friction, g is the gravitational acceleration and m_c is the mass of the car.

3. FEM modelling of the chassis

The finite element method (FEM) was used for the stress analysis. For the analysis two types of mechanical models of chassis were applied: a simple beam-bar model and a more complex shell model [12].

The beam-bar model consists of beam-bar elements modelled by their central lines and additional concentrated mass elements that are modelling the effect of the other parts of the car. From these additional mass elements the chassis obtains further loading, weight $\vec{G} = m\vec{g}$ and body forces $\vec{T} = -m\vec{a}$.
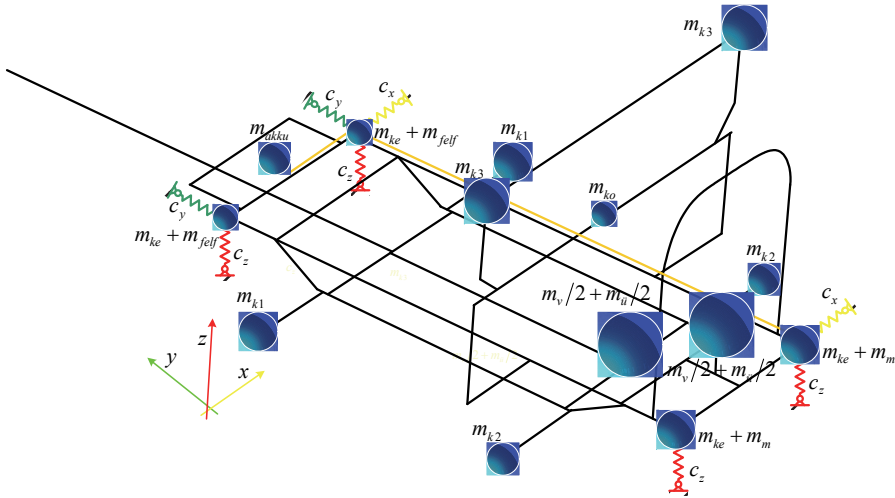

Figure 5. The FEM beam-bar model with additional masses and boundary conditions

At the wheels the chassis is supported by spring elements that are modelling the elastic support of the wheels. The spring coefficients c_x, c_y and c_z are different in reality. The coefficient c_z could only be determined by measurement (Fig. 6.) and the same value $c_z = 6 \cdot 10^{-3}$ mm/N was applied for the tangential coefficients c_x and c_y.
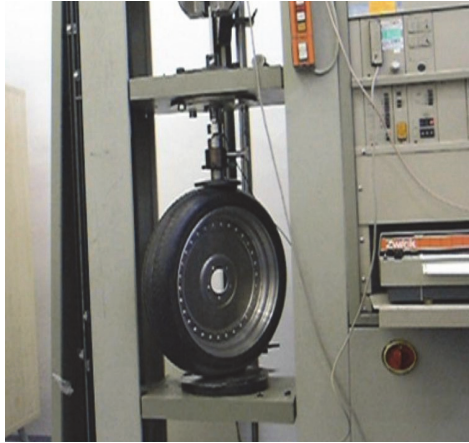
Figure 6. Measurement of the spring coefficient c_z

The beam-bar model can only be applied to analyse of the chassis made of steel and aluminium alloy because the traditional beam and bar elements cannot model thin walled laminated beam-bar structures.

The chassis made of laminated fibre reinforced composite material can be modelled by laminated orthotropic shell elements. As mentioned before the chassis is built up of thin walled closed profile beams. For the shell model the beams of the chassis are modelled with their outer surfaces. The wall thickness and the number of layers were design parameters. The shell FEM model and the additional mass elements are shown in Fig. 7. The additional mass elements and the boundary condition were the same by the beam-bar FEM model.
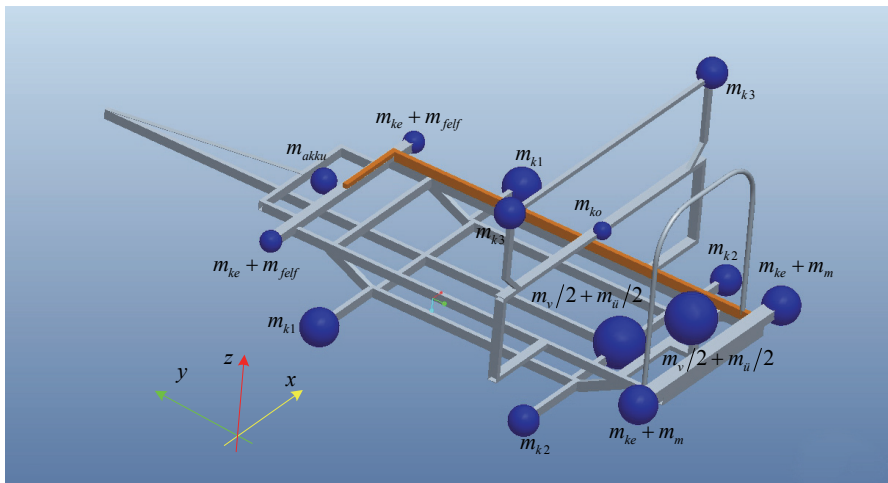

Figure 7. The shell FEM model with additional mass elements

In the laminated shell model two weave types (plain weave and twill weave) of reinforcement were used for the two different laminated composite structures. The material properties of these two composites are shown in Table 2.

By composites it is very important to define the optimal fibre orientations and layer structures. In our models one composite layer's thickness was 0.25 mm. The shell models were built from eight layers e.g. 1.5 mm wall thickness and the warp fibre orientations were parallel to the beam-axes.

The *von Mises* equivalent stress theory was used to evaluate the strength analysis and to dimension for isotropic materials. The *von Mises* equivalent stress and the safety criterion are defined by (13)

$$\sigma_{eq}(von\ Mises) = \sqrt{\frac{1}{2}\left[\left(\sigma_I - \sigma_{II}\right)^2 + \left(\sigma_{II} - \sigma_{III}\right)^2 + \left(\sigma_I - \sigma_{III}\right)^2\right]} \leq R_{p0,2}, \qquad (13)$$

where $\sigma_I, \sigma_{II}, \sigma_{III}$ are principal stresses and $R_{p0,2}$ is the yield stress of material.

Composites are not isotropic materials so they need different failure criteria then metals. In this analysis the *Tsai-Wu* failure criterion was applied (14). If the left side value is bigger than the one in the criterion, it means the starting point of the failure of the composite structure [3] [4]

$$\left(\frac{\sigma_1^2}{X_T X_C}\right) + \left(\frac{\sigma_2^2}{Y_T Y_C}\right) + \left(\frac{1}{X_T} - \frac{1}{X_C}\right)\sigma_1 + \left(\frac{1}{Y_T} - \frac{1}{Y_C}\right)\sigma_2 + \left(\frac{\tau_{12}}{S}\right)^2 +$$

$$+ \left(\frac{\tau_{23}}{S}\right)^2 + \left(\frac{\tau_{13}}{S}\right)^2 - \sqrt{\frac{1}{X_T X_C}\frac{1}{Y_T Y_C}}\sigma_1 \sigma_2 < 1 \qquad (14)$$

where σ identify the normal stress coordinates and τ identify the shear stress coordinates in the fibre direction coupled natural coordinate system x_1, x_2, x_3. Directions x_1, x_2 lie in the plane of layer and direction x_1 shows the orientation of the fibres. In the (13) *Tsai-Wu* criterion $E_1 = E_2$ and E_3 are *Young* moduli in direction *1, 2* and *3*, $v_{12} = v_{21}$ and $v_{13} = v_{23}$ *Poisson*'s coefficients, G_{12} and $G_{23} = G_{13}$ are independent shear moduli, $X_T = Y_T$ is the tension strength in direction *1* and *2*, $X_C = Y_C$ is the compression strength in direction *1* and *2*, S is the shear strength and ρ is the mass density.

4. Results of FEM stress analysis

4.1. Results of beam-bar modeling

Laminated composite structures cannot be analysed with the usual beam-bar modelling so the following beam-bar computations were made only for these two homogeneous metal materials, i.e. steel and aluminium alloy.

The beam-bar FEM analysis was performed with the material properties in Table 1. In the table E is the *Young* modulus, ν is the *Poisson*'s coefficient, G_{12} is the shear modulus, $R_{p0,2}$ is the yield stress and ρ is the mass density of the material.

Table 1. Material properties of metal materials

		Steel	Aluminium alloy
E	$[\text{MPa}]$	206800	72000
ν	$[-]$	0.29	0.2963
G	$[\text{MPa}]$	80155	27771.3
$R_{p0,2}$	$[\text{MPa}]$	250	280
ρ	$[\text{t/m}^3]$	$7.82 \cdot 10^{-9}$	$2.7 \cdot 10^{-9}$

The most important results of FEM strength analysis of beam modelling provide maximum *von Mises* equivalent stress distributions along the beams which can be used for the comparison of the three critical load cases and for the dimensioning of the chassis. The maximum equivalent *von Mises* stress distribution for the three critical load cases are shown for the aluminium alloy chassis in Fig. 7.-9. The character of maximum stress distribution is the same for the steel chassis as well, only the stress values are different.
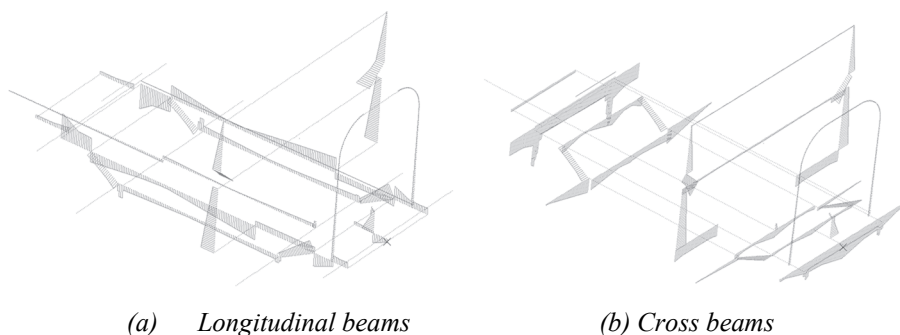

(a)   Longitudinal beams   *(b) Cross beams*
Figure 7. Maximum von Mises stress distribution for the CL1 critical loading case (aluminium chassis)

The comparison of the three critical load cases CL1-CL3 showed that the CL3 is the most dangerous (Fig. 9.). In this complex critical load case the maximum equivalent *von Mises* stress is on the chassis 116 MPa for aluminium alloy and 135.78 MPa for steel. These maximum stress values occur in the same place. This is the front axle and the chassis connection area.
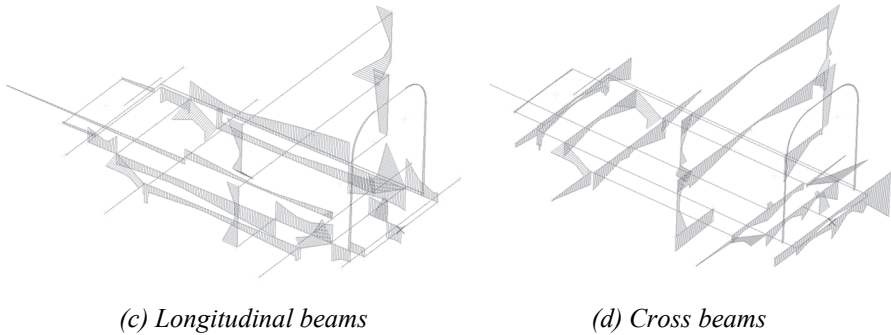
(c) Longitudinal beams   *(d) Cross beams*

Figure 8. Maximum von Mises stress distribution for the CL2 critical loading case (aluminium chassis)

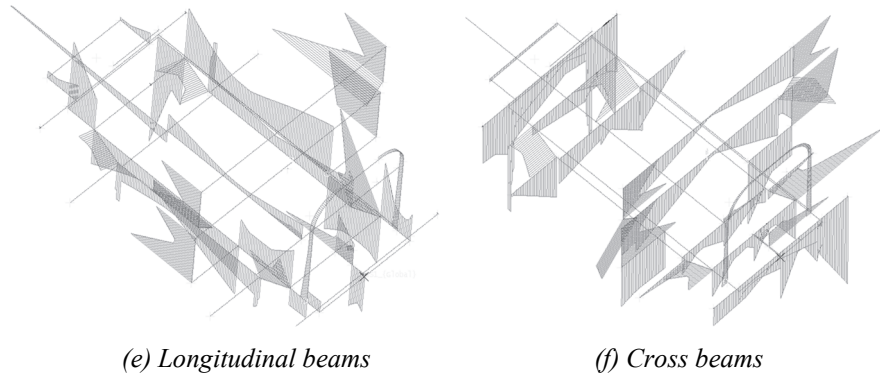*(e) Longitudinal beams*   *(f) Cross beams*

Figure 9. Maximum von Mises stress distribution for the CL3 critical loading case (aluminium chassis)

From one part of the chassis the loads can pass over to other areas and this can cause local maximum values. The connection area is a small area and this causes these critical high values. But if the safety factor for case *CL3* is calculated with both metal materials these show that the chassis fulfils the safety criterion (13) for both materials.

Using the (15) expression the *n* safety factor is 1.84 for steel material and 2.41 for aluminium alloy

$$n = \frac{R_{p0,2}}{\sigma_{eq\,max}}$$

(15)

where $R_{p0,2}$ is the yield strength and $\sigma_{eq\,max}$ is the maximum equivalent stress value.

The safety factor values show that there is emergency reserve in both structures of metal materials. This fact enables the weight to be reduced, for example reducing the wall thickness of beams.

Also an important task is to find the critical cross sections and critical points of the chassis. From the results of beam-bar modelling, shown in Fig. 7-9., the critical points and the critical cross sections of the chassis have been obtained. The results show good correlation between steel and aluminium-alloy so the places of critical points are the same for both metal structures. These critical points are assigned with numbers and they are shown in Fig. 10.
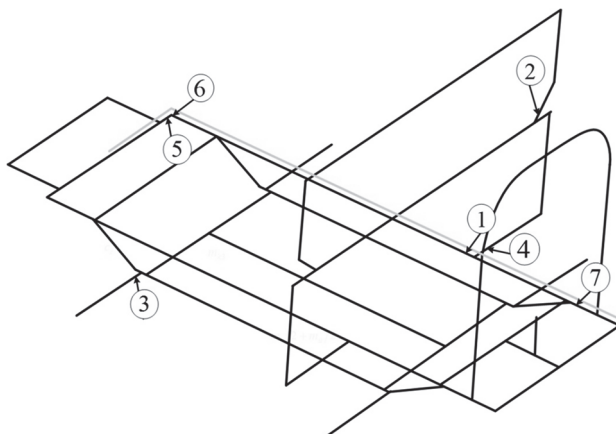
Figure 10. Critical points of the aluminium alloy chassis

4.2. Results of laminated shell modeling

The material and strength properties of applied composite materials are summarized in Table 2.

Table 2. Material properties of the applied composite material

	Twill weave	Plain weave
$E_1 = E_2$ [MPa]	35734	45514
E_3 [MPa]	3500	3500
$\nu_{12} = \nu_{21}$ [-]	0.07	0.06
$\nu_{13} = \nu_{23}$ [-]	0.35	0.35
G_{12} [MPa]	2115	2722
$G_{23} = G_{13}$ [MPa]	1296.3	1296.3
$X_T = Y_T$ [MPa]	321	562
$X_C = Y_C$ [MPa]	521	580
S [MPa]	23	26,3
ρ [t/mm^3]	$1.19 \cdot 10^{-9}$	$1.23 \cdot 10^{-9}$

In Table 2. $E_1 = E_2$ and E_3 are *Young* moduli in direction *1, 2* and *3*, $v_{12} = v_{21}$ and $v_{13} = v_{23}$ Poisson's coefficients, G_{12} and $G_{23} = G_{13}$ are independent shear moduli, $X_T = Y_T$ is the tension strength in direction *1* and *2*, $X_C = Y_C$ is the compression strength in direction *1* and *2*, S is the shear strength and ρ is the mass density.

The basic shell model which includes two different weave types and 8 plies, was analysed first. Every beam of the chassis was identified by a number for easier evaluation. In the chassis 38 beams were identified (Fig. 11-12.).
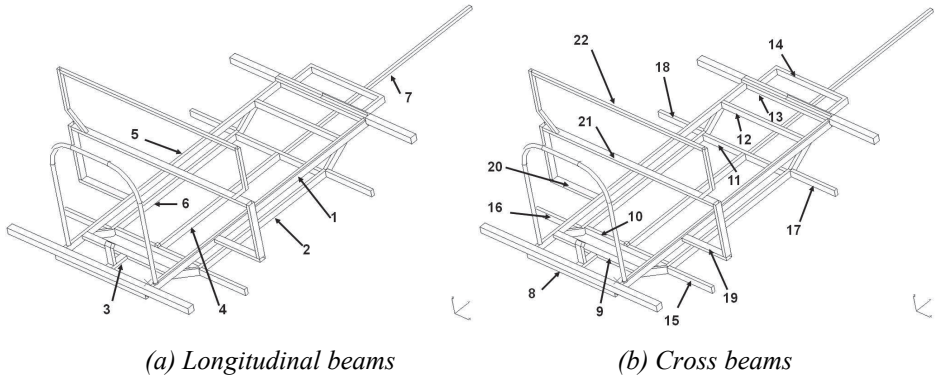

| (a) Longitudinal beams | (b) Cross beams |

Figure 11. Identification or numbering of the composite chassis beams
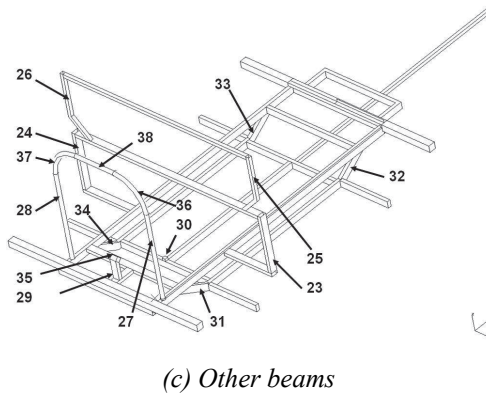

(c) Other beams

Figure 12. Identification or numbering of the composite chassis beams

For evaluating the computational results the maximum *Tsai-Wu* failure criterion values were used. The FEM software determines the failure value in the three parts of the composite laminate (top, middle, bottom). From these results the maximum *Tsai-Wu* failure indices were obtained.

In Table 3. the results of the model made from plain weave plies were summarized. This weave type has shown better strength properties from the strength analysis so the laminated composite structure of final model is made from plain weave carbon

composite. As seen in Table 3. the *Tsai-Wu* failure indices are over one value in a lot of beams. This means these beams will fail in the given critical load case.

This case is not acceptable because the chassis cannot fail in any of the critical load cases. However (see Table 3.) the maximum failure index is much less than one in a lot of beams. This is also not an ideal case because this means that the beams are over-dimensioned. For the optimal chassis structure a homogeneous stress state has to be achieved, i.e., a homogeneous maximum failure index distribution should be realized.

The results of Table 3. were considered as the starting point for the construction of the laminated shell structure. In those beams where the failure index is greater than one the ply number has to be increased. The ply number could be decreased where the failure index is less than one. The chassis is a symmetrical structure and this fact was also considered in the laminate plies construction. As a result the beams with the same function on the right side and on the left side of the chassis were constructed with the same ply number.

*Table 3. Values of failure criteria of plain weave composite chassis*

	1	2	3	4	5	6	7	8	9	10
CL1	1.0158	1.829	0.54	0.008	0.91	1.88	0.0009	0.0169	0.012	0.0067
CL2	1.32	0.85	0.749	0.0097	0.76	0.7	0.0028	0.295	0.479	0.0416
CL3	1.642	1.149	1.272	0.008	1.58	1.159	0.002	0.349	0.48	0.048
	11	12	13	14	15	16	17	18	19	20
CL1	0.049	0.04	3.37	0.31	0.00837	0.00997	0.15	0.15	2.72	2.68
CL2	0.167	0.12	0.66	0.0036	0.02	0.011	0.15	0.15	2.84	0.26
CL3	0.188	0.24	3.49	0.23	0.017	0.018	0.319	0.26	3.63	0.76
	21	22	23	24	25	26	27	28	29	30
CL1	1.06	0.0036	0.43	0.33	1.84	2.14	0.013	0.0013	0.051	0.025
CL2	1.01	0.15	4.14	3.14	0.7	0.71	0.014	0.011	0.14	0.046
CL3	1.2	0.16	4.52	3.1	0.68	0.71	0.017	0.008	0.29	0.044
	31	32	33	34	35	36	37	38		
CL1	0.024	0.71	0.63	0.028	0.034	0.00001	0.000009	0.000008		
CL2	0.27	0.1	0.14	0.063	0.048	0.0013	0.0017	0.0002		
CL3	0.24	0.2	0.29	0.072	0.068	0.0014	0.0015	0.00026		

The above computations showed that the shear-force is a dangerous internal force in lot of beams. Therefore, two plies with 45° warp orientation in the inner and outer side of the laminate structure were used in the beams where high shear-forces occur (see Fig. 13.). With this modification the following modified laminate plies structure was built up:

- Beams with 2 plies: 4, 7, 10, 11, 14, 15, 16, 22, 27, 28, 30, 35, 36, 37, 38
- Beams with 3 plies: 31, 34,
- Beams with 4 plies: 12, 17, 18,
- Beams with 5 plies: 29, 32, 33
- Beams with 6 plies: 3, 8, 9 ➔ 45°
- Beams with 8 plies: 1, 5, 2, 6, 23, 24, 25, 26 ➔ 45°
- Beams with 9 plies: 19, 20, 21 ➔ 45°
- Beams with 18 plies: 13 ➔ 45°

The list above shows that the ply numbers could be decreased in most of the beams compared to the basic initial laminate structure. Also it can be observed that there are a few beams in which the ply number has to be increased. From the computation one can state that the beam number 13 is the most critical beam in the chassis. In this beam the ply number has to be increased up to 18.

The final weight of this above modified laminate structure is 7.17 kg. This means that the chassis is approximately 12 kg applying composite materials, which means 62.5 % weight reduction compared to the aluminium alloy chassis.
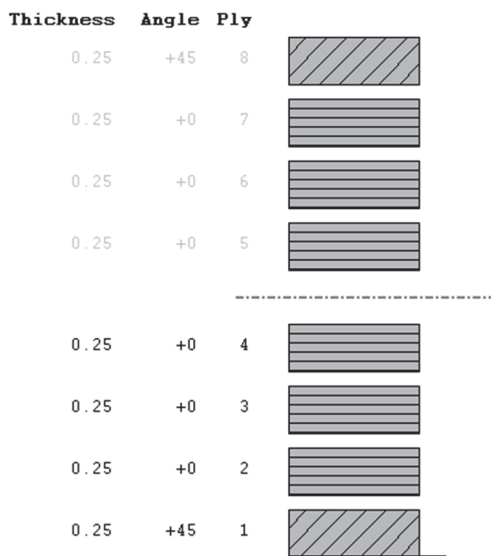
Figure 13. Ply structure in beams number 1, 5, 2, 6, 23, 24, 25, 26

5. Conclusion

The paper focused on a real-life engineering problem experienced in the Shell Eco marathon competition. In this race the weight of the car is key to reach success like in any other car race. The paper shows an example of how laminated composite materials can be used for a chassis structure and therefore to reduce weight. For the FEM strength analysis different mechanical models of the chassis with critical load cases were applied. The FEM strength analysis was carried out in three critical load cases.

From the results of beam-bar modelling an overview was obtained concerning stress distribution in metal chassis. From the initial identical wall-thickness of the composite construction a laminate composite chassis was constructed with nearly homogeneous stress distribution. The weight of the basic laminated composite structure with 8 plies was 8.77 kg. From the results of FEM analysis of the basic laminated model an optimal layer structure was achieved. This means that every beam has nearly a homogeneous

stress state. The weight of the final laminated composite chassis structure was only 7.17 kg.

Having compared the final laminated carbon composite chassis to the aluminium alloy chassis the weight reduction is almost 12 kg, namely 62.5 %. Using this chassis the fuel consumption can be significantly reduced.

References

[1] Mat MH, Ab Ghani AR: Design and Analysis of 'Eco' Car Chassis. Procedia Engineering, Vol. 41, pp. 1756–1760, 2012.
DOI: 10.1016/j.proeng.2012.07.379

[2] Heißing B, Ersoy M: Chassis Handbook. Vieweg+Teubner Verlag, Berlin, 2011.

[3] Hinton MJ, Kaddour AS, Soden PD: Failure criteria in fibre-reinforced-polymer composites. Elsevier Science Ltd, Oxford, 2004.

[4] Kollár PL, Springer GS: Mechanics of Composite Structures. Cambridge University Press, Cambridge, 2003.

[5] Matthews FL, Davies GAO, Hitchings D, Soutis C: Finite element modelling of composite materials and structures. Woodhead Publishing Limited, Cambridge, 2000.

[6] Hyer MW: Stress analysis of fiber-reinforced composite materials. McGraw-Hill Companies, Singapore, 1998.

[7] Kaw AK: Mechanics of composite materials. CRC Press LLC, New York, 1997.

[8] Elhajjar R, Saponara LV, Muliana A: Smart composites Mechanics and Design. CRC Press, New York, 2014.

[9] Herakovich CT: Mechanics of fibrous composites. John Wiley & Sons Inc., New York, 1998.

[10] Barbero EJ: Finite element analysis of composite materials. CRC Press, Boca Raton, 2008.

[11] Barbero EJ: Finite element analysis of composite materials using ANSYS. CRC Press, Boca Raton, 2014.

[12] Aczél Á, Bojtár G, Fehér L, Keresztes D: FE Analysis of the 3D bar model of the chassis of a racing car (Versenyautó alváz térbeli rúdmodelljének végeselemes analízise). GÉP, Vol. 7-8, pp. 3-8., 2011.

[13] Égert J, Aczél Á, Fehér L, Körmendy Á: Modeling possibilities and critical loading of the chassis of a racing car (Versenyautó alváz mechanikai modellezési lehetőségei és kritikus terhelései). GÉP, Vol. 7-8, pp. 43-48., 2011.

[14] Lawrence M: Colin Chapman Wayward Genius. Breedon Books Publishing, 2003.

Upgrading Multilevel Full Cost Allocation: Case Transport

Z. Bokor[1], B. Horváth[2]

[1]Budapest University of Technology and Economics, Department of Transport Technology and Economics
Műegyetem rkp. 3., 1111 Budapest, Hungary
Phone: +36 1 463 1054
e-mail: zbokor@kgazd.bme.hu

[2] Széchenyi István University, Department of Transport
Egyetem tér 1., 9026 Győr, Hungary

Abstract: Multilevel full cost allocation is an effective alternative of traditional, accounting oriented costing regimes and is applied in the transport sector as well. Nevertheless, its basic algorithm may fail when the cost calculation model contains so called performance feedbacks. This is often the case in transport management practice, especially in complex operational systems. Hence effective solutions are to be found to cope with this methodological problem. This paper aims at contributing to the improvement of multilevel full cost allocation method when performance feedbacks hinder the exact calculation. The application of estimated cost ratios or specific costs is proposed and tested on a theoretical transport costing example. If some methodological requirements can be met both parameter estimations may be employed effectively in calculation tasks with performance feedbacks.

Keywords: *multilevel full cost allocation, performance feedbacks, transport costing*

1. Introduction

Activity-based costing (ABC) is a widely used costing tool. It aims at allocating indirect costs, i.e. costs that can not be assigned to products or services directly, to products or services on the basis of measured performance consumptions. Indirect costs are traced to activities first by using so called resource drivers. Then activity costs are distributed among products or services by using cost drivers, i.e. various performance indicators. By doing so a transparent and traceable cost allocation system can be set up instead of using arbitrary allocation techniques distorting cost information [7].

Multilevel full cost allocation (MFCA) is a similar method having the same advantages. Here, however, organisational units are used instead of activities for the allocation of indirect costs. These units are arranged into a multilevel hierarchy

depicting the operational structure of the company [2]. There are transport related applications reported in the literature for both methods.

The most detailed transport oriented ABC application was carried out at a road freight transport company. Here the costs of service packages were determined and the results were compared to the outputs of traditional costing regime [1]. Road transport was in the focus of other, more complex ABC analyses as well [13]. ABC was used to calculate the costs of elementary air transport services, i.e. individual flights [14]. Furthermore, ABC was combined with other decision making techniques where the financial efficiency of air transport services was evaluated [11].

The cost structure of third-party logistics service providers was also analysed by ABC with special regard to transport activities [10]. The activity-based cost structure of logistics service providers focusing on road freight transport was studied by highlighting the role of fixed and variable costs [6].

MFCA has a moderate application intensity compared to the one of ABC. Sample transport cost calculation models have been developed for the cases of rail infrastructure management [2], bus transport [3], rail freight transport [4] and rail passenger transport [5].

This paper focuses on multilevel full cost allocation by making also use of available research results. Besides its advantages, MFCA models used in practice still have some limitations. One of the limitations is the fact that it neglects performance feedbacks, which may reduce accuracy. Performance feedbacks mean that performance flows facilitating the operation and governing the allocation of indirect costs are not always straightforward and there may be performances that stream backwards in the system. This is the case when some organisational units have mutual performance services and their operation depends on each other's outputs. Thus the allocation of indirect costs can not be carried out exactly. This is a real problem in transport costing as the operational models establishing the calculation models, mainly in complex systems, often consist of objects interacting with each other in various ways.

Before analysing the impacts of performance feedbacks, the basic MFCA calculation model is described also with view on the latest relevant research results [4, 5]. Note that fixed and variable costs are not separated in the applied approach as the further analysis of performance feedbacks does not require this distinction. Nevertheless, in the full scale cost calculation models the segmentation of costs into fixed and variable parts is also recommended provided the segmentation can be carried out in an exact way [6].

The basic MFCA model consists of the cost centres ($k, i = 1…n$) and the profit objects ($j = 1…m$). The former ones are the organisational units and perform for the profit objects or for other cost centres. Nevertheless, performance feedbacks, i.e. when cost centres serve each other mutually, may not be incorporated into the basic model so that the calculation can be performed in an exact way. The cost of cost centres can be calculated in the following way:

$$C_k = C_k^p + \sum_{i=1}^{n} C_i \frac{P_{ki}}{P_i} \tag{1}$$

where:

C_k: total cost of cost centre k;

C_k^p: primary cost of cost centre k;

C_i: total cost of service cost centre i;

P_{ki}: performance consumption of cost centre k at service cost centre i;

P_i: total performance of service cost centre i.

Primary costs are assigned to cost centres directly and are obtained from the general ledger with certain transformations, if needed. Performance data are provided by technology information systems or by dedicated surveys.

Profit objects are the ultimate products or services in the examined business-technology system. The main goal of MFCA is to calculate the costs of profit objects. The basic formula of this calculation is the following:

$$C_j = C_j^d + \sum_{i=1}^{n} C_i \frac{P_{ji}}{P_i} \qquad (2)$$

where:

C_j: total cost of profit object j;

C_j^d: direct cost of profit object j;

P_{ji}: performance consumption of profit object j at service cost centre i.

Direct costs are assigned to profit objects directly and are obtained from the general ledger, too. As MFCA is a full cost allocation method it is required that all indirect costs, i.e. the costs of cost centres, are fully allocated by the end of the calculation process:

$$\sum_{k=1}^{n} \frac{P_{ki}}{P_i} + \sum_{j=1}^{m} \frac{P_{ji}}{P_i} = 1 \qquad (3)$$

As mentioned, the calculation is always based on an operational model which depicts the business and technology processes of the investigated company or organisational system. The following analysis highlights what methodological problems can be faced when the operational model and so the calculation model contain performance feedbacks.

2. Initial problem

The problem of performance feedbacks is illustrated using a simple transport costing example. Fig. 1 shows a simplified operation model of a transport company operating two kinds of services, i.e. regular and irregular transport services. These are the profit objects (j = A, B) in the corresponding calculation model. Four organisational units take part in the production of these services: two of them directly, i.e. vehicle pools I and II, while other two ones, i.e. operation and maintenance, indirectly. These units are regarded as cost centres (k, i = 1…4). The possible performance indicators (and their dimensions) can be for example:

- operation: disposition (number of orders);

- maintenance: service time (hours);

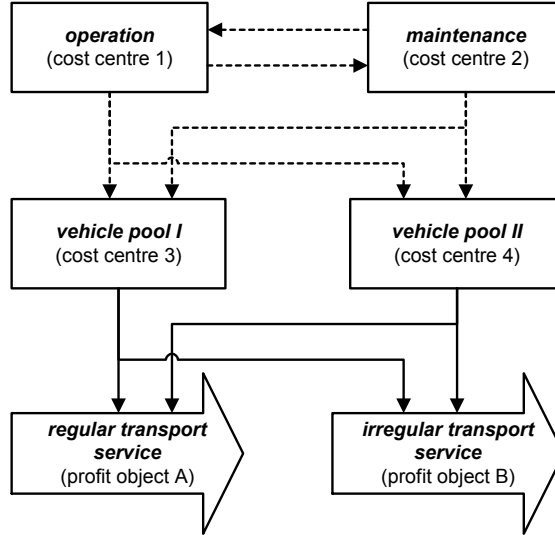- vehicle pools I and II: running (vehicle kilometres).

Figure 1. Sample operation model

The vehicle pools produce the transport services while the operation and the maintenance units serve the vehicle pools. The latter units serve each other as well. It means that the model contains a performance feedback in the relation of cost centre *1* and cost centre *2*.

The calculation formulas based on the extraction of (1) and (2) are the following:

$$C_1 = C_1^p + C_2 \frac{P_{12}}{P_2} \tag{4}$$

$$C_2 = C_2^p + C_1 \frac{P_{21}}{P_1} \tag{5}$$

$$C_3 = C_3^p + C_1 \frac{P_{31}}{P_1} + C_2 \frac{P_{32}}{P_2} \tag{6}$$

$$C_4 = C_4^p + C_1 \frac{P_{41}}{P_1} + C_2 \frac{P_{42}}{P_2} \tag{7}$$

$$C_A = C_A^d + C_3 \frac{P_{A3}}{P_3} + C_4 \frac{P_{A4}}{P_4} \tag{8}$$

$$C_B = C_B^d + C_3 \frac{P_{B3}}{P_3} + C_4 \frac{P_{B4}}{P_4} \tag{9}$$

Let us calculate the cost of profit object *A*:

$$C_A = C_A^d + C_3^P \frac{P_{A3}}{P_3} + C_1 \frac{P_{31}}{P_1} \frac{P_{A3}}{P_3} + C_2 \frac{P_{32}}{P_2} \frac{P_{A3}}{P_3} + C_4^p \frac{P_{A4}}{P_4} + C_1 \frac{P_{41}}{P_1} \frac{P_{A4}}{P_4} + C_2 \frac{P_{42}}{P_2} \frac{P_{A4}}{P_4} \tag{10}$$

It is obvious that the calculation cannot be executed on a regular basis because C_1 and C_2 depend on each other as can be seen in (4) and (5). Additional considerations are needed to overcome this methodological problem and to calculate the costs of profit objects.

3. Proposed solutions

3.1. Estimation of cost ratios

One of the possible solutions is to introduce the parameter of total cost ratio (r_{ki}) for the case of cost centres affected by performance feedbacks:

$$r_{ki} = \frac{c_k}{c_i} \tag{11}$$

for *ki* representing pairs of cost centres with performance feedbacks. Let us introduce $r_{12} = C_1 / C_2$ in our example. Substituting by r_{12} the missing cost elements can be determined from (4) and (5):

$$C_1 = \frac{c_1^p}{1 - \frac{P_{12}}{r_{12}P_2}} \tag{12}$$

$$C_2 = C_2^p + \frac{c_1^p P_{21}}{(1 - \frac{P_{12}}{r_{12}P_2})P_1} \tag{13}$$

Using (12) and (13) for C_1 and C_2 respectively, the final cost calculation of profit objects, e.g. the one of profit object *A* in formula (10), can be executed exactly.

Another task is to estimate the value of the parameter applied:

- the simplest estimation for total cost ratio is the ratio of corresponding primary costs. It can be an acceptable solution when the total costs do not differ from primary costs significantly or are approximately proportional to total costs. Usually this is not the case;

- an alternative estimation can be the ratio of measurable natural indicators, like staff numbers, numbers of machines or appliances, etc., of corresponding cost centres. It shall be ensured that the indicator chosen is relevant for both cost centres concerned, and furthermore, that the indicator can be regarded as a real influencing factor of costs. The latter condition can be justified by correlation analyses on the basis of historical data, provided these data are available;

- combining and extending the former approaches by using the ratios of multiple factors (*l* = 1…q) which may be primary costs or selected natural indicators. Theoretically this leads to the best estimation thanks to the multiple factors considered, provided the conditions mentioned are observed and the input data are available:

$$\hat{r}_{ki} = \prod_{l=1}^{q} \frac{f_{l_k}}{f_{l_i}} \tag{14}$$

where:

$\hat{r}_{ki}$: estimated value of total cost ratio ki;

f_{l_k}: factor l at cost centre k;

f_{l_i}: factor l at cost centre i.

In our example the total costs of operation and maintenance are to be compared. The corresponding primary cost data are available but it is not sure that they reflect the total costs. Nevertheless, some natural indicators are likely to be found for the estimation. For example, the material costs of maintenance might be recorded in the served cost centres and so the main driver of maintenance costs will be labour, staff numbers or working hours. These may be appropriate indicators for estimating the cost ratio as operation is also a labour-intensive business area. Of course, the suitability of the indicator chosen depends on the operational characteristics of the company assessed. Last but not least, if more than one indicator is found appropriate or we would like to combine the ratio of primary costs and the ratio of natural indicators the complex estimation formula (14) can also be applied here.

3.2. Estimation of specific costs

Another possible solution for handling the methodological problem mentioned is to introduce the parameter of specific cost (c_i) for the case of cost centres affected by performance feedbacks:

$$c_i = \frac{C_i}{P_i} \tag{15}$$

for i representing cost centres having performance feedbacks. We can use c_1 or c_2 in our example, let us introduce $c_2 = C_2 / P_2$. Substituting by this parameter the missing cost elements can be determined from (4) and (5) again:

$$C_1 = C_1^p + c_2 P_{12} \tag{16}$$

$$C_2 = C_2^p + C_1^p \frac{P_{21}}{P_1} + c_2 \frac{P_{12} P_{21}}{P_1} \tag{17}$$

Using (16) and (17) for C_1 and C_2 respectively, the final cost calculation of profit objects, e.g. the one of profit object A in formula (10), can also be executed exactly. The question arises again: how to estimate the value of the parameter. There are two alternatives:

- using the corresponding primary specific cost values increased by estimated percentages based on relevant practical expertise. The subjectivity of this approach can be reduced by applying group decision making methods like analytic hierarchy process (AHP) or by conducting sensitivity analyses;
- using benchmark values, e.g. the average prices of similar activities or services. As benchmark values reflect the actual prices and are probably less subjective this approach seems to be a better solution, provided the benchmark data are available and the prices offered are cost-based.

In our example the specific cost of maintenance shall be determined with the dimension of (EUR / service hour). If we have experts with reliable expertise in the field we can rely on their estimations. It is, however, reasonable to obtain benchmark values

as prices of maintenance services may easily be obtained, e.g. by calling for bids. Nevertheless, the quotations shall be reviewed carefully to ensure their applicability in the costing model.

Concerning AHP it shall be noted that this method has an extensive application potential in transport. For example, it was successfully applied in the quality analysis of public transport systems [8, 9]. These research results may also be useful when AHP is utilised for supporting cost estimations.

4. Discussion and conclusions

The additional procedures proposed may help improve the accuracy of multilevel full cost allocation models containing performance feedbacks, with special regards to transport related applications. Comparing the two solutions it can be concluded that the application of estimated specific costs involves, in general, less risks than the one of estimated cost ratios. This can be explained by the better availability of input data, i.e. benchmark values. Nevertheless, if the estimation is company specific, then cost ratios may be more favourable as they use multiple factors describing the operational characteristics of the particular business-technology system. So it depends on the data availability and the circumstances of the estimation which approach is worth being preferred.

Although the proposed solutions have been tested on a simple example they can also systematically be applied in multilevel cost calculation models with multiple feedbacks. Of course the complexity of additional calculations is in line with the number of feedbacks.

When using the improved MFCA model the limitations of estimations shall also be considered as the estimations may contain some errors. These errors can be reduced by using further statistical or decision supporting methods like correlation analysis or AHP. Furthermore, there are some conditions which are to be observed when applying estimated cost ratios or specific costs during the calculation process. These requirements, i.e. the suitability of selected factors or benchmarks, etc., have been discussed in detail.

It shall be noted that iterative calculation procedures may also contribute to solve the problems of performance feedbacks in MFCA models. Here initial values for the missing parameters are set and the calculation programme runs until the calculated values do not change significantly. Nevertheless, the approach proposed in this paper aims at making advantages of the available information instead of iteration. However, the two different approaches, i.e. iteration and the use of estimated cost ratios or specific costs, may be effective complementary tools of cost calculation tasks that are not calculable in exact ways. This may often be the case in complex transport systems.

And at last it is worth mentioning that if the general or overall efficiency of complex transport systems is to be assessed single cost calculations may not be sufficient. Cost is namely one of the factors determining efficiency. Here more comprehensive and non-parametric methods like data envelopment analysis (DEA) can be utilised [12].

References

[1] Baykasoglu A, Kaplanoglu V: Application of Activity-based Costing to a Land Transportation Company: a Case Study. International Journal of Production Economics, Vol. 116, No. 2, pp. 308-324, 2008.
DOI: 10.1016/j.ijpe.2008.08.049

[2] Bokor Z: Rail Infrastructure Costing Based on Multi-level Full Cost Allocation. Periodica Polytechnica. Transportation Engineering, Vol. 40, No. 1, pp. 3-9. 2012.
DOI: 10.3311/pp.tr.2012-1.01

[3] Bokor Z: Cost Calculation in Transport Companies. Acta Technica Jaurinensis, Vol. 5, No. 3, pp. 253-262. 2012.

[4] Bokor Z: Cost Calculation in Complex Transport Systems. LOGI – Scientific Journal on Transport and Logistics, Vol. 4, No. 1, pp. 5-22, 2013.

[5] Bokor Z: Multilevel Full Cost Allocation Model in Transport (in Hungarian). Budapest Management Review (Vezetéstudomány), Vol. 45, No. 1, pp. 55-68, 2014.

[6] Bokor Z, Markovits-Somogyi R: Applying Activity-based Costing at Logistics Service Providers. Periodica Polytechnica. Transportation Engineering, Vol. 43, No. 2, pp. 98-105, 2015.
DOI: 10.3311/PPtr.7700

[7] Cooper R, Kaplan R S: How Cost Accounting Systematically Distorts Product Costs. In Accounting and Management – Field Study Perspectives, Editors: Burns W J, Kaplan R S, Harvard Business School Press, Boston, pp. 204-228, 1987.

[8] Duleba Sz: A Hierarchical Model to Evaluate Public Transport's Supply Quality. Acta Technica Jaurinensis, Vol. 3, No. 3, pp. 377-382, 2010.

[9] Duleba Sz, Shimazaki Y, Mishina T: An Analysis on the Connections of Factors in a Public Transport System by AHP-ISM. Transport, Vol. 28, No. 4, pp. 404-412, 2013.
DOI: 10.3846/16484142.2013.867282

[10] Griful-Miquela C: Activity-based Costing Methodology for Third-party Logistics Companies. International Advances in Economic Research, Vol. 7, No. 1, pp. 133-146, 2001.
DOI: 10.1007/BF02296598

[11] Lin W-Ch: Financial Performance and Customer Service: an Examination Using Activity-based Costing of 38 International Airlines. Journal of Air Transport Management, Vol. 19, No. 1, pp. 13-15, 2012.
DOI: 10.1016/j.jairtraman.2011.12.002

[12] Markovits-Somogyi R: Data Envelopment Analysis and its Key Variants Utilized in the Transport Sector. Periodica Polytechnica ser. Transportation Engineering, Vol. 39, No. 2, pp. 63-68. 2011.
DOI: 10.3311/pp.tr.2011-2.03

[13] Nurminen T, Korpunen H, Uusitalo J: Applying the Activity-based Costing to Cut-to-length Timber Harvesting and Trucking. Silva Fennica, Vol. 43, No. 5, pp. 847-870, 2009.

[14] Tsai W-H, Kuo L: Operating Costs and Capacity in the Airline Industry. Journal of Air Transport Management. Vol. 10, No. 4, pp. 271-277, 2004.
DOI: 10.1016/j.jairtraman.2004.03.004

Measurement and Analysis of Deformation Shapes on Corrugated Cardboard Logistical Boxes under Static and Dynamic Compression

P. Böröcz

Széchenyi István University, Department of Logistics and Forwarding
Egyetem tér 1., 9026 Győr, Hungary
Phone: +36 70 335 2260
e-mail: boroczp@sze.hu

Abstract: On the field of compression strength of packaging the researchers, studies and standards mainly focus on the short-time measurements. The derived results from these accelerated tests can be used with a large tolerance in reality. These measurements do not give response that how long the packaging is able to tolerate the given calculated or measured stress during a long-term storage in the supply chain without damages in the aspect of product or packaging. The tests in this paper were done to analyse the possible deformation shapes, which could be a first step into the process of final damage. The samples were tested under both static and dynamic compression. The deformations were observed by TRITOP system in order to prove the different deformation processes between the static and dynamic compression. This paper is a first phase of a research program, which deals with the effect of real compression and physical environment of logistics boxes on long time storage and stacking.

Keywords: packaging, compression test, TRITOP

1. Introduction

Tests of various international standards (FEFCO, ISO, ASTM, TAPPI) use only short-time measurements to assess the compression strength of packaging materials made of corrugated cardboard sheet (HPL) and this is partly when they have contents during the time of the measurement [1][2][3][4]. Historically, HPL industry has established the nature and scope of compression strength measurements. Experts believe that the measurement results derived from BCT (Box Compression Test) allow them to evaluate the compressive behaviour of the boxes. Compression strength measurements were mainly developed for quality control purposes. Furthermore, large production volumes do not allow any of the test methods to be applied for a relatively long duration.

Some previous studies dealing with the mechanical behaviours of these packaging structures under static compression [5][6][7], and some other papers show results and

brief reviews of designing and modelling the stack-ability of cardboard boxes using finite element method (FEM) [8][9]. However, deformation images and shapes occurring during quasi-dynamic laboratory testing and static load have not been dealt with so far. This paper is aimed to examine deformation events occurring through static load and deformation events occurring during laboratory measurement, and then compare them. These kind of quasi-dynamic and static tests are usually applied to observe the cushioning and strength effect of other packaging materials [10][11] such as plastic foams. The results of these tests are applied as input data to decision support models and processes to choose the right protective packaging system [12]. However, it is well known that the long time distribution environment and storage can affect the real strength of the packaging especially the corrugated paper box packaging [13].

This study begun with the conduction of traditional BCT tests, and box sidewalls were inspected for deformations. Load was then gradually reduced on the packaging material all the way until the packaging material had the ability to carry the static load for at least 1 hour. This method was also used to analyse the deformation processes of the sidewalls under static load.

During the tests, a press recognised in the standards was used for the analysis of the deformation images, and also a system called TRITOP. Type and quality of the corrugated cardboard box used is FEFCO 0201 - 26C. The main flutes are composed of 175KL (Kraftliner), 150HP (semi-chemical fluting) and 200TL1 (Testliner 1) with a weight of 597 g/m2. Box dimensions were 600 mm x 400 mm x 300 mm (L x W x H).

Tests and results presented in this paper constitute the first stage of a research that will eventually focus on the more accurate estimation of static compressive strength of corrugated cardboard boxes as a function of load and compressive resistance. The research may be able to answer the question of how much the results of the methods used in laboratory practice are suitable for the evaluation of actual compression strength.

2. Measurement system and setup

The deformation processes were analysed with the mentioned TRITOP system above. This measurement system is an industrial optical measurement system for three-dimensional acquisition of discrete object points. While in operation, the instrument marks every essential object point and employs a photogrammetric camera to capture images about the object, and then it records the measurement from various angles. Subsequently, the instrument software automatically calculates the three-dimensional coordinates and physical properties of the markers using the digital images.
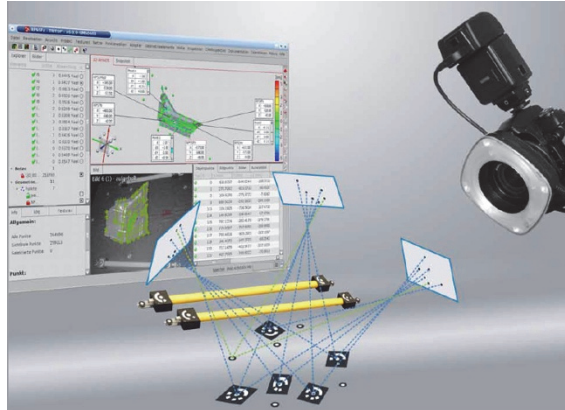
Figure 1 – Structure of the TRITOP system

The basic idea of measuring light intensity is that it views points from multiple directions and uses light rays to calculate the three-dimensional coordinates. The reference points shown in the figure below must be fixed relative to each other. Consequently, camera position relative to the reference points can be calculated for each specific case using images captured from various angles. During the acquisition of the image set, the aim is to record reference points from several angles of view so that they can show one another the largest possible angle. This is demonstrated in Fig. 2.
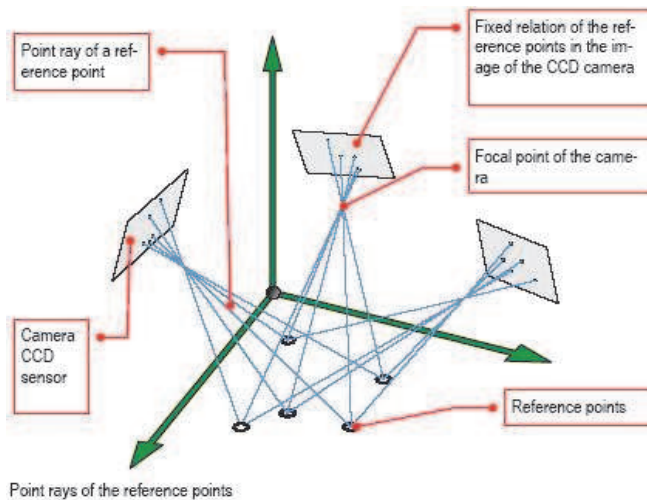
Figure 2 – Spatial operating principles of the TRITOP system

The legends, shown in the figure, have the following meaning:

Point ray of a reference point: Point ray of a reference point

Camera CCD sensor: Camera CCD sensor

Fixed relation of the reference points in the image of the CCD camera

	Fixed relation of the reference points in the image of the CCD camera
Focal points of the camera:	Focal points of the camera
Reference points:	Reference points

The main goal of the software is to accurately locate omissions and gaps (a bird's-eye view of the reference points) in each image within the image set and in 3D directions of the image. Measurement data can be evaluated in the TRITOP system (CAD comparison). The coded reference points of TRITOP allow the setting of an image that evaluates and automatically calculates the locations of the camera. When capturing an image, the aim is to see as many reference points as possible and preferably reach beyond the entire object so as to obtain an accurate measurement.

Figure 3 – ID marks of the uncoded points of the TRITOP system

The orientation crosses are factory-equipped with coded points. Uncoded reference points (Fig. 3.) are used to determine the 3D coordinates and are identified automatically by the instrument. If the uncoded reference points are positioned on the object, the measurement method determines the location of these points. For determining the dimensions reference is provided by scale bars (Fig. 4). The software package of the system then also allows you to determine how much each single point has migrated from their original location in the X:Y:Z coordinates. The extent of how much each single point has migrated from its original location can be determined with an accuracy of $1/1000^{th}$ of a millimetre.

Figure 4 - Scale bars of the system

3. Measurement and observation method

The following measurement programme was devised for observing the deformation processes:

- 25 boxes of identical corrugated cardboard material and identical geometric size were selected and kept under identical climatic conditions.
- 5 boxes were measured for standard BCT values.
- 5 boxes loaded into the press based on the BCT value were provided with the markers of the TRITOP system and subjected to load with quasi-permanent feed until the boxes collapsed. But since the TRITOP system operates on the principle of imaging, we stopped increasing load while the pictures were being taken. Load was set so as to increase compression force with 50N increments using constant feed rate, then feeding was stopped, the images regarding compression were taken (an approximately 5-minute period of time), then load was again increased with the standard feed rate of 50N, feeding was stopped, images were taken and this process continued until the box got completely damaged.
- Based on the BCT value, 5 boxes were respectively loaded with an evenly distributed static compression that equals 80%, 70% and 60% of the BCT value and the TRITOP system was used to take pictures every 24 hours for deformation analysis. Naturally, the box was kept under the same climatic conditions during static compression.
- Under static compression, TRITOP images were only evaluated for boxes, which were able to withstand static compression for at least 1 hour.

The measurement results were compared with and evaluated against deformation images captured during the test. During the evaluation of the deformation images, it was determined from within the inside of the box whether the deformation appeared concave or convex for each side.

4. Results and discussion

The original BCT value of the box was measured at 1340N on average. During static testing, the box was able to resist loading for longer than 1 hour if the compression was as low as 80 kg – any higher compression caused the box being tested to collapse right away. The measurement results are shown in Table 1.

Table 1. The measured BCT values and the period to collapse during static testing besides varied BCT load

S.	BCT (N) Measured values	Static 80% BCT		Static 70% BCT		Static 60% BCT	
		Time to collapse (min, hour)					
1	1 360	< 1 min	-	< 1 min	-	> 1 hour	✔
2	1 320	< 1 min	-	< 1 hour	-	> 1 hour	✔
3	1 240	< 1 min	-	< 1 min	-	> 1 hour	✔
4	1 390	< 1 min	-	< 1 hour	-	> 1 hour	✔
5	1 310	< 1 min	-	< 1 min	-	> 1 hour	✔

Deformation under 80 kg of compression within the first 24 hours and deformation produced on the last day of compression (day 18) is presented in Fig. 5a and Fig. 5b respectively. Intermediate states are not published for considerations of space.
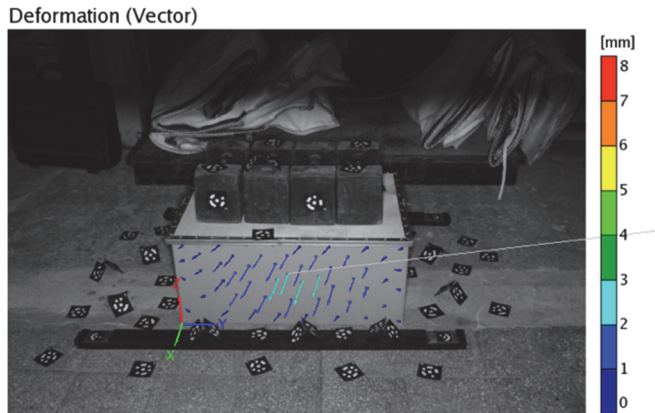
Figure 5a – Deformations occurring as a result of 800N static load after the first 24 hours of measurement
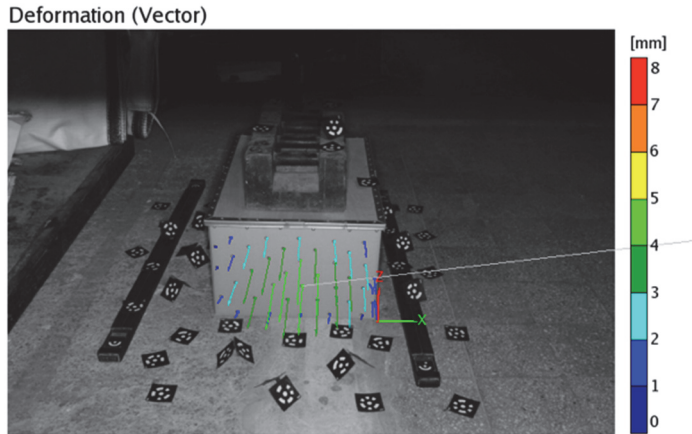
Figure 5b – Deformation occurring as a result of 800N static load after 18 days

Deformation occurring at a load of 1200N during quasi-standard BCT testing is shown in Fig. 6a and 6b. As the reduction of the feed rate reduces the peak load of BCT measurement to a significant extent, this phenomenon occurred here as well, and the box cracked under a load of 1250N.
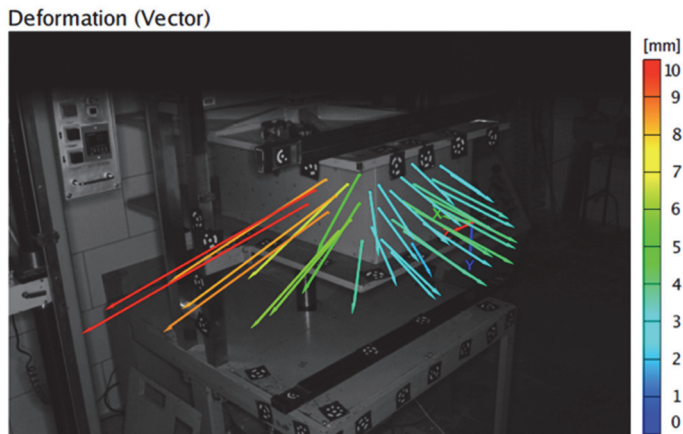

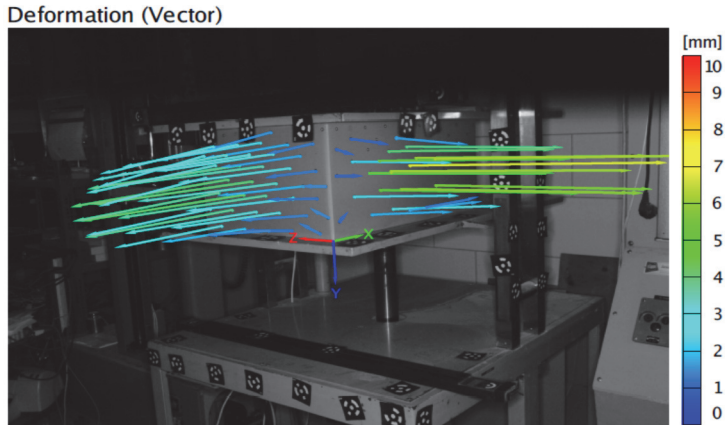*Figure 6a – Deformation occurring at a load of 1200N during quasi BCT testing (left side)*

Deformation (Vector)

Figure 6b. – Deformation occurring at a load of 1200N during quasi BCT testing (right side)

Table 2 shows the shape of deformation of each sidewall of the boxes under static and quasi-dynamic loads. Deformation is characterised by either a convex or a concave shape depending on whether the deformation of the sidewall is evaluated from inside the boxes. The results show that it was only in a total of 3 cases that the deformation of a sidewall was concave out of the 20 sidewalls of the five samples. However, in the case of quasi-dynamic measurements, each pair of the opposing walls almost always suffered concave deformation viewed from the inside.

Table 2. Shape of deformation during the tests

Sample number			Quasi dynamic	Static
1	Short sides	left / right	convex / convex	convex / convex
	Long sides	left / right	concave / concave	convex / convex
2	Short sides	left / right	convex / concave	convex / convex
	Long sides	left / right	convex / convex	convex / concave
3	Short sides	left / right	concave / convex	convex / concave
	Long sides	left / right	concave / concave	convex / convex
4	Short sides	left / right	convex / concave	concave / convex
	Long sides	left / right	concave / concave	convex / convex
5	Short sides	left / right	concave / concave	convex / convex
	Long sides	left / right	convex / convex	convex / convex

5. Conclusion

The measurements described above verify the following:

- There is a significant difference between the deformation images of BCT measurements and static load measurements. The two measurements are not considered compatible with each other.

- The deformation images of the BCT value measured in laboratory practice are not in accordance with the deformation images of the boxes captured in real life circumstances.
- The two different measurement methods show significantly varying skews with the alternating side walls bending in both convex and concave direction seen from the inside in the case of BCT, and with each side wall bending in convex direction seen from the inside in the case of static load.
- The results of this study can be adapted to determine the possible damages under a real storage period, thereby to reduce the protective packaging costs, and also to provide more sustainable solutions.

References

[1] Test Method TAPPI T804 (om-12): Compression test of fiberboard shipping containers, 2012.

[2] ASTM D642 – 00: Standard Test Method for Determining Compressive Resistance of Shipping Containers, Components, and Unit Loads, 2010.

[3] FEFCO TM 50, FEFCO Testing Method Determination of the Compression Resistance of corrugated fibreboard containers, 1997.

[4] ISO 12048 - Complete, filled transport packages – Compression and stacking tests using a compression tester, 1994.

[5] Aboure, Z, Talbi N, Allaoui S, and Benzeggagh M: Elastic behavior of corrugated cardboard: experiments and modelling. Composite Structures, Vol. 63, No. 1, pp. 53-62, 2004.
DOI: 10.1016/S0263-8223(03)00131-4

[6] Sek, M. A, Kirkpatrick J: Prediction of the cushioning properties of corrugated fibreboard from static and quasi-dynamic compression data. Packaging Technology and Science, Vol. 10, No. 2, pp. 87-94, 1997.
DOI: 10.1002/(SICI)1099-1522(199703/04)10:2<87::AID-PTS389>3.0.CO;2-L

[7] Marcondes JA: Effect of load history on the performance of corrugated fibreboard boxes. Packaging technology and science, Vol. 5, No. 4, pp. 179-187, 1992.
DOI: 10.1002/pts.2770050403

[8] Talbi N, Batti A, Ayad R, Guo Y, An analytical homogenization model for finite element modelling of corrugated cardboard. Composite Structures, Vol. 88, No. 2, pp. 280-289, 2009.
DOI: 10.1016/j.compstruct.2008.04.008

[9] Biancolini ME, Brutti C, Porziani S: Corrugated board containers design methods. International Journal of Computational Materials Science and Surface Engineering, Vol. 3, No. 2, pp. 143-163, 2010.
DOI: 10.1504/IJCMSSE.2010.033150

[10] Mojzes Á, Böröcz P: □Predicting cushion characteristic on new type of environmental friendly foam. Acta Technica Jaurinensis, Vol. 3, No. 3, pp. 395-404, 2010.

[11] Mojzes Á, Földesi P, Böröcz P: Define Cushion Curves For Environmental Friendly Packaging Foam. Annals of faculty of engineering hunedoara - International Journal of Engineering, Vol. 10, No. 1, pp. 113-118, 2012.

[12] Mojzes Á, Böröcz P: Decision Support Model to Select Cushioning Material for Dynamics Hazards During Transportation. Acta Technica Jaurinensis, Vol. 8, No. 2, pp. 188-200, 2015.
DOI: 10.14513/actatechjaur.v8.n2.369

[13] Böröcz P, Singh S. P, Singh J: Evaluation of Distribution Environment in LTL Shipment between Central Europe and South Africa. Journal of Applied Packaging Research, Vol. 7, No. 2, 2015.

Finite Element Analysis of Glass Vacuum Windows of the *"COMPASS"* Tokamak in Prague

Z. Molnár[1], J. Égert[1], M. Berta[2,3]

**[1]Széchenyi István University, Department of Applied Mechanics
Egyetem tér 1., 9026 Győr, Hungary
E-mail: molnarz@sze.hu**

**[2]Széchenyi István University, Department of Physics and Chemistry
Egyetem tér 1., 9026 Győr, Hungary**

**[3]Institute of Plasma Physics AS CR, v.v.i., Tokamak Department
Za Slovankou 1782/3, 182 00 Prague, Czech Republic**

Abstract: In this paper is described the mechanical stress analysis of glass vacuum windows of the *"COMPASS"* tokamak using Finite Element Method. As a reference test problem the problem of uniformly loaded glass vacuum window with circular shape has been chosen, which problem can be solved also analytically. In the case of Finite Element analysis we used two different mechanical models: 3-dimensional (3D) approach and shell modelling. For 3D approach an optimal finite element type - quadratic tetrahedron element with appropriate size - was chosen, and for shell modelling a thin axisymmetric shell element type - shell 209 - was applied. The analytic and both numerical results have been compared and summarized.

Keywords: stress analysis, Finite Element Method (FEM), glass vacuum windows, the "COMPASS" tokamak, shell element, deformation, equivalent (von Mises) stress

1. Introduction

The purpose of analysis presented in this paper is to verify mechanical resistance of vacuum windows on the *"COMPASS"* tokamak through numerical stress analysis. This analysis has been carried out using Finite Element (FE) modelling. As the FE modelling tool we used the programme package ANSYS.

The main structure of this paper is the following:

- after the general introduction (Chapter 1.) we introduce the *"COMPASS"* tokamak in Prague and the geometrical model of its vacuum windows as objects of our tests (Chapters 2.,3.);

- as a reference problem we remind the analytical solution for displacement and stress in the case of window with circle shape fixed along its perimeter and with uniformly distributed pressure as the load (Chapter 4.);

- in Chapter 5. we solve the problem for different shapes numerically (as a 3D model and also as a shell model);

- and finally, Chapter 6. consists of the summary of results and conclusions.

2. The *"COMPASS"* tokamak and its glass vacuum windows

The *"COMPASS"* tokamak of IPP-CR in Prague is a medium size fusion device producing plasma with ITER-like shape [1]. Proper operation of such fusion devices need to fulfil some necessary prerequisites (Fig. 1.):

- enough energy to produce and maintain high magnetic fields and huge electric currents;

- very high vacuum in the reaction chamber;

- and very clean first wall.

Requirement of the vacuum inside the tokamak chamber during operation is very strict. Pressure inside the vessel must be below of $5 \cdot 10^{-8}$ Pa during regular operation. The fulfilment of this requirement needs powerful vacuum pumps and to seal all ports on the vacuum system properly.

Some ports were designed as diagnostic ports for optical observation of plasma inside the chamber. These ports are equipped with high quality SIO_2 glass vacuum windows.

The glass vacuum windows on diagnostic ports for optical observation were welded using electron beam welding technology which allows welding of steel to glass [2, 3] and results are high quality vacuum windows with extremely good vacuum sealing. Even the smallest displacement of the glass in this type of windows is excluded because of the required high vacuum in the reaction chamber.
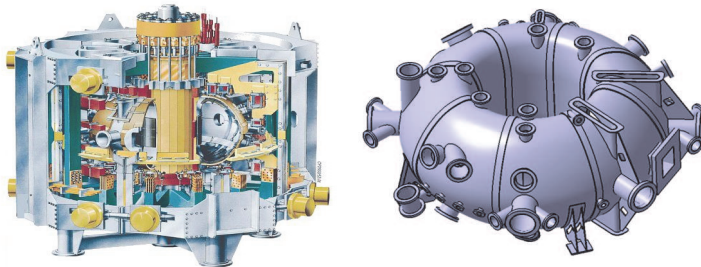

Figure 1. View of the "COMPASS" tokamak (left) and its vacuum chamber with diagnostic ports (right)

3. Mechanical model

The vacuum chamber of the *"COMPASS"* tokamak (Fig. 2.) is equipped with two types of vacuum windows of different shapes, *circular* and *oval* (Fig. 3.).
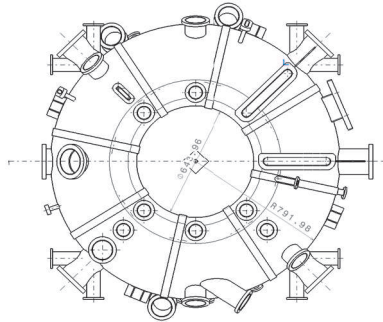

Figure 2. Locations of vacuum windows on the "COMPASS" tokamak

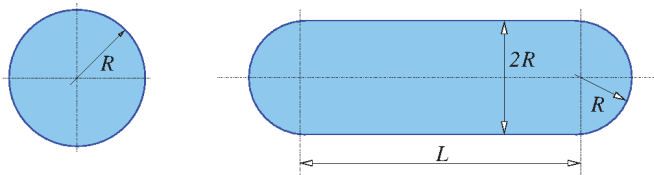The geometric model of both types of vacuum windows is shown in Fig. 3.

Figure 3. Shapes (circular and oval) of vacuum windows of the "COMPASS" tokamak

The thickness of glass in vacuum windows is $h = 8.5$ mm. Windows were made from normal SIO_2 glass with the following mechanical properties: Young's modulus is $E = 72$ GPa, Poisson's ratio is $v = 0.23$ and tensile strength is $R_m = 50$ MPa (which strongly depends on the state of surface [2, 4]). The other (geometric) parameters (R, L) of these windows are collected in Tab. 5.

The pressure outside the device is the normal atmospheric pressure ($\sim 10^5$ Pa) and inside of it is in the range of : $5 \cdot 10^{-8} \approx 0$ Pa ($\sim$ vacuum). The difference of these two pressures is the load on the outer surface of vacuum windows $\left(p = 10^5 \text{ Pa} \right)$. All windows are fixed along their perimeters. The optimal 3D and shell meshing for FE analysis has been chosen (see Chapter 5.) on the base of existing analytical solution for the window with circular shape.

4. Reference problem with analytical solution

Firstly, we remind the analytical solution for displacement and stress in the case of circular window as a reference problem [5, 6, 7].

We consider a vacuum window of circular shape (Fig. 4.), which is fixed along its perimeter, the load is a uniformly distributed (atmospheric) pressure perpendicular to the surface of the window. The target of the analysis is to calculate the radial distribution of the deformation (in direction *z*) and the equivalent (von Mises) stress.

4.1. Analytical solution for circular case

Consider the glass plate of circular shape with radius *R* and thickness *h* fixed along the perimeter. The load is an uniformly distributed pressure *p* perpendicular to the surface of the plate. The mathematical model of the problem can be easily formulated in cylindrical coordinate system which origin is in the centre of the plate, and the direction of axis *z* is perpendicular to the plane of the window (see Fig. 4.).
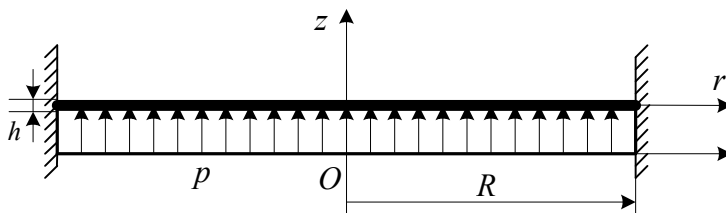
Figure 4. Mechanical model of uniformly loaded circular glass window

The problem is described with the biharmonic differential equation [5, 6, 8]:

$$\Delta^2 w(r) = \frac{p}{D} = \text{constant} , \text{ where } D = \frac{Eh^3}{12(1-\nu^2)} , \tag{1}$$

and *E* is the modulus of elasticity (Young's modulus) of glass (SIO_2) and ν is its Poisson's ratio. The radial component of equation (1) is:

$$\frac{1}{r}\frac{d}{dr}\left\{ r\frac{d}{dr}\left[\frac{1}{r}\frac{d}{dr}\left(r\frac{dw(r)}{dr} \right) \right] \right\} = \frac{p}{D} . \tag{2}$$

The analytical solution for circular shape is:

$$w(r) = \frac{p}{64D}\left(R^2 - r^2\right)^2 , \tag{3}$$

where *w (r)* is the radial distribution of the deformation in direction *z*. (see Fig. 5)

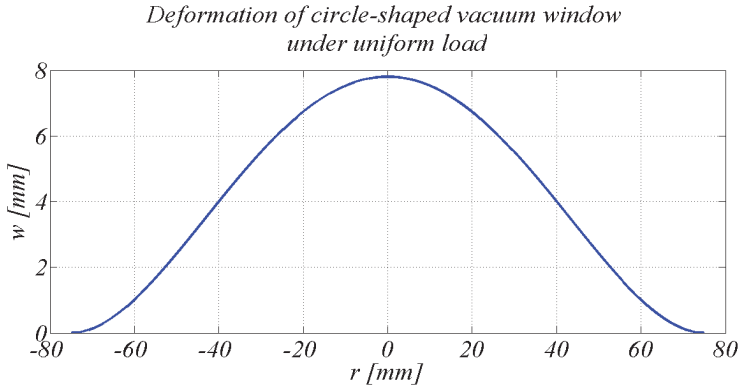*Figure 5. Analytical solution of displacement w(r) in direction z for circular window under homogeneous pressure load (R=75 mm, h=1 mm, p=10⁵ Pa)*

The calculated displacement from the formula (3) at the centre of the window is $w_{max} = 7.8040$ mm in the test case (*R=75 mm, h=1 mm*). The maximal displacement is $w_{max} = 0.0215$ mm in the real case (*R=85.575 mm, h=8.5 mm*).

4.2. Analytical solution for stresses

Radial and tangential torques M_r, M_t can be calculated from following expressions (see [8]):

$$M_r(r) = -D\left(\frac{d^2 w}{dr^2} + v\frac{1}{r}\frac{dw}{dr}\right) = D\left(\frac{d\varphi_t}{dr} + v\frac{1}{r}\varphi_t\right),\qquad(4)$$

$$M_t(r) = -D\left(v\frac{d^2 w}{dr^2} + \frac{1}{r}\frac{dw}{dr}\right) = D\left(v\frac{d\varphi_t}{dr} + \frac{1}{r}\varphi_t\right)\quad,\qquad(5)$$

where φ_t is the rotation angle around *t* axis.

From formulas (4) and (5) maximum values of torques and stresses at $z = \pm\frac{h}{2}$ can be calculated:

$$M_r(r) = \frac{p}{16}\left[R^2(1+v) - r^2(3+v)\right],\qquad \sigma_{r\,max}(r) = 6\frac{M_r(r)}{h^2},\qquad(6)$$

$$M_t(r) = \frac{p}{16}\left[R^2(1+v) - r^2(1+3v)\right],\qquad \sigma_{t\,max}(r) = 6\frac{M_t(r)}{h^2}\;.\qquad(7)$$

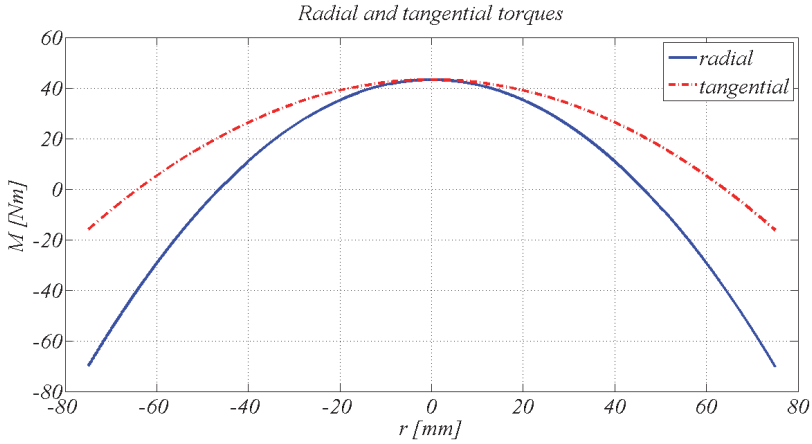In Fig. 6. the distributions of torques $M_r(r)$, $M_t(r)$ are shown:

Figure 6. Distribution of radial and tangential torques $M_r(r)$, $M_t(r)$ in the case of circular window with testing parameters R=75 mm, h=1mm

At the perimeter of the window $(r = R)$ one can get:

$$M_r\big|_{r=R} = -\frac{pR^2}{8} < 0, \qquad M_t\big|_{r=R} = -v\frac{pR^2}{8} < 0, \qquad \left|M_r\big|_{r=R}\right| > \left|M_t\big|_{r=R}\right|, \tag{8}$$

and in the centre of the window $(r = 0)$:

$$M_r\big|_{r=0} = M_t\big|_{r=0} = \frac{pR^2}{16}(1+v) > 0. \tag{9}$$

The maximum value of torque in both cases is along the fixed perimeter of the window, therefore there is the *critical location*.

The equivalent stress σ_{eq} (von Mises) is defined by the formula (10):

$$\sigma_{eq}\left(von\,Mises\right) = \sqrt{\left\{\frac{1}{2}\left[\left(\sigma_x - \sigma_y\right)^2 + \left(\sigma_y - \sigma_z\right)^2 + \left(\sigma_x - \sigma_z\right)^2 + 6\left(\tau_{xy}^2 + \tau_{yz}^2 + \tau_{xz}^2\right)\right]\right\}}, \tag{10}$$

where σ_i are normal stresses and τ_{ij} are shear stresses from the stress tensor $\underline{\underline{F}}$:

$$\underline{\underline{F}}_{xyz} = \begin{bmatrix} \sigma_x & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \sigma_y & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \sigma_z \end{bmatrix}. \tag{11}$$

The stress tensor can be transformed into the coordinate system defined by its eigenvectors (so called main axes coordinate system), and in this system the stress tensor is diagonal:

$$\underset{(e_1 e_2 e_3)}{\underset{=}{F}} = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix}. \tag{12}$$

The definition of the equivalent stress in the main axis coordinate system is:

$$\sigma_{eq}\left(von \ Mises\right) = \sqrt{\left\{\frac{1}{2}\left[\left(\sigma_1-\sigma_2\right)^2+\left(\sigma_2-\sigma_3\right)^2+\left(\sigma_1-\sigma_3\right)^2\right]\right\}}, \tag{13}$$

where for the stresses $\sigma_1, \sigma_2, \sigma_3$ we have a convention $\sigma_1 > \sigma_2 > \sigma_3$ and we call them stress maximum principal, middle principal and minimum principal.

In our case $\sigma_1 = \sigma_r$, $\sigma_2 = \sigma_t$ and $\sigma_3 = \sigma_z = 0$. Principal stresses σ_1, σ_2 can be calculated as follows:

$$\sigma_1 = \sigma_r = -\frac{6}{h^2}M_r\left(r=R\right) = -\frac{6}{h^2}\frac{p}{16}\left[R^2\left(1+v\right)-R^2\left(3+v\right)\right] = \frac{3}{4}\frac{p}{h^2}R^2, \tag{14}$$

$$\sigma_2 = \sigma_t = -\frac{6}{h^2}M_t\left(r=R\right) = -\frac{6}{h^2}\frac{p}{16}\left[R^2\left(1+v\right)-R^2\left(1+3v\right)\right] = \frac{3}{4}\frac{p}{h^2}vR^2, \tag{15}$$

and

$$\sigma_{eq\,max}\left(von\,Mises\right) = \sqrt{\frac{1}{2}\left[\left(\sigma_r-\sigma_t\right)^2+\sigma_t^2+\sigma_r^2\right]}\Big|_{r=R} = \sqrt{\left(1-v+v^2\right)}\frac{3}{4}\frac{p}{h^2}R^2. \tag{16}$$

Finally for safe operation must be valid:

$$\sigma_{eq\,max}\left(von\,Mises\right) \leq R_m. \tag{17}$$

The equivalent stress σ_{eq} reaches its maximum along the perimeter of the window.

For the glass (SiO_2) is $R_m = 50$ MPa and the maximum value of equivalent stress for real size window (R=85.575 mm, h=8.5 mm) is $\sigma_{eq\,max} = 6.256$ MPa. *This value is almost ten times less than tensile strength of SiO₂ glasses.*

5. Numerical solution

5.1. FE meshing and convergence

The numerical analysis we started with FE stress analysis [9] using *ANSYS Workbench* programme package [10, 11] for 3D FE - model, and using *ANSYS APDL* (Ansys Parameter Design Language) package for shell FE - model.

At the beginning of the numerical analysis we tested the convergence of FE solutions [12] in the case of 3D FE - model of vacuum window with circular shape (*R = 75 mm, h*

= *1 mm*) [12, 13]. As the 3D FE - type has been chosen the *quadratic tetrahedron element*. We have investigated the convergence of numerical solution to the analytic solution as a function of FE size (see Fig 7. and Fig. 8.).
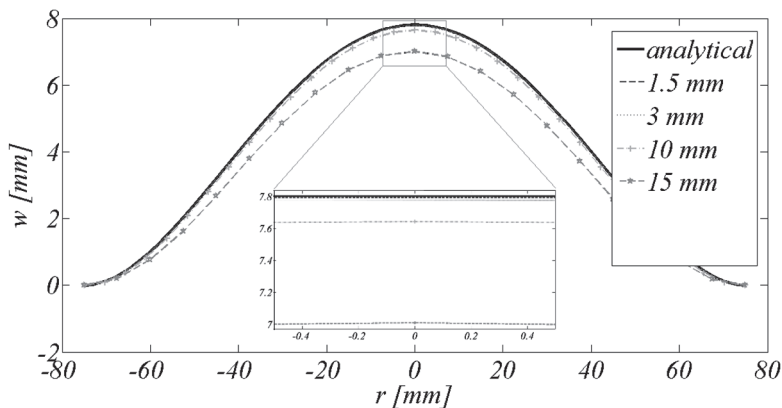

Figure 7. Convergence of FE solutions to the analytic solution in the sense of displacement for the circular test window (R=75 mm, h=1mm) in 3D model with different FE sizes

From Fig. 7. one can see, the numerical solution calculated from 3D model with decreasing element size converge to analytic solution.
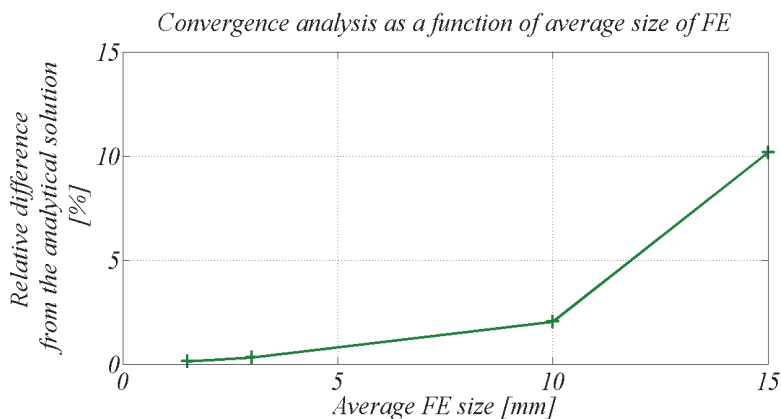
Figure 8. Relative difference between numerical and analytic solution as a function of average FE size at the radial position r=0 mm for the circular test window (R=75 mm, h=1mm) in 3D model

Fig. 8. shows the relative difference between numerical and analytic solution with decreasing FE size rapidly converge to zero value, thus it is reasonable to use this FE - type and FE size also in the cases of oval windows without analytic solutions.

5.2. Results of FE analysis for the case of test window

In this subchapter we will present numerical results for the case of the test window with circular shape and comparison of numerical results with analytic solutions.

First of all, we need to note our problem is cylindrically symmetric therefore we present in all figures just the quarter of the window for the 3D model, and in the shell model it is enough to consider only the radial intersection of the circular plate.

In Fig. 9. one can see meshes with two different FE sizes. Fig.10. shows the distribution of displacement w in direction z and the distribution of equivalent stress σ_{eq} calculated from 3D model. Same two quantities (w and σ_{eq}) calculated from shell model are shown in Fig. 11.
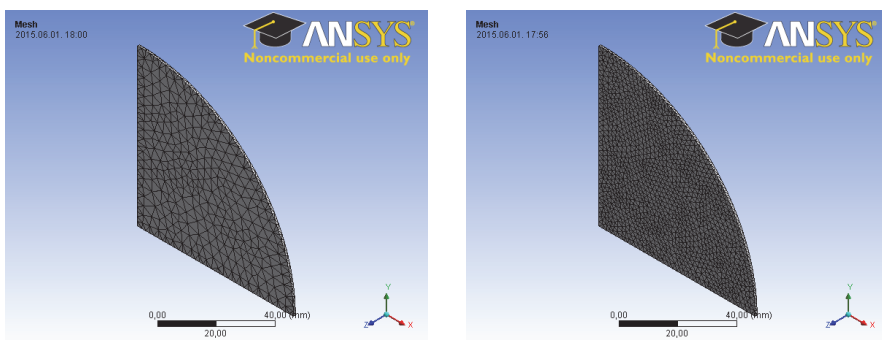

Figure 9. Two different meshes in 3D model for circular window with test parameters R=75 mm, h=1 mm and with FE sizes 4 mm (left) and 2 mm (right).

Table 1. Number of nodes and elements of two applied meshing in 3D model for the meshes with bigger and smaller FE sizes for the test window of circular shape (R= 75 mm and h=1 mm)

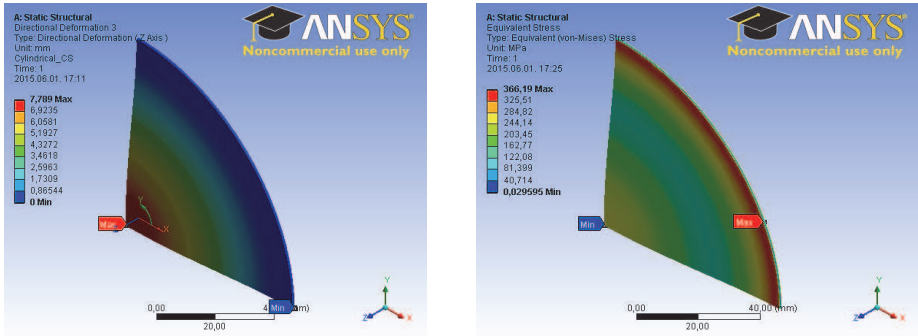|	Element size $[\text{mm}]$	Number of Nodes	Number of Elements
Mesh 1	4	4336	1975
Mesh 2	2	14692	7202

Figure 10. Deformation w in direction z (left) and equivalent stress σ_{eq} (right) from 3D model for circular test window (R=75 mm, h=1 mm) with average FE size 2mm.
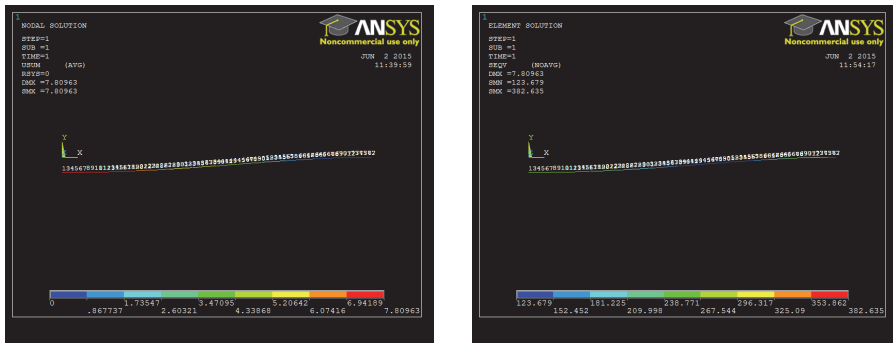

Figure 11. Deformation w in direction z (left) and equivalent stress σ_{eq} (right) calculated from shell model for circular window with testing parameters R =75 mm, h =1 mm with FE size 2mm.

Table 2. Comparison of results (displacements and equivalent stresses) for the circular test window (R=75 mm, h=1 mm) with average FE size 2 mm

w_{max} [mm]			$\sigma_{eq\,max}$ [MPa]		
Analytic solution	3D model	Shell model	Analytic solution	3D model	Shell model
7.804	7.789	7.809	347.161	366.19	382.635

In Tab. 2. one can see that differences between three different solutions (analytic solution, 3D model and shell model) for displacement are small, thus our model for displacement is appropriate, but in cases of equivalent stresses the model is less precise.

5.3. Results of FE analysis for the case of real window sizes

This chapter presents results for the case of real window sizes for both models (3D and shell) and for both shapes of vacuum windows (circular and oval) of the *"COMPASS"* tokamak. The stress study of all types of vacuum windows on *"COMPASS"* tokamak has been provided using meshing with FE types of *quadratic tetrahedron* in 3D model and FE type *shell 209* in shell model with FE sizes equal 4.25 mm. All windows have equal thicknesses *h*=8.5 mm.

5.3.1. The 3D model

5.3.1.1. The case of circular window

Fig. 12. shows FE meshes from 3D model for the circular vacuum window with radius *R = 85.575 mm* and with thickness *h = 8.5 mm*:
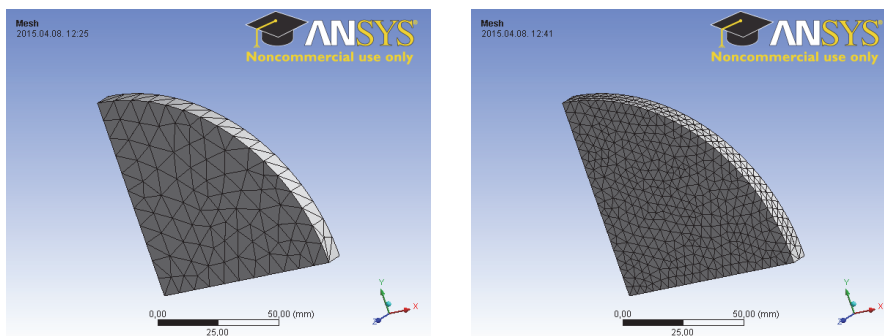

Figure 12. Meshing from 3D model for the case of circular window using average FE size 8.5 mm (left) and 4.25 mm (right) with quadratic tetrahedron elements

Table 3. Number of nodes and elements of applied meshing from 3D model for the meshes with bigger and smaller FE sizes for the vacuum window of circular shape (R= 85.575 mm, h = 8.5 mm)

	Element size [mm]	Number of Nodes	Number of Elements
Mesh 1	8.5	1598	856
Mesh 2	4.25	9934	6060

In Fig 13. one can see results of FE analysis (deformation and stress) for the vacuum window of circular shape calculated from 3D model.

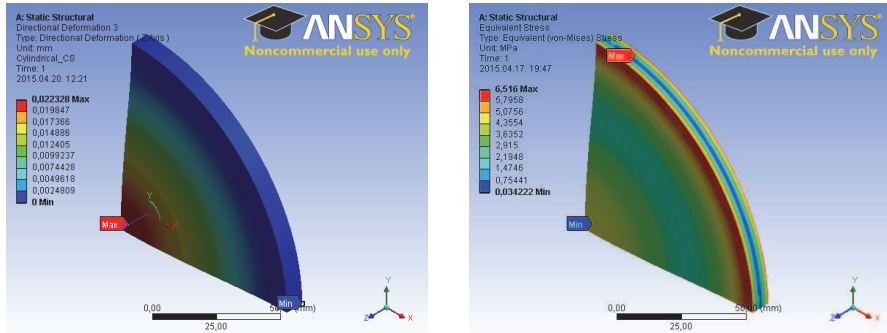*Figure 13. Results of FE analysis (the deformation in the direction of axis z (left) and the equivalent stress (right)) calculated from 3D model for circular vacuum window (R = 85.575 mm, h=8.5mm) with FE size 4.25mm*

Both of investigated physical quantities (deformation and equivalent stress) for the circular vacuum window calculated from 3D model meet very well with analytic results.

5.3.1.2. The case of oval window

Consider the oval window with geometric parameters R = 34.5 mm, L = 389 mm and the thickness h = 8.5 mm.

Fig. 14. shows the mesh of the model. We used the same mesh as in the case of circular window and the quadratic tetrahedron elements with average element size of 8.5 mm.
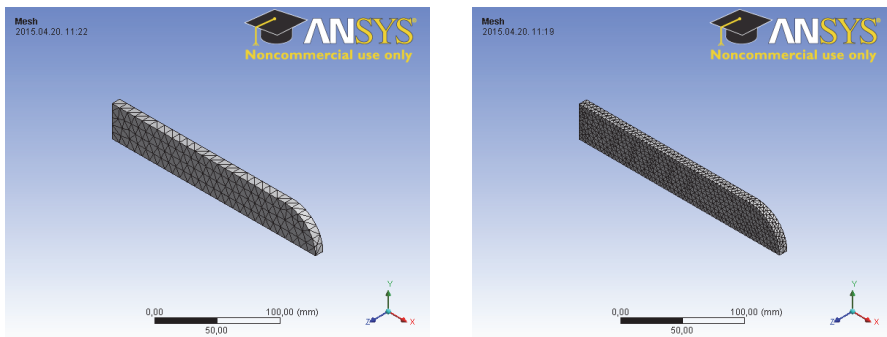

Figure 14. Meshing from 3D model for the case of oval window using average FE size 8.5 mm (left) and 4.25 mm (right) with quadratic tetrahedron elements

Table 4. Two different meshes of 3D model with bigger and smaller FE sizes for the biggest window of oval shape with parameters L = 389.0 mm and R = 34.5 mm

Element size $[\text{mm}]$	Number of Nodes	Number of Elements
8.5	2168	1166
4.25	12130	7383

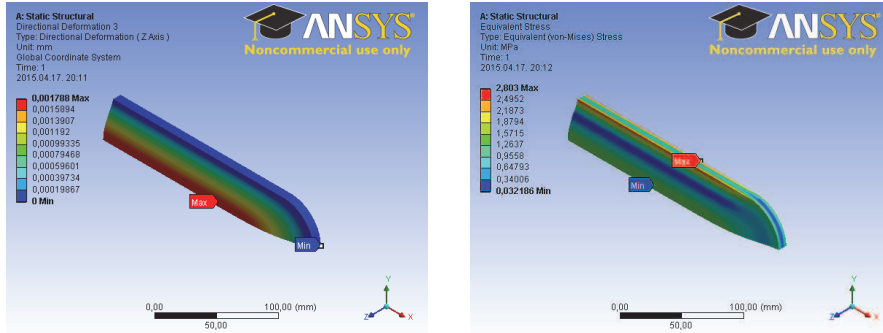*Figure 15. Results of FE analysis (the deformation in the direction of axis z (left) and the equivalent stress (right)) from 3D model for oval window with L=389.0 mm and R = 34.5 mm*

One can see from Fig. 15. the maximum of deformation is in the middle of the window.

The result of stress analysis for oval window is shown in Fig. 15. We can see that the equivalent stress again has its maximum at the perimeter of the window, where the window is fixed, but its distribution is not symmetric. Along the straight part of the perimeter it is significantly higher in comparison with equivalent stress along the round part of the perimeter.

5.3.1.3. Summary of results of FE analysis from 3D model

In this subchapter we will summarize the results of FE analysis calculated from 3D model for both shapes of windows of real sizes.

Table 5. Summary table of FE analysis calculated from 3D model with real geometric parameters of windows (real window sizes)

Shape	Size [mm]		w_{max} [µm]	$\sigma_{eq.max}$ [MPa]
Round	$R_1 = 24.175$		0.204	0.676
	$R_2 = 41.275$		1.357	1.634
	$R_3 = 60.350$		5.731	3.456
	$R_4 = 85.575$		22.328	6.516
Oval	$L_1 = 90.0$	$R_1 = 17.25$	0.162	0.784
	$L_2 = 123.5$	$R_2 = 17.25$	0.163	0.809
	$L_3 = 389.0$	$R_3 = 34.5$	1.788	2.803

In Tab. 5. one can see the maximum of displacement in direction *z* and the maximum of equivalent stress (for both shapes of window and for cases of different radii *R* and length *L*) calculated from 3D model.

Deformations in the direction of axis z and stresses are small. Equivalent stresses, which are the main design parameters for these vacuum windows, are safely below of tensile strength for SiO₂ glass.

5.3.2. The shell model

One dimension - the thickness of windows - is significantly smaller than other two dimensions, therefore we decided about the FE analysis with *shell elements* for comparison with 3D results. Appropriate FE shell element is included in *ANSYS APDL,* thus we provided FE analysis in this part of *ANSYS* programme package.

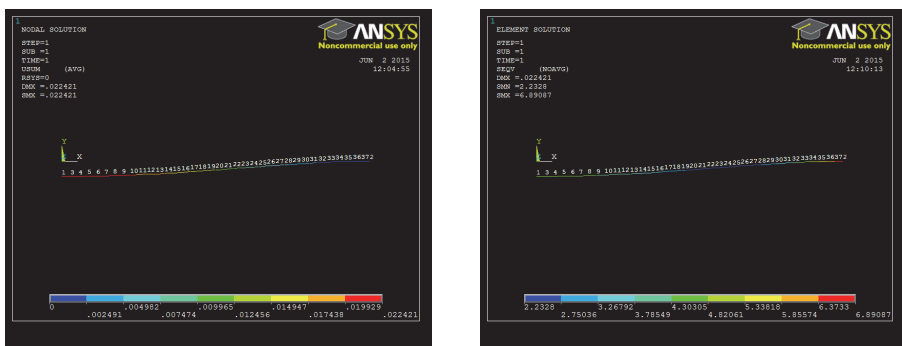Results of this analysis is summarized in Fig. 16.

Figure 16. Deformation w in direction z (left) and the equivalent stress σ_{eq} (right) of shell model for circular window with real size parameters R = 85.575 mm and h= 8.5 mm

One can see from Fig. 16. and also from Tab. 6 given below, the maximum value of the deformation is again at the centre of the window, and its value is near to values from analytic calculations.

The equivalent stress has its maximum again along the perimeter of the window and one can see, the 3D model and the shell model give basically the same results.

Table 6. Summary table of FE analysis calculated from shell model for real thickness of windows h = 8.5 mm and with FE size 4.25 mm

Shape	Size [mm]	w_{max} [μm]	$\sigma_{eq.max}$ [MPa]
Round	$R_1 = 24.175$	0.208	0.547
	$R_2 = 41.275$	1.371	1.601
	$R_3 = 60.350$	5.767	3.426
	$R_4 = 85.575$	22.421	6.890

Our analysis showed, the 3D model and shell model are practically equivalent for stress analysis of vacuum windows of the *"COMPASS"* tokamak.

6. Summary and Conclusions

Our task - agreed with operators of the "COMPASS" tokamak - was to perform detailed FE stress analysis for all types of vacuum windows of the device.

In the case of windows with circular shape fortunately exists an analytic solution. This situation gave us the possibility to test properties of numerical solutions from FE analysis (convergence, element type, element size).

It was reasonable to use the same FE parameters also for oval windows.

Detailed analysis showed, deformations are small and equivalent stresses are safely below of tensile strength for SiO_2 glasses.

From the FE analysis point of view it is interesting, practically no difference between results from 3D and shell model.

Table 7. Summary table: comparison of results (maximum of displacements in direction z and maximum of equivalent stresses) of analytic solution, the 3D and shell model for the circular window (R = 85.575 mm, h = 8.5 mm, FE size 4.25 mm)

w_{max} [mm]			$\sigma_{eq.max}$ [MPa]		
Analytical solution	3D model	Shell model	Analytic solution	3D model	Shell model
0.0215	0.0223	0.0224	6.256	6.5163	6.891

Acknowledgement

References

[1] Pánek R et. al.: Reinstallation of the COMPASS-D tokamak in IPP ASCR. Czechoslovak Journal of Physics, Vol. 56, Issue 2 Supplement, pp B125-B137, 2006. DOI: 10.1007/s10582-006-0188-1

[2] Le Bourhis E: Glass - Mechanics and Technology. WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, 2006.

[3] Sun Z, Karppi R: The application of electron beam welding for the joining of dissimilar metals: an overview. Journal of Materials Processing Technology, Vol. 59, Issue 3, pp. 257–267. DOI: 10.1016/0924-0136(95)02150-7

[4] Renlund GM, Prochazka S, Doremus RH: Silicon oxycarbide glasses: Part II. Structure and properties. Journal of Materials Research, Vol. 6, Issue 12, pp 2723-2734, 1991. DOI: 10.1557/JMR.1991.2723

[5] Landau LD, Lifshitz EM: Theory of Elasticity. Butterworth-Heinemann, 1986.

[6] Brenner SC, Ridgway SL: The Mathematical Theory of Finite Element Methods. Third Edition, Springer, 2008.

[7] Zienkievicz OC, Taylor RL: The Finite Element Method. Volume 2: Solid Mechanics. Fifth Edition, Butterworth-Heinemann, 2000.

[8]   Timoshenko S, Woinowsky-Krieger S: Theory of plates and shells. McGraw-Hill 1959.

[9]   Fries TP, Belytschko T: The extended / generalized finite element method: An overview of the method and its applications. Int. J. Numer. Meth. Engng, 00:1-6, 2000.

[10] Nakasone Y, Yoshimoto S, Stolarski TA: Engineering Analysis With ANSYS Software, Elsevier Butterworth-Heinemann, 2006.

[11] ANSYS Workbench User's Guide. Ansys Inc., Release 12.1, November 2009.

[12] Hughes TJR: The Finite Element Method. Linear Static and Dynamic Finite Element Analysis, Prentice-Hall, Englewood Cliffs, New Jersey 07632, 1987.

[13] Logan DL: A First Course in the Finite Element Method. Fourth Edition, Thomson Canada, 2007.

Overview of the Finite Element Method

M. Kuczmann

Széchenyi István University, Department of Automation
Egyetem tér 1, H9026, Győr, Hungary
E-mail: kuczmann@sze.hu

Abstract: By using scalar and vector potentials, Maxwell's equations can be transformed into partial differential equations. Generally, the partial differential equations can be solved by numerical methods. One of these numerical methods is the finite element method, which is based on the weak formulation of the partial differential equations. The basis of numerical techniques is to reduce the partial differential equations to algebraic ones whose solution gives an approximation of the unknown potentials and electromagnetic field quantities. This reduction can be done by discretizing the partial differential equations in time if necessary and in space. The potential functions, the approximation method and the generated mesh distinguish the numerical field solvers. This paper summarize the finite element method as a CAD technique in electrical engineering to obtain the electromagnetic field quantities in the case of static magnetic field and eddy current field problems. Here, we show how to discretize the analyzed domain with finite elements, how to approximate potential functions with nodal and vector shape functions and how to build up the system of equations, which solution obtain the unknown potentials.

Keywords: Maxwell's equations, weak formulation, finite element method

1.   Introduction

This paper is based on the book [1].

The Finite Element Method (FEM) is the most popular and the most flexible numerical technique to determine the approximate solution of the partial differential equations in engineering. For example, commercially available FEM software package is COMSOL Multiphysics, which is able to solve one, two and three-dimensional problems. A free mesh generator with a built-in CAD engine and post-processor is Gmsh.

The main steps of simulation with FEM are illustrated in Fig. 1. Firstly, in the model specification phase, the model of the real life problem, which simulation require electromagnetic field calculations must be set up, i.e. we have to find out the partial differential
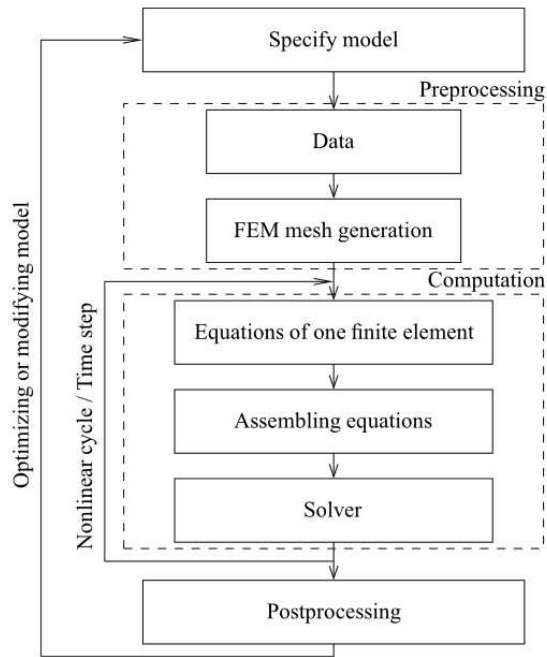
Figure 1. Steps of simulation by FEM.

equations, which must be solved with prescribed boundary and continuity conditions. We have to find out, whether it is a linear or a nonlinear problem and how the characteristics look like. After selecting potentials, the weak formulation of these partial differential equations must be worked out as well. It is depending on the problem, of course, but the chosen mathematical model of the arrangement should be adequate to calculate electromagnetic field quantities in the given accuracy. The geometry of the problem must be defined by a CAD software tool, e.g. by using a user friendly interface, see e.g. Fig. 2.

The next step is the preprocessing task. Here we have to give the values of different parameters, such as the material properties, i.e. the constitutive relations, the excitation signal and the others. The geometry can be simplified according to symmetries or axial symmetries.

The geometry of the problem must be discretized by a FEM mesh. The fundamental idea of FEM is to divide the problem region to be analyzed into smaller finite elements with given shape. A finite element can be e.g. triangle or quadrangle in 2D, e.g. tetrahedron or hexahedra in 3D. A triangle has three vertices 1, 2 and 3 numbered counter-clockwise and has 3 edges. The quadrangle element has 4 nodes and 4 edges. A tetrahedral element has 4
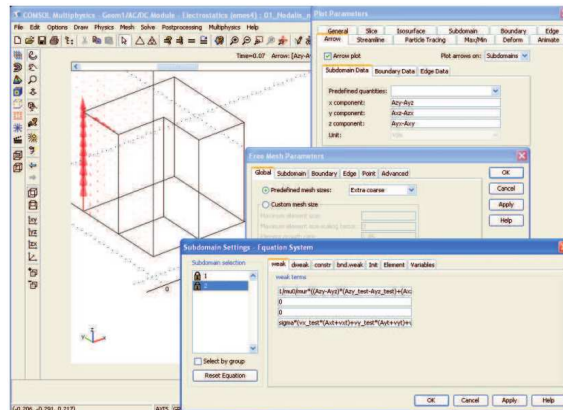
Figure 2. COMSOL Multiphysics, a CAD environment to solve electromagnetic field problems.

vertices and 6 edges and a hexahedral element has 8 nodes and 12 edges.

There are some simple rules, how to generate a mesh. Neither overlapping nor holes are allowed in the generated finite element mesh. If material interface are included in the problem region, the configuration of mesh must be adapted to these boundaries, i.e. interfaces coincide with finite element interfaces.

FEM mesh, as two illustrations, generated by COMSOL Multiphysics can be seen in Fig. 3 and in Fig. 4. The first 2D illustration (Fig. 3) shows the mesh of a horseshoe-shaped permanent magnet. The two ends are pre-magnetized in different directions. The second illustration (Fig. 4) presents a model of a micro-scale square inductor, used for LC bandpass filters in microelectromechanical systems. The model geometry consists of the spiral-shaped inductor and the air surrounding it (the mesh in air is not shown). The outer dimensions of the model geometry are around 0.3 mm. These illustrations are cited from the Model Documentation of COMSOL Multiphysics.

The next step in FEM simulations is solving the problem. The FEM equations, based on the weak formulations, must be set up in the level of one finite element, then these equations must be assembled through the FEM mesh. Assembling means that the global system of equations is built up, which solution is the approximation of the introduced potential. The obtained global system of algebraic equations is linear, or nonlinear but linearized, depending on the medium to be analyzed. Then this global system of equations must be solved by a solver. The computation may contain iteration if the constitutive equations are nonlinear. This is the situation when simulating ferromagnetic materials with nonlinear characteristics. Iteration means that the system of equations must be set up and
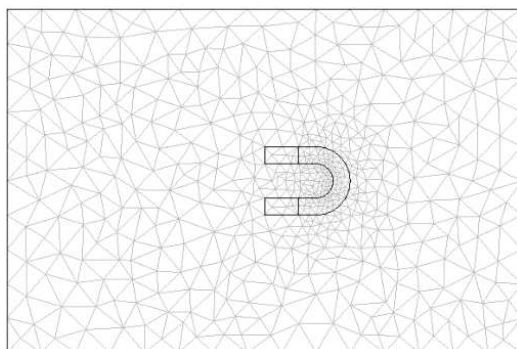
Figure 3. COMSOL model of a permanent magnet, geometry is meshed by triangles.
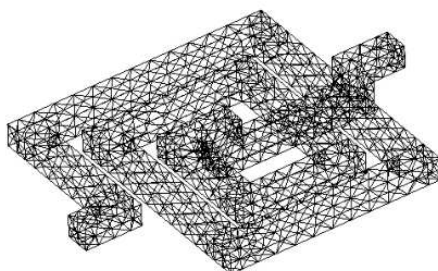

Figure 4. COMSOL model of a micro-scale square inductor, geometry is discretized by tetrahedral shape finite elements.
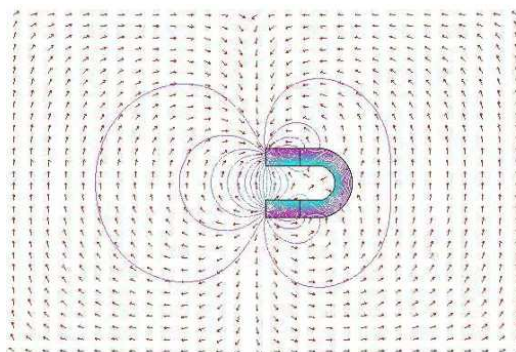

Figure 5. COMSOL solution of the static magnetic field generated by a permanent magnet.
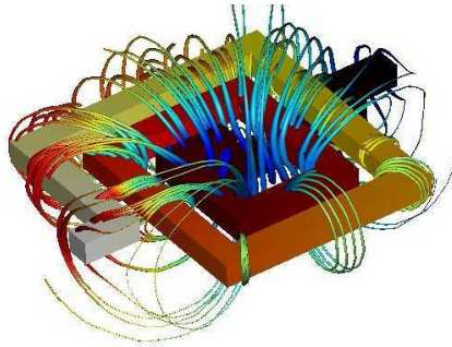
Figure 6. Electric potential in the device and magnetic flux lines around the device, the problem has been solved by COMSOL.

must be solved step by step until convergence is reached. If the problem is time dependent, then the solution must be worked out at every discrete time instant.

The result of computations is the approximated potential value in the FEM mesh. Any electromagnetic field quantity (e.g. magnetic field intensity, or magnetic flux density, etc.) can be calculated by using the potentials at the postprocessing stage. Capacitance, inductance, energy, force and other quantities can also be calculated. The postprocessing give a chance to modify the geometry, the material parameters or the FEM mesh to get more accurate result. The COMSOL Multiphysics has been used to show two examples about postprocessing. The pattern of the magnetic field around the permanent magnet is well known through experiments (see Fig. 5). Figure 6 shows the electric potential in the inductor and the magnetic flux lines. The thickness of the flow lines represents the magnitude of the magnetic flux.

2. Approximating potentials with shape functions

The potential function can be scalar valued (e.g. the magnetic scalar potential Φ, or the electric scalar potential V), or vector valued (e.g. the current vector potential $\boldsymbol{T}$, or the magnetic vector potential $\boldsymbol{A}$).

The scalar potential functions can be approximated by so-called nodal shape functions and the vector potential functions can be approximated by either nodal or so-called vector shape functions, also called edge shape functions. Generally, a shape function is a simple continuous polynomial function defined in a finite element and it is depending on the type of the used finite element.

Shape functions have the following general properties:

(*i*) Each shape function is defined in the entire problem region;

(*ii*) Each scalar shape function corresponds to just one nodal point and each vector shape function corresponds to just one edge;

(*iii*) Each scalar shape function is nonzero over just those finite elements that contain its nodal point and equals to zero over all other elements. Each vector shape function in nonzero over just those finite elements that contain its edge and equals to zero over all other elements;

(*iv*) The scalar shape function has a value unity at its nodal point and zero at all other nodal points. The line integral of a vector shape function is equal to one along its edge and the line integral of it is equal to zero along the other edges;

(*v*) The shape functions are linearly independent, i.e. no shape function equals a linear combination of the other shape functions.

The accuracy of solution obtained by FEM can be increased in three ways. The first one is increasing the number of finite elements, i.e. decreasing the element size. It is called h-FEM. The second way is to increase the degree of polynomials building up a shape function (e.g. using Lagrange or Legendre interpolation functions). This is the so-called p-FEM. The mixture of these methods results in hp-FEM. Potentials approximated by h-version or p-version are assigned in the indices of the potentials.

2.1. Nodal finite elements

Scalar potential functions can be represented by a linear combination of shape functions associated with nodes of the finite element mesh. Within a finite element, a scalar potential function $\Phi = \Phi(\boldsymbol{r},t)$ is approximated by

$$\Phi \simeq \sum_{i=1}^{m} N_i \Phi_i, \tag{1}$$

where $N_i = N_i(\boldsymbol{r})$ and $\Phi_i = \Phi_i(t)$ are the nodal shape functions and the value of potential function corresponding to the ith node, respectively. The number of degrees of freedom is $m = 2$ in 1D problems, $m = 3$ in a 2D problem using triangular FEM mesh and $m = 4$ in a 3D arrangement meshed by tetrahedral elements and the shape functions are linear. The nodal shape functions can be defined by the relation

$$N_i = \begin{cases} 1, & \text{at the node } i, \\ 0, & \text{at other nodes.} \end{cases} \tag{2}$$

(*i*) In 1D, the linear shape functions can be build up by

$$N_1 = \frac{x_2 - x}{x_2 - x_1}, \quad \text{and} \quad N_2 = \frac{x - x_1}{x_2 - x_1}, \tag{3}$$

where x_1 and x_2 are the coordinates of the boundaries of one finite element. The linear shape functions are plotted in Fig. 7. It is easy to control the equation (2).
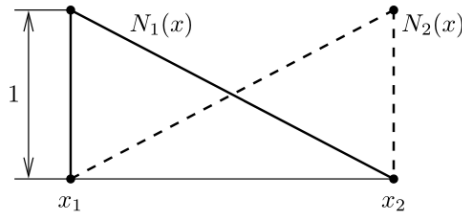

Figure 7. The 1D linear shape functions $N_1(x)$ and $N_2(x)$.

If the values of the potential are known in the two boundary points x_1 and x_2, then the potential can be determined easily inside the finite element $x_1 \leq x \leq x_2$ as (see Fig. 8)

$$\Phi = N_1 \, \Phi_1 + N_2 \, \Phi_2 = \frac{x_2 - x}{x_2 - x_1} \Phi_1 + \frac{x - x_1}{x_2 - x_1} \Phi_2. \tag{4}$$

Of course, it is valid in the other finite elements as well, e.g. if $x_2 \leq x \leq x_3$, then

$$\Phi = N_1 \, \Phi_2 + N_2 \, \Phi_3 = \frac{x_3 - x}{x_3 - x_2} \Phi_2 + \frac{x - x_2}{x_3 - x_2} \Phi_3, \tag{5}$$

and N_1, N_2 are shifted to the second finite element.

The scalar potential is continuous in the whole 1D region. It is noted here that the accuracy of approximation can be increased by decreasing the length of the elements,
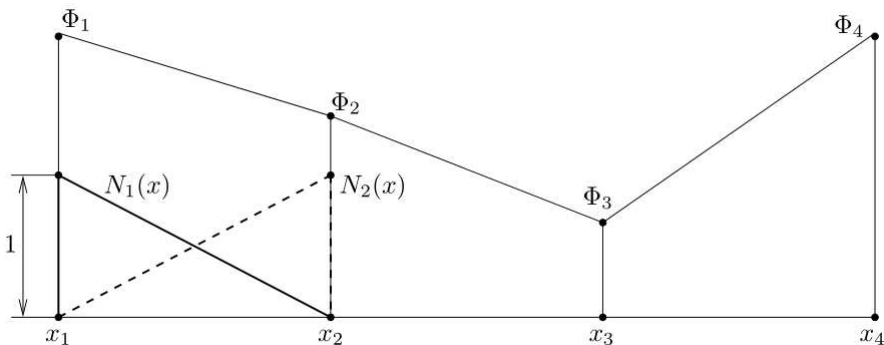

Figure 8. Known potential values are approximated by linear functions.

especially where the rate of change of the solution is large, e.g. between x_3 and x_4 in Fig. 8. Here, the mesh can be very fine and higher order approximation can results in better solution.

One way to build up higher order shape functions is using Lagrange interpolation functions, defined by the formula

$$N_i(x) = \prod_{j=1,\, j\neq i}^{m} \frac{x - x_j}{x_i - x_j}. \tag{6}$$

The order is $m-1$ and $N_i(x)$ is equal to one in the node i and equal to zero in all the other nodes. Here, second and third order approximations are shown.

The second order approximation can be defined by 3 quadratic shape functions (i.e. $m = 3$ in (1), see Fig. 9),

$$N_1 = \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)}, \tag{7}$$

$$N_2 = \frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_2 - x_3)}, \tag{8}$$

$$N_3 = \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)}, \tag{9}$$

and the new point x_3 is placed in the center of the element,

$$x_3 = \frac{x_1 + x_2}{2}. \tag{10}$$

The third order approximation can be defined by 4 cubic shape functions ($m = 4$ in (1), see Fig. 10),

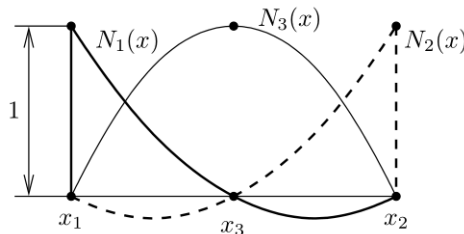$$N_1 = \frac{(x - x_2)(x - x_3)(x - x_4)}{(x_1 - x_2)(x_1 - x_3)(x_1 - x_4)}, \tag{11}$$

Figure 9. The 1D quadratic shape functions $N_1(x)$, $N_2(x)$ and $N_3(x)$.

$$N_2 = \frac{(x - x_1)(x - x_3)(x - x_4)}{(x_2 - x_1)(x_2 - x_3)(x_2 - x_4)}, \tag{12}$$

$$N_3 = \frac{(x - x_1)(x - x_2)(x - x_4)}{(x_3 - x_1)(x_3 - x_2)(x_3 - x_4)}, \tag{13}$$

$$N_4 = \frac{(x - x_1)(x - x_2)(x - x_3)}{(x_4 - x_1)(x_4 - x_2)(x_4 - x_3)}, \tag{14}$$

and the new points x_3 and x_4 are placed inside the element as

$$x_3 = \frac{1(x_1 + x_2)}{3}, \quad x_4 = \frac{2(x_1 + x_2)}{3}. \tag{15}$$
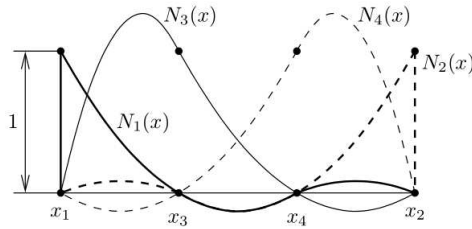
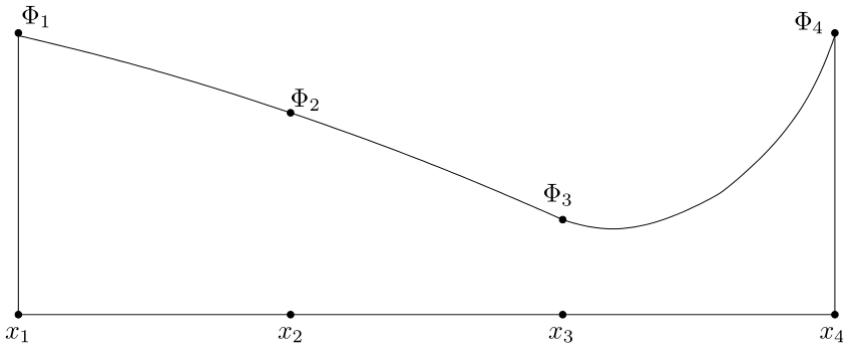*Figure 10. The 1D cubic shape functions $N_1(x)$, $N_2(x)$, $N_3(x)$ and $N_4(x)$.*

Figure 11. Known potential values are approximated by quadratic functions.

With this technique, the interpolation functions of any order can be defined and the equation (2) can be controlled.

Figure 11 shows the higher order approximation of the potential plotted in Fig. 8. This illustration shows the applicability of higher order functions.

(***ii***) 2D linear shape functions can be built up as follows when using a finite element mesh with triangular finite elements. Linear basis functions can be introduced by using the so-called barycentric coordinate system in a triangle as follows. The area of a triangle is denoted by Δ and it can be calculated as

$$\Delta = \frac{1}{2} \begin{vmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{vmatrix}, \tag{16}$$

where (x_1,y_1), (x_2,y_2) and (x_3,y_3) are the coordinates of the three nodes of the triangle in the global coordinate system building an anticlockwise sequence. The area functions (see Fig. 12) of a given point inside the triangle with coordinates (x,y) can be calculated as

$$\Delta_1 = \frac{1}{2} \begin{vmatrix} 1 & x & y \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{vmatrix}, \Delta_2 = \frac{1}{2} \begin{vmatrix} 1 & x_1 & y_1 \\ 1 & x & y \\ 1 & x_3 & y_3 \end{vmatrix}, \Delta_3 = \frac{1}{2} \begin{vmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x & y \end{vmatrix}, \tag{17}$$

i.e. $\Delta_1 = \Delta_1(x,y)$, $\Delta_2 = \Delta_2(x,y)$ and $\Delta_3 = \Delta_3(x,y)$ are depending on the coordinates x and y.

The barycentric coordinates $L_i = L_i(x,y)$ can be defined by the above area functions as

$$L_i = \frac{\Delta_i}{\Delta}, \qquad i = 1,2,3. \tag{18}$$

Three linear shape functions $N_i = N_i(x,y)$ can be described as

$$N_i = L_i, \qquad i = 1,2,3. \tag{19}$$

The shape function N_i is equal to 1 at the ith node of the triangle and it is equal to zero at the other two nodes, because Δ_i is equal to Δ at node i and it is equal to zero at the other two nodes. That is why the relation (2) is satisfied. It is obvious that the three shape functions are linearly independent.

The linear shape functions N_i ($i = 1,2,3$) vary linearly over the triangle, because the fraction Δ_i/Δ measures the perpendicular distance of the point (x,y) toward the vertex opposite to node i as it is illustrated in Fig. 13 and the linear shape function is constant along such a line. The three linear shape functions are shown in Fig. 14.

If the potential at the nodes is known, then a linear approximation of the potential function can be represented by (1). The derivative of a first order approximation is zeroth order, i.e. constant. The magnetic field intensity $\boldsymbol{H}$, or the magnetic flux density $\boldsymbol{B}$ are constant within a triangle, if these are obtained from a first order approximation by $\boldsymbol{H} = -\nabla\Phi$, or $\boldsymbol{B} = \nabla \times \boldsymbol{A}$. This may results in inaccurate solution. This is the reason why higher order approximations are studied. Here, only the second and the third order approximations are shown.
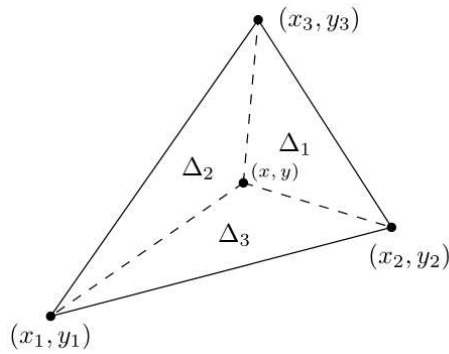
Figure 12. The area function of a triangle.
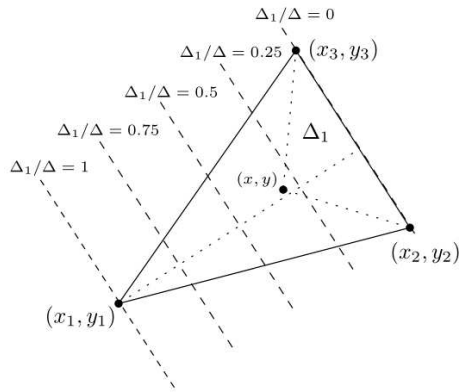
Figure 13. Fraction Δ_i/Δ measures the perpendicular distance of the point (x,y) toward the vertex opposite to node i (here $i = 1$).

Higher order shape functions can also be built up by using the barycentric coordinates L_1, L_2 and L_3 introduced above in (18).

(a) $N_1(x,y)$

(b) $N_2(x,y)$

(c) $N_3(x,y)$
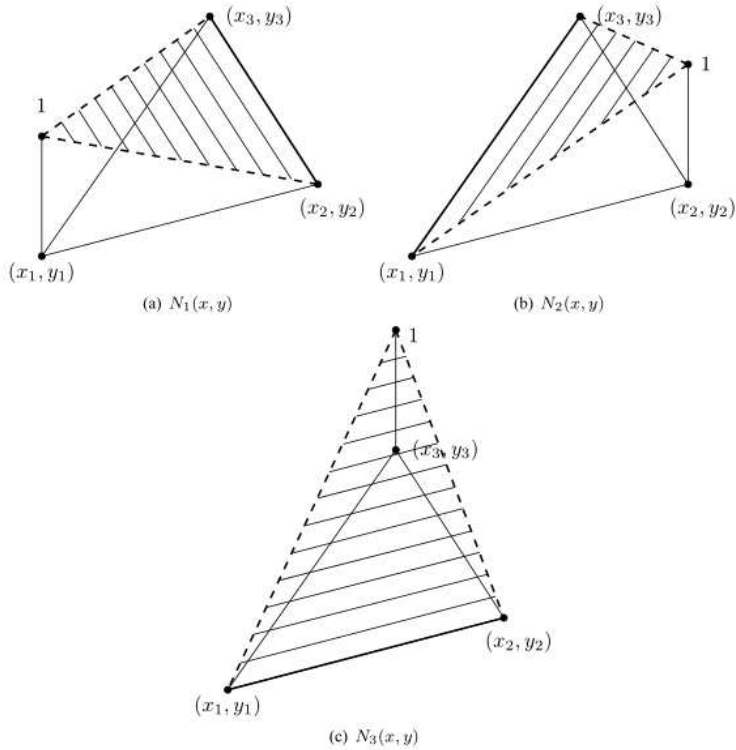
Figure 14. The 2D linear shape functions $N_1(x,y)$, $N_2(x,y)$ and $N_3(x,y)$.

A polynomial of order n must contain all possible terms $x^p\, y^q$, $0 \le p + q \le n$, as it is presented by Pascal's triangle,

$$1$$
$$x \quad y$$
$$x^2 \quad xy \quad y^2$$
$$x^3 \quad x^2y \quad xy^2 \quad y^3 \qquad \cdots$$

The first row contains the only one term of the zeroth order polynomials, the second, third and fourth rows contain the terms of the first, second and third order polynomials. Pascal's triangle can be used to generate the elements of a polynomial with given order. Such a polynomial contains

$$m = \frac{(n+1)(n+2)}{2} \tag{20}$$

elements altogether, i.e. $m = 1$, $m = 3$, $m = 6$ and $m = 10$ in the case of zeroth, first, second and third order polynomials. It means that m coefficients must be expressed, finally m points must be placed within a triangle. Pascal's triangle can be continued, of course.

The interpolation function of order n can be constructed as

$$N_i = P_I^n(L_1)\, P_J^n(L_2)\, P_K^n(L_3), \quad \text{where} \quad I + J + K = n, \tag{21}$$

and the integers I, J and K label the nodes within the triangle, resulting in a numbering scheme. Figure 15, Fig. 16 and Fig. 17 illustrate the numbering scheme of the first, the second and the third order approximations. It is noted that points must be inserted not only the edges, but inside the triangle, if $n > 2$.
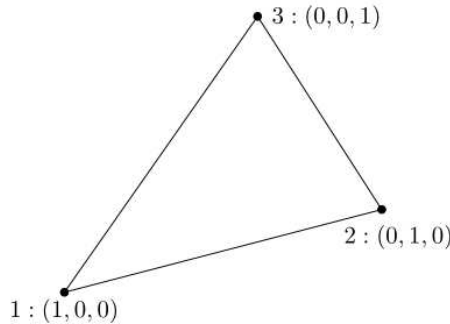
Figure 15. Numbering scheme for linear element, $n = 1$.

The polynomials $P_I^n(L_1)$, $P_J^n(L_2)$ and $P_K^n(L_3)$ are defined as

$$P_I^n(L_1) = \prod_{p=0}^{I-1} \frac{n\,L_1 - p}{I - p} = \frac{1}{I!} \prod_{p=0}^{I-1} (n\,L_1 - p), \quad \text{if} \quad I > 0, \tag{22}$$

$$P_J^n(L_2) = \prod_{p=0}^{J-1} \frac{n\,L_2 - p}{J - p} = \frac{1}{J!} \prod_{p=0}^{J-1} (n\,L_2 - p), \quad \text{if} \quad J > 0, \tag{23}$$

$$P_K^n(L_3) = \prod_{p=0}^{K-1} \frac{n\,L_3 - p}{K - p} = \frac{1}{K!} \prod_{p=0}^{K-1} (n\,L_3 - p), \quad \text{if} \quad K > 0, \tag{24}$$

and as a definition

$$P_0^n = 1. \tag{25}$$

If $n = 1$, then $m = 3$, i.e. (see Fig. 15)

$$N_1 = P_1^1(L_1)\, P_0^1(L_2)\, P_0^1(L_3) = L_1, \tag{26}$$

$$N_2 = P_0^1(L_1)\, P_1^1(L_2)\, P_0^1(L_3) = L_2, \tag{27}$$

$$N_3 = P_0^1(L_1)\, P_0^1(L_2)\, P_1^1(L_3) = L_3, \tag{28}$$

since

$$P_1^1(L_i) = \prod_{p=0}^{1-1} \frac{1}{1-p}\frac{L_i - p}{1} = L_i, \tag{29}$$

as it was mentioned in (19).
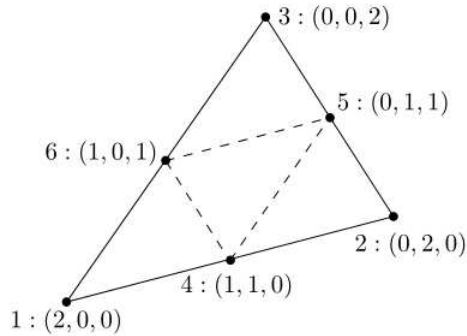
Figure 16. Numbering scheme for quadratic element, $n = 2$.
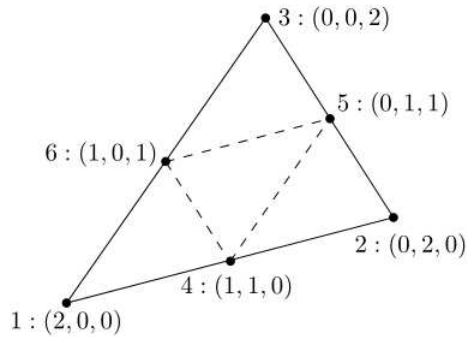
Figure 17. Numbering scheme for cubic element, $n = 3$.

If $n = 2$, then $m = 6$, i.e. (see Fig. 16)

$$N_1 = P_2^2(L_1)\, P_0^2(L_2)\, P_0^2(L_3) = L_1(2\, L_1 - 1), \tag{30}$$

$$N_2 = P_0^2(L_1)\, P_2^2(L_2)\, P_0^2(L_3) = L_2(2\, L_2 - 1), \tag{31}$$

$$N_3 = P_0^2(L_1)\, P_0^2(L_2)\, P_2^2(L_3) = L_3(2\, L_3 - 1), \tag{32}$$

$$N_4 = P_1^2(L_1)\, P_1^2(L_2)\, P_0^2(L_3) = 4\, L_1\, L_2, \tag{33}$$

$$N_5 = P_0^2(L_1)\, P_1^2(L_2)\, P_1^2(L_3) = 4\, L_2\, L_3, \tag{34}$$

$$N_6 = P_1^2(L_1)\, P_0^2(L_2)\, P_1^2(L_3) = 4\, L_1\, L_3, \tag{35}$$

because

$$P_1^2(L_i) = \prod_{p=0}^{1-1} \frac{2\, L_i - p}{1 - p} = 2\, L_i, \tag{36}$$

and

$$P_2^2(L_i) = \prod_{p=0}^{2-1} \frac{2\, L_i - p}{2 - p} = \frac{2\, L_i}{2}\, \frac{2\, L_i - 1}{1} = L_i(2\, L_i - 1). \tag{37}$$

Figure 18 shows the shape functions N_1 and N_4. The other shape functions look like these, N_2 and N_3 are the same as N_1, moreover N_5 and N_6 are the same as N_4, but they must be rotated to the corresponding nodes.

(a) $N_1(x, y)$ (b) $N_4(x, y)$
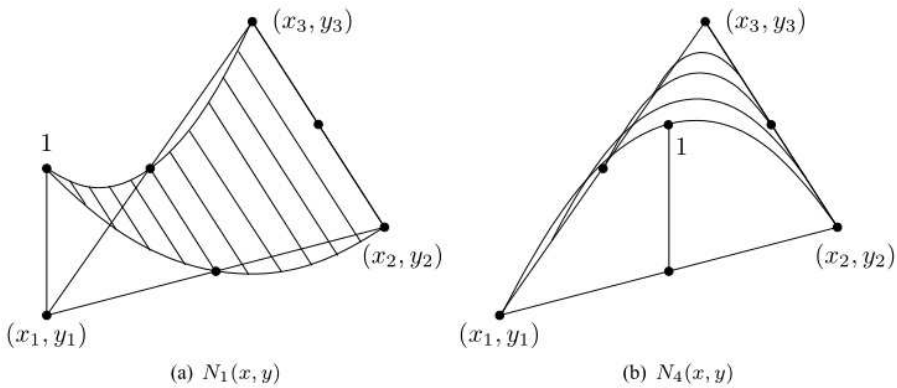
Figure 18. The 2D quadratic shape functions $N_1(x,y)$ and $N_4(x,y)$.

Finally, if $n = 3$, $m = 10$, shape functions can be constructed as (see Fig. 19)

$$N_1 = P_3^3(L_1)\, P_0^3(L_2)\, P_0^3(L_3) = \frac{1}{2}L_1(3\,L_1 - 1)(3\,L_1 - 2), \tag{38}$$

$$N_2 = P_0^3(L_1)\, P_3^3(L_2)\, P_0^3(L_3) = \frac{1}{2}L_2(3\,L_2 - 1)(3\,L_2 - 2), \tag{39}$$

$$N_3 = P_0^3(L_1)\, P_0^3(L_2)\, P_3^3(L_3) = \frac{1}{2}L_3(3\,L_3 - 1)(3\,L_3 - 2), \tag{40}$$

$$N_4 = P_2^3(L_1)\, P_1^3(L_2)\, P_0^3(L_3) = \frac{9}{2}L_1(3\,L_1 - 1)L_2, \tag{41}$$

$$N_5 = P_1^3(L_1)\, P_2^3(L_2)\, P_0^3(L_3) = \frac{9}{2}L_2(3\,L_2 - 1)L_1, \tag{42}$$

$$N_6 = P_0^3(L_1)\, P_2^3(L_2)\, P_1^3(L_3) = \frac{9}{2}L_2(3\,L_2 - 1)L_3, \tag{43}$$

$$N_7 = P_0^3(L_1)\, P_1^3(L_2)\, P_2^3(L_3) = \frac{9}{2}L_3(3\,L_3 - 1)L_2, \tag{44}$$

$$N_8 = P_1^3(L_1)\, P_0^3(L_2)\, P_2^3(L_3) = \frac{9}{2}L_3(3\,L_3 - 1)L_1, \tag{45}$$

$$N_9 = P_2^3(L_1)\, P_0^3(L_2)\, P_1^3(L_3) = \frac{9}{2}L_1(3\,L_1 - 1)L_3, \tag{46}$$

$$N_{10} = P_1^3(L_1)\, P_1^3(L_2)\, P_1^3(L_3) = 27\,L_1\,L_2\,L_3, \tag{47}$$

because

$$P_1^3(L_i) = \prod_{p=0}^{1-1} \frac{3\,L_i - p}{1 - p} = 3\,L_i, \tag{48}$$

Figure 19. The 2D cubic shape functions $N_1(x,y)$ and $N_5(x,y)$.

$$P_2^3(L_i) = \prod_{p=0}^{2-1} \frac{3\,L_i - p}{2 - p} = \frac{3\,L_i}{2}\,\frac{3\,L_i - 1}{1} = \frac{3}{2}L_i(3\,L_i - 1), \tag{49}$$

$$P_3^3(L_i) = \prod_{p=0}^{3-1} \frac{3\,L_i - p}{3 - p} = \frac{3\,L_i}{3}\,\frac{3\,L_i - 1}{2}\,\frac{3\,L_i - 2}{1}$$
$$= \frac{1}{2}L_i(3\,L_i - 1)(3\,L_i - 2). \tag{50}$$

These functions satisfy the condition (2). Figure 19 shows the shape functions N_1 and N_5, as examples. The other shape functions look like these, N_2 and N_3 are the same as N_1, N_4, N_6, N_7, N_8 and N_9 look like $N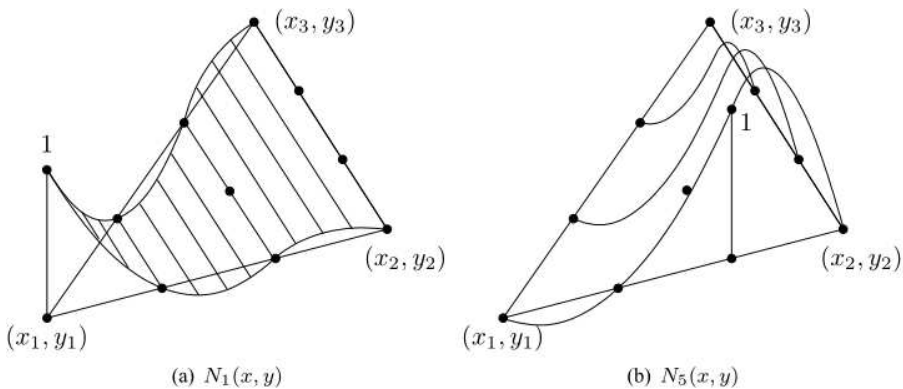_5$, but they must be imagined at the corresponding nodes. The shape function N_{10} is equal to one at the center of mass of the triangle and equal to zero on the other nine nodes.

The scalar potential along any edge of a triangle is the linear combination of the values defined in the points of this edge (see Fig. 15, Fig. 16, Fig. 17), so that if two triangles share the same vertice, the potential will be continuous across the interface element boundary. This means that the approximate solution is continuous everywhere, however, its normal derivate is not.

It is easy to see that the 1D shape functions are the same as the functions along the edges of a triangle.

(*iii*) 3D linear shape functions can be worked out as follows when using tetrahedral finite elements. Linear basis functions can be introduced again by using the barycentric coordinate system. The volume of a tetrahedron is denoted by V and it can be expressed as

$$V = \frac{1}{6} \begin{vmatrix} x_4 - x_1 & y_4 - y_1 & z_4 - z_1 \\ x_4 - x_2 & y_4 - y_2 & z_4 - z_2 \\ x_4 - x_3 & y_4 - y_3 & z_4 - z_3 \end{vmatrix}, \tag{51}$$

where (x_1,y_1,z_1), (x_2,y_2,z_2), (x_3,y_3,z_3) and (x_4,y_4,z_4) are the coordinates of the four nodes of the tetrahedron as shown in Fig. 20. The volume functions according to a given point inside the tetrahedron with coordinates (x,y,z) can be calculated as

$$V_1 = \frac{1}{6} \begin{vmatrix} x_4 - x & y_4 - y & z_4 - z \\ x_4 - x_2 & y_4 - y_2 & z_4 - z_2 \\ x_4 - x_3 & y_4 - y_3 & z_4 - z_3 \end{vmatrix}, \tag{52}$$

$$V_2 = \frac{1}{6} \begin{vmatrix} x_4 - x_1 & y_4 - y_1 & z_4 - z_1 \\ x_4 - x & y_4 - y & z_4 - z \\ x_4 - x_3 & y_4 - y_3 & z_4 - z_3 \end{vmatrix}, \tag{53}$$
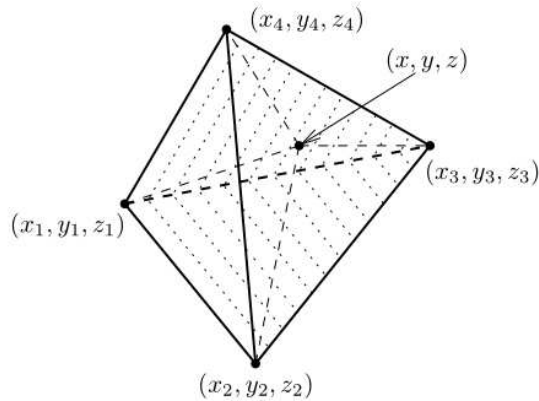
Figure 20. The volume functions in a tetrahedron.

$$V_3 = \frac{1}{6} \begin{vmatrix} x_4 - x_1 & y_4 - y_1 & z_4 - z_1 \\ x_4 - x_2 & y_4 - y_2 & z_4 - z_2 \\ x_4 - x & y_4 - y & z_4 - z \end{vmatrix}, \tag{54}$$

$$V_4 = \frac{1}{6} \begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ x - x_2 & y - y_2 & z - z_2 \\ x - x_3 & y - y_3 & z - z_3 \end{vmatrix}. \tag{55}$$

The barycentric coordinates $L_i = L_i(x,y,z)$ of a tetrahedron can be formulated as

$$L_i = \frac{V_i}{V}, \qquad i = 1,2,3,4. \tag{56}$$

Four linear shape functions $N_i = N_i(x,y,z)$ correspondingly to the four nodes are

$$N_i = L_i, \qquad i = 1,2,3,4. \tag{57}$$

A shape function N_i is equal to 1 at the ith node of the tetrahedron, moreover it is equal to zero at the other three nodes and it is varying linearly within the tetrahedron, because the fraction V_i/V measures the perpendicular distance of the point (x,y,z) toward the facet opposite to node i as it is illustrated in Fig. 21 and the linear shape function is constant along such a surface. That is why the relation (2) is satisfied. It is obvious that the four shape functions are linearly independent.
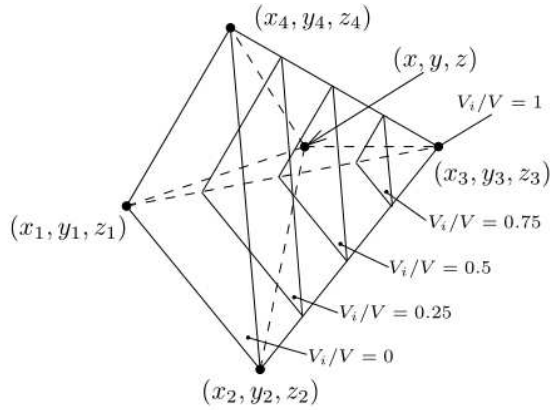
Figure 21. Fraction V_i/V measures the perpendicular distance of the point (x,y,z) toward the facet opposite to node i (here $i = 3$).

The higher order shape functions can be worked out similarly as it was mentioned in the case of triangular elements. The barycentric coordinates L_1, L_2, L_3 and L_4 can be used. A polynomial of order n must contain all possible terms $x^p y^q z^r$, $0 \leq p + q + r \leq n$ and a polynomial contains

$$m = \frac{(n+1)(n+2)(n+3)}{6} \tag{58}$$

elements altogether, i.e. $m = 1$, $m = 4$, $m = 10$ and $m = 20$ in the case of zeroth, first, second and third order polynomials. It means that m coefficients must be expressed and m points must be placed within a tetrahedron.

The interpolation function of order n can be constructed as

$$N_i = P_I^n(L_1)\, P_J^n(L_2)\, P_K^n(L_3)\, P_L^n(L_4), \quad \text{where} \quad I + J + K + L = n, \tag{59}$$

where the integers I, J, K and L label the nodes within the tetrahedra, resulting in a numbering scheme. Figure 22, Fig. 23 and Fig. 24 illustrate the numbering scheme of the first, the second and the third order approximations.

The polynomials $P_I^n(L_1)$, $P_J^n(L_2)$, $P_K^n(L_3)$ and $P_L^n(L_4)$ are defined in the same way as it was presented in the 2D situation, see definitions (22)–(25).

If $n = 1$, then $m = 4$, i.e. (see Fig. 22)

$$N_1 = P_1^1(L_1)\, P_0^1(L_2)\, P_0^1(L_3)\, P_0^1(L_4) = L_1, \tag{60}$$

$$N_2 = P_0^1(L_1)\, P_1^1(L_2)\, P_0^1(L_3)\, P_0^1(L_4) = L_2, \tag{61}$$
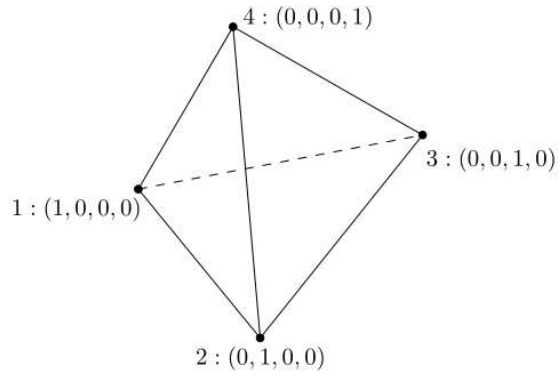
Figure 22. Numbering scheme for linear element, $n = 1$.

$$N_3 = P_0^1(L_1)\, P_0^1(L_2)\, P_1^1(L_3)\, P_0^1(L_4) = L_3, \tag{62}$$

$$N_4 = P_0^1(L_1)\, P_0^1(L_2)\, P_0^1(L_3)\, P_1^1(L_4) = L_4, \tag{63}$$

since (29) as it was mentioned in (57).

If $n = 2$, then $m = 10$, i.e. (see Fig. 23)

$$N_1 = P_2^2(L_1)\, P_0^2(L_2)\, P_0^2(L_3)\, P_0^2(L_4) = L_1(2\,L_1 - 1), \tag{64}$$

$$N_2 = P_0^2(L_1)\, P_2^2(L_2)\, P_0^2(L_3)\, P_0^2(L_4) = L_2(2\,L_2 - 1), \tag{65}$$
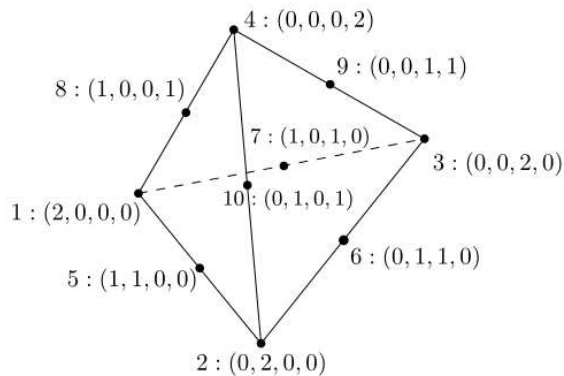
Figure 23. Numbering scheme for quadratic element, $n = 2$.

$$N_3 = P_0^2(L_1)\, P_0^2(L_2)\, P_2^2(L_3)\, P_0^2(L_4) = L_3(2\,L_3 - 1), \tag{66}$$

$$N_4 = P_0^2(L_1)\, P_0^2(L_2)\, P_0^2(L_3)\, P_2^2(L_4) = L_4(2\,L_4 - 1), \tag{67}$$

$$N_5 = P_1^2(L_1)\, P_1^2(L_2)\, P_0^2(L_3)\, P_0^2(L_4) = 4\,L_1\,L_2, \tag{68}$$

$$N_6 = P_0^2(L_1)\, P_1^2(L_2)\, P_1^2(L_3)\, P_0^2(L_4) = 4\,L_2\,L_3, \tag{69}$$

$$N_7 = P_1^2(L_1)\, P_0^2(L_2)\, P_1^2(L_3)\, P_0^2(L_4) = 4\,L_1\,L_3, \tag{70}$$

$$N_8 = P_1^2(L_1)\, P_0^2(L_2)\, P_0^2(L_3)\, P_1^2(L_4) = 4\,L_1\,L_4, \tag{71}$$

$$N_9 = P_0^2(L_1)\, P_0^2(L_2)\, P_1^2(L_3)\, P_1^2(L_4) = 4\,L_3\,L_4, \tag{72}$$

$$N_{10} = P_0^2(L_1)\, P_1^2(L_2)\, P_0^2(L_3)\, P_1^2(L_4) = 4\,L_2\,L_4, \tag{73}$$

because (36) and (37).

Finally, if $n = 3$, $m = 20$ shape functions can be constructed as (see Fig. 24)

$$N_1 = P_3^3(L_1)\, P_0^3(L_2)\, P_0^3(L_3)\, P_0^3(L_4) = \frac{1}{2} L_1(3\,L_1 - 1)(3\,L_1 - 2), \tag{74}$$

$$N_2 = P_0^3(L_1)\, P_3^3(L_2)\, P_0^3(L_3)\, P_0^3(L_4) = \frac{1}{2} L_2(3\,L_2 - 1)(3\,L_2 - 2), \tag{75}$$

$$N_3 = P_0^3(L_1)\, P_0^3(L_2)\, P_3^3(L_3)\, P_0^3(L_4) = \frac{1}{2} L_3(3\,L_3 - 1)(3\,L_3 - 2), \tag{76}$$

$$N_4 = P_0^3(L_1)\, P_0^3(L_2)\, P_0^3(L_3)\, P_3^3(L_4) = \frac{1}{2} L_4(3\,L_4 - 1)(3\,L_4 - 2), \tag{77}$$

$$N_5 = P_2^3(L_1)\, P_1^3(L_2)\, P_0^3(L_3)\, P_0^3(L_4) = \frac{9}{2} L_1(3\,L_1 - 1)L_2, \tag{78}$$
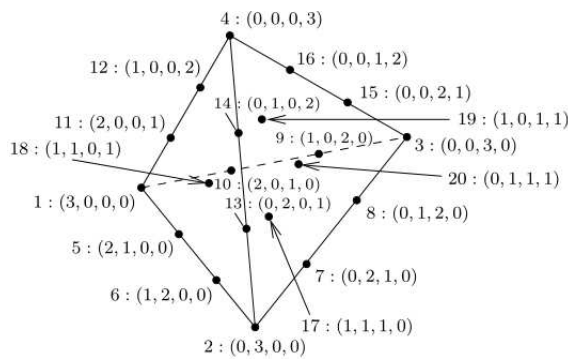

Figure 24. Numbering scheme for cubic element, $n = 3$.

$$N_6 = P_1^3(L_1)\, P_2^3(L_2)\, P_0^3(L_3)\, P_0^3(L_4) = \frac{9}{2} L_2(3\, L_2 - 1)L_1, \tag{79}$$

$$N_7 = P_0^3(L_1)\, P_2^3(L_2)\, P_1^3(L_3)\, P_0^3(L_4) = \frac{9}{2} L_2(3\, L_2 - 1)L_3, \tag{80}$$

$$N_8 = P_0^3(L_1)\, P_1^3(L_2)\, P_2^3(L_3)\, P_0^3(L_4) = \frac{9}{2} L_3(3\, L_3 - 1)L_2, \tag{81}$$

$$N_9 = P_1^3(L_1)\, P_0^3(L_2)\, P_2^3(L_3)\, P_0^3(L_4) = \frac{9}{2} L_3(3\, L_3 - 1)L_1, \tag{82}$$

$$N_{10} = P_2^3(L_1)\, P_0^3(L_2)\, P_1^3(L_3)\, P_0^3(L_4) = \frac{9}{2} L_1(3\, L_1 - 1)L_3, \tag{83}$$

$$N_{11} = P_1^3(L_1)\, P_0^3(L_2)\, P_0^3(L_3)\, P_2^3(L_4) = \frac{9}{2} L_4(3\, L_4 - 1)L_1, \tag{84}$$

$$N_{12} = P_2^3(L_1)\, P_0^3(L_2)\, P_0^3(L_3)\, P_1^3(L_4) = \frac{9}{2} L_1(3\, L_1 - 1)L_4, \tag{85}$$

$$N_{13} = P_0^3(L_1)\, P_1^3(L_2)\, P_0^3(L_3)\, P_2^3(L_4) = \frac{9}{2} L_4(3\, L_4 - 1)L_2, \tag{86}$$

$$N_{14} = P_0^3(L_1)\, P_2^3(L_2)\, P_0^3(L_3)\, P_1^3(L_4) = \frac{9}{2} L_2(3\, L_2 - 1)L_4, \tag{87}$$

$$N_{15} = P_0^3(L_1)\, P_0^3(L_2)\, P_1^3(L_3)\, P_2^3(L_4) = \frac{9}{2} L_4(3\, L_4 - 1)L_3, \tag{88}$$

$$N_{16} = P_0^3(L_1)\, P_0^3(L_2)\, P_2^3(L_3)\, P_1^3(L_4) = \frac{9}{2} L_3(3\, L_3 - 1)L_4, \tag{89}$$

$$N_{17} = P_1^3(L_1)\, P_1^3(L_2)\, P_1^3(L_3)\, P_0^3(L_4) = 27\, L_1\, L_2\, L_3, \tag{90}$$

$$N_{18} = P_1^3(L_1)\, P_1^3(L_2)\, P_0^3(L_3)\, P_1^3(L_4) = 27\, L_1\, L_2\, L_4, \tag{91}$$

$$N_{19} = P_1^3(L_1)\, P_0^3(L_2)\, P_1^3(L_3)\, P_1^3(L_4) = 27\, L_1\, L_3\, L_4, \tag{92}$$

$$N_{20} = P_0^3(L_1)\, P_1^3(L_2)\, P_1^3(L_3)\, P_1^3(L_4) = 27\, L_2\, L_3\, L_4, \tag{93}$$

because of the equations (48), (49) and (50).

The scalar potential along any edge of a tetrahedron is the linear combination of the values defined on the points of the given edge, so that if two tetrahedra share the same facet, the potential will be continuous across this interface. This means that the approximate solution is continuous everywhere, however, its normal derivate is not.

If potentials at the nodes are known, then a linear approximation of the potential function can be represented by (1).

The sum of all nodal shape functions is equal to 1, hence the sum of their gradient is zero,

$$\sum_{i=1}^{m} N_i = 1, \quad \text{and} \quad \sum_{i=1}^{m} \nabla N_i = 0. \tag{94}$$

This means that the maximal number of linearly independent nodal basis functions is m and the maximal number of linearly independent gradients of the nodal basis functions is $m - 1$, i.e. shape functions are linearly independent but their gradients are not.

2.2. Edge finite elements

Vector potentials can be represented either by nodal shape functions or by so-called edge shape functions. Edge shape functions are also called vector shape functions.

The natural approach is to treat the vector field $\boldsymbol{T} = \boldsymbol{T}(\boldsymbol{r},t)$ as two or three coupled scalar fields $T_x = T_x(\boldsymbol{r},t)$, $T_y = T_y(\boldsymbol{r},t)$ and $T_z = T_z(\boldsymbol{r},t)$, i.e.

$$\boldsymbol{T} = T_x \boldsymbol{e}_x + T_y \boldsymbol{e}_y, \tag{95}$$

and

$$\boldsymbol{T} = T_x \boldsymbol{e}_x + T_y \boldsymbol{e}_y + T_z \boldsymbol{e}_z \tag{96}$$

in 2D and in 3D situations, respectively, $\boldsymbol{e}_x$, $\boldsymbol{e}_y$ and $\boldsymbol{e}_z$ are the orthogonal unit vectors in the $x - y$ and in the $x - y - z$ plane.

Nodal shape functions can be used in this case as well, as it was presented for scalar potentials in the previous section, however, each node has two or three unknowns. Nodal shape functions can be applied to approximate the scalar components of the vector field $\boldsymbol{T}$. For example in 3D, $\boldsymbol{T}$ can be approximated as

$$\begin{aligned}
\boldsymbol{T} &\simeq \sum_{i=1}^{m} N_i \left(T_{x,i}\, \boldsymbol{e}_x + T_{y,i}\, \boldsymbol{e}_y + T_{z,i}\, \boldsymbol{e}_z \right) \\
&= \sum_{i=1}^{m} N_i T_{x,i}\, \boldsymbol{e}_x + \sum_{i=1}^{m} N_i T_{y,i}\, \boldsymbol{e}_y + \sum_{i=1}^{m} N_i T_{z,i}\, \boldsymbol{e}_z.
\end{aligned} \tag{97}$$

Here $N_i = N_i(\boldsymbol{r})$ are the usual nodal shape functions defined by (2) and $T_{x,i} = T_{x,i}(t)$, $T_{y,i} = T_{y,i}(t)$, $T_{z,i} = T_{z,i}(t)$ are the values of components of the approximated vector potential at node i. The number of degrees of freedom is $2m$ in a 2D problem using triangular mesh and $3m$ in a 3D arrangement meshed by tetrahedral elements.

Nodal shape functions are used to approximate gauged vector potentials, which was the first in the history of finite element method in electromagnetics. Unfortunately, there are some problems when the usual nodal based finite elements are used to interpolate vector potentials. The lack of enforcement of the divergence condition (lack of gauging) results in a system of algebraic equations, which has infinite number of solution and the application of iterative solvers sometimes fails. We have to take care about the Coulomb gauge. There are problems on the iron/air interface when using the magnetic vector potential approximated by nodal elements and extra interface conditions must be set up to solve this problem.

Fortunately, vector shape functions have been developed in the last decades, which application in static and eddy current field problems is more and more popular, because of their advantages. The use of edge shape functions solves the problems described above. It will be shown that the divergence of vector shape functions is equal to zero, that is why, gauging is satisfied automatically. The ungauged potential functions are approximated by vector elements. Vector shape functions are usually called edge shape functions, because they are associated to the edges of the FEM mesh. Vector shape functions are more and more popular in wave problems, too.

Instead of scalar shape functions, vector shape functions (or edge shape functions) $\boldsymbol{W}_i = \boldsymbol{W}_i(\boldsymbol{r})$ can be applied to approximate a vector potential $\boldsymbol{T}$,

$$\boldsymbol{T} \simeq \sum_{i=1}^{k} \boldsymbol{W}_i\, T_i, \tag{98}$$

where $T_i = T_i(t)$ is the line integral of the vector potential $\boldsymbol{T}$ along the edge i. First order vector shape functions are defined by the line integral

$$\int_l \boldsymbol{W}_i \cdot \mathrm{d}\boldsymbol{l} = \left\{ \begin{array}{ll} 1, & \text{along edge } i, \\ 0, & \text{along other edges,} \end{array} \right. \tag{99}$$

i.e. the line integral of the vector shape function $\boldsymbol{W}_i$ along the ith edge is equal to one. In other words, the vector shape function $\boldsymbol{W}_i$ has tangential component only along the ith edge and it has only normal component along the other edges, because $\boldsymbol{W}_i \cdot \mathrm{d}\boldsymbol{l}$ is equal to zero only if the vectors $\boldsymbol{W}_i$ and $\mathrm{d}\boldsymbol{l}$ are perpendicular to each others and $|\boldsymbol{W}_i||\mathrm{d}\boldsymbol{l}| > 0$. Moreover, in 3D case, the vector shape function $\boldsymbol{W}_i$ has zero tangential component along every facet of the 3D finite element, which not share the edge i.

If two triangles share the same vertices, the tangential component of the approximated vector potential will be continuous across the interface element boundary. This is true in 3D case as well, moreover, if two tetrahedra share the same facet, the tangential component of the vector potential will be continuous across this interface. This means that the tangential component of the approximate solution is continuous everywhere, however, its normal component is not. In the words of equations, according to the definition (99), the line integral of the vector potential along the mth edge is equal to T_m, i.e.

$$\begin{aligned} \int_{l_m} \boldsymbol{T} \cdot \mathrm{d}\boldsymbol{l} &= \int_{l_m} \left(\sum_{i=1}^{k} \boldsymbol{W}_i\, T_i \right) \cdot \mathrm{d}\boldsymbol{l} = \sum_{i=1}^{k} \int_{l_m} (\boldsymbol{W}_i\, T_i) \cdot \mathrm{d}\boldsymbol{l} \\ &= T_m \int_{l_m} \boldsymbol{W}_m \cdot \mathrm{d}\boldsymbol{l} = T_m. \end{aligned} \tag{100}$$

That is why, edge shape functions are also called tangentially continuous shape functions.

The gradients of the nodal shape functions are in the function space spanned by the edge basis functions, that is

$$\nabla N_j = \sum_{i=1}^{k} c_{ji} \boldsymbol{W}_i, \quad j = 1, \cdots, m-1, \tag{101}$$

where $\sum_{i=1}^{k} c_{ji}^2 > 0$. Taking the curl of each equation in (101) results in

$$\sum_{i=1}^{k} c_{ji} \nabla \times \boldsymbol{W}_i = \boldsymbol{0}, \quad j = 1, \cdots, m-1, \tag{102}$$

because $\nabla \times (\nabla \varphi) \equiv 0$. This shows that the maximal number of linearly independent curls of the edge basis functions is $k - (m-1)$. The interdependence of the curls of the edge basis functions means that an ungauged formulation leads to a singular, positive semidefinit finite element curl-curl matrix. Singular systems can be solved by iterative methods, if the right-hand side of the system of equations is consistent. We took care about it when obtaining the weak formulations of the ungauged version of potentials, because excitation current density has been taken into account by the use of impressed current vector potential, $\boldsymbol{T}_0$.

The vector function

$$\boldsymbol{w}_{ij} = L_i \nabla L_j - L_j \nabla L_i \tag{103}$$

will be applied to construct the edge shape functions, because it can be used in functions, which satisfies (99) and (100). In 2D, L_i ($i = 1,2,3$) are the barycentric coordinates of the triangle defined by (18). In 3D, L_i ($i = 1,2,3,4$) are the barycentric coordinates of
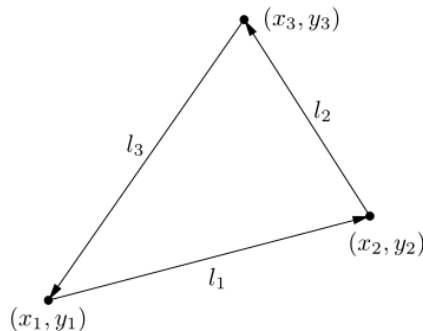

Figure 25. The definition of edges with local directions of the triangular finite element.
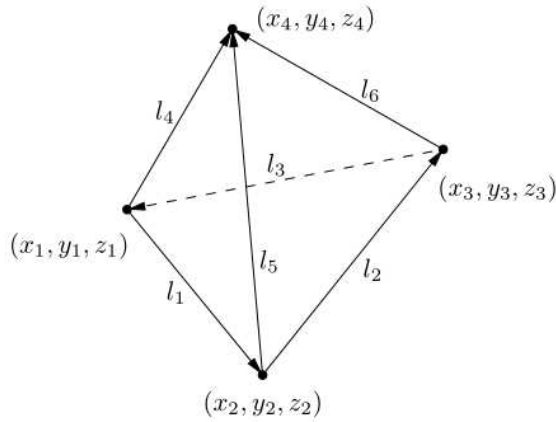
Figure 26. The definition of edges with local directions of the tetrahedral finite element.

the tetrahedron defined by (56). According to the notations in (103), the edges of a finite element are pointing from node i to node j, as it can be seen in Fig. 25 and in Fig. 26.

The vector field $\boldsymbol{w}_{ij}$ has the following important properties, which proofs the use of vector function $\boldsymbol{w}_{ij}$ as vector shape function.

(*i*) Let $\boldsymbol{e}_{ij}$ is a unit vector pointing from node i to node j, then

$$\boldsymbol{e}_{ij} \cdot \boldsymbol{w}_{ij} = \frac{1}{l_{ij}}, \tag{104}$$

where l_{ij} is the length of edge $\{i,j\}$. This means that $\boldsymbol{w}_{ij}$ has constant tangential component along the edge $\{i,j\}$.

Since L_i and L_j are linear functions that vary from node i to node j from 1 to 0 and from 0 to 1, respectively, we have $\boldsymbol{e}_{ij} \cdot \nabla L_i = -1/l_{ij}$ and $\boldsymbol{e}_{ij} \cdot \nabla L_j = 1/l_{ij}$, finally

$$\boldsymbol{e}_{ij} \cdot \boldsymbol{w}_{ij} = L_i \frac{1}{l_{ij}} + L_j \frac{1}{l_{ij}} = \frac{L_i + L_j}{l_{ij}} = \frac{1}{l_{ij}}, \tag{105}$$

because $L_i + L_j = 1$ along the edge $\{i,j\}$. See, for example Fig. 14 and let $i = 1$, $j = 2$, so $N_1 = L_1$ is decreasing along edge $\{1,2\}$ and $N_2 = L_2$ is increasing along the same edge. See also Fig. 7, from which it is easy to see the gradients $\boldsymbol{e}_{ij} \cdot \nabla L_i = -1/l_{ij}$ and $\boldsymbol{e}_{ij} \cdot \nabla L_j = 1/l_{ij}$.

(*ii*) In 2D, the function L_i varies linearly from node i to the opposite edge $\{j,k\}$ (see e.g. N_1 in Fig. 14, $i = 1$, $j = 2$, $k = 3$), i.e. the vector field ∇L_i is perpendicular to this edge, but L_i is zero there, that is why $\boldsymbol{w}_{ij}$ is perpendicular to the edge $\{j,k\}$,

$$\boldsymbol{w}_{ij} = -L_j \nabla L_i, \quad \text{on the edge} \quad \{j,k\}, \tag{106}$$

and the length of this vector is decreasing from node j to k according to L_j. On the other hand, the function L_j varies linearly from node j to the opposite edge $\{k,i\}$ (see N_2 in Fig. 14), i.e. ∇L_j is perpendicular to this edge, but L_j is zero there and $\boldsymbol{w}_{ij}$ is perpendicular to the edge $\{k,i\}$,

$$\boldsymbol{w}_{ij} = L_i \nabla L_j, \quad \text{on the edge} \quad \{k,i\}, \tag{107}$$

and the length of this vector is decreasing from node i to k according to L_i.
This with item (i) means that the vector function $\boldsymbol{w}_{ij}$ has tangential component only on the edge $\{i,j\}$ and it is perpendicular to the other edges.
In 3D, this is valid to the whole triangular facet with the bounding edges opposite to a node, see e.g. Fig. 21.

(*iii*) The vector field $\boldsymbol{w}_{ij}$ is divergence-free,

$$\begin{aligned}
\nabla \cdot \boldsymbol{w}_{ij} &= \nabla \cdot (L_i \nabla L_j - L_j \nabla L_i) = \nabla \cdot (L_i \nabla L_j) - \nabla \cdot (L_j \nabla L_i) \\
&= \nabla L_i \cdot \nabla L_j + L_i \nabla \cdot \nabla L_j - \nabla L_j \cdot \nabla L_i - L_j \nabla \cdot \nabla L_i = 0,
\end{aligned} \tag{108}$$

by using the identity

$$\nabla \cdot (\varphi \, \boldsymbol{v}) = \nabla \varphi \cdot \boldsymbol{v} + \varphi \nabla \cdot \boldsymbol{v} \tag{109}$$

with the notations $\varphi = L_i$, $\boldsymbol{v} = \nabla L_j$ in the second and $\varphi = L_j$, $\boldsymbol{v} = \nabla L_i$ in the last term. The barycentric coordinates are linear functions of the coordinates and their gradient is constant, which divergence is equal to zero, i.e. the second and fourth terms are vanishing. The first and the third terms are equal, finally, $\nabla \cdot \boldsymbol{w}_{ij} = 0$.

(*iv*) The vector field $\boldsymbol{w}_{ij}$ has constant curl,

$$\begin{aligned}
\nabla \times \boldsymbol{w}_{ij} &= \nabla \times (L_i \nabla L_j - L_j \nabla L_i) = \nabla \cdot (L_i \nabla L_j) - \nabla \cdot (L_j \nabla L_i) \\
&= L_i \nabla \times \nabla L_j - \nabla L_j \times \nabla L_i - L_j \nabla \times \nabla L_i + \nabla L_i \times \nabla L_j \\
&= 2 \nabla L_i \times \nabla L_j,
\end{aligned} \tag{110}$$

by using the identity

$$\nabla \times (\varphi \, \boldsymbol{v}) = \varphi \nabla \times \boldsymbol{v} - \boldsymbol{v} \times \nabla \varphi \tag{111}$$

with the notations $\varphi = L_i$, $\boldsymbol{v} = \nabla L_j$ in the first and $\varphi = L_j$, $\boldsymbol{v} = \nabla L_i$ in the second term. The first and the third terms are equal to zero because of the identity $\nabla \times \nabla \varphi \equiv 0$ for any function φ. The second term can be reformulated by $\boldsymbol{a} \times \boldsymbol{b} = -\boldsymbol{b} \times \boldsymbol{a}$, finally, the result is constant, because the gradients of the barycentric coordinates are constant.

First, the edge shape functions defined on triangles based on (103) are collected. The basic 2D vector shape functions $\boldsymbol{W}_i$ can be constructed by using the first order nodal shape functions,

$$\boldsymbol{W}_1 = l_1(N_1 \nabla N_2 - N_2 \nabla N_1)\delta_1, \tag{112}$$

$$\boldsymbol{W}_2 = l_2(N_2 \nabla N_3 - N_3 \nabla N_2)\delta_2, \tag{113}$$

$$\boldsymbol{W}_3 = l_3(N_3 \nabla N_1 - N_1 \nabla N_3)\delta_3. \tag{114}$$

Here l_i (Fig. 25) denotes the length of the ith edges of the triangle and it is used to normalize the edge shape function according to (104). The edge basis function $\boldsymbol{W}_i$ ($i = 1,2,3$) has tangential component only along the ith edge and it is perpendicular to the other two edges as represented in Fig. 27(a)-27(c). It is easy to see that an edge shape function has magnitude and direction. The value of δ_i is equal to ± 1, depending on whether the local direction of the edge is the same as the global direction or opposite (see Fig. 25 for local direction). This set of vector functions is called zeroth order vector shape functions.

If the approximation of the vector function $\boldsymbol{T}$ is known along the edges of the mesh, then (98) can be used to interpolate the function anywhere and in linear case $K = 3$.

Higher order vector shape functions can be constructed by using the vector function $\boldsymbol{w}_{ij}$ defined by (103), too. This vector function must be multiplied by a complete interpolatory polynomial, which results in the higher order vector shape functions. First and second order polynomials will be used to build up first and second order vector shape functions. Here, we follow, the method is as follows.

First of all, an indexing sequence must be set up, which is similar to the method used to build up the scalar shape functions, because the higher order vector shape functions are based on the Lagrange polynomials and (103). In the case of first order approximation, the numbering scheme of the third order scalar interpolation can be used and the points are shown in Fig. 28, Fig. 29 and Fig. 30 must be used to represent first order vector shape functions associated to the edge $\{1,2\}$, $\{2,3\}$ and $\{3,1\}$, respectively. In the case of second order approximation, the numbering scheme of the fourth order scalar interpolation can be used and the interpolation points shown in Fig. 31, Fig. 32 and Fig. 33 must be used to represent second order vector shape functions associated to the edge $\{1,2\}$, $\{2,3\}$ and $\{3,1\}$, respectively. The interpolation points have been selected in this special way, because the interpolation of field vectors along vertices has been avoided, i.e. the points have been shifted inside the triangle and the indexing scheme of order $n + 2$ is used to represent the vector interpolation of order n. This is called global numbering and denoted
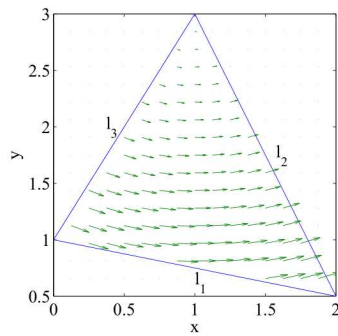
(a) The edge shape function $\boldsymbol{W}_1$
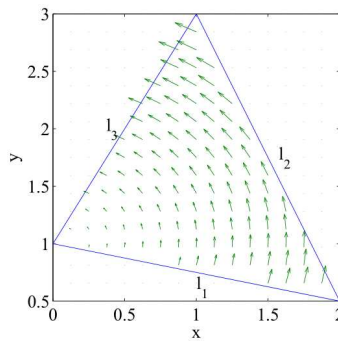
(b) The edge shape function $\boldsymbol{W}_2$
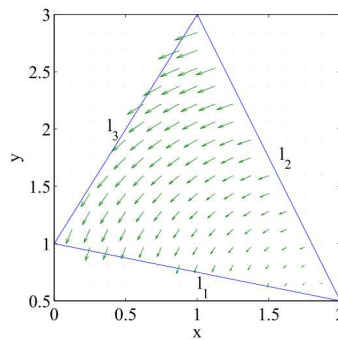
(c) The edge shape function $\boldsymbol{W}_3$

Figure 27. The 2D edge shape functions.

Figure 28. Numbering scheme for the first order vector element associated with $\boldsymbol{w}_{12}$.
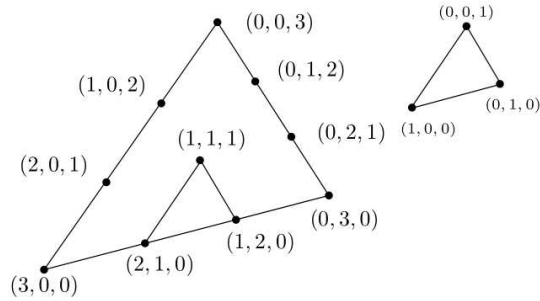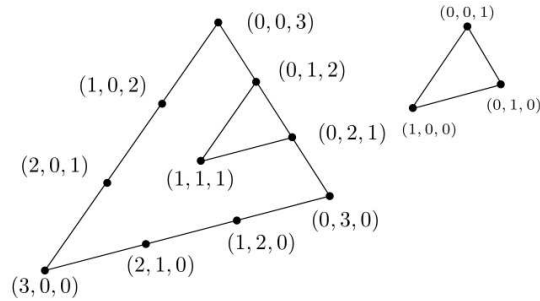

Figure 29. Numbering scheme for the first order vector element associated with $\boldsymbol{w}_{23}$.
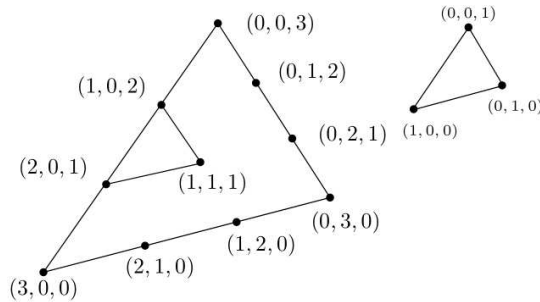
Figure 30. Numbering scheme for the first order vector element associated with $\boldsymbol{w}_{31}$.

by (I,J,K) on the 'big' triangle, local numbering means the numbering scheme with the real order (i,j,k) defined over the 'small' triangle.

Figure 31. Numbering scheme for the second order vector element associated with $\boldsymbol{w}_{12}$.

Figure 32. Numbering scheme for the second order vector element associated with $\boldsymbol{w}_{23}$.

Figure 33. Numbering scheme for the second order vector element associated with $\boldsymbol{w}_{31}$.
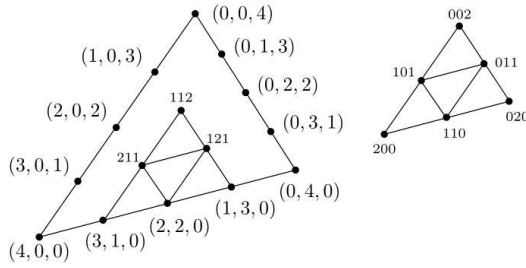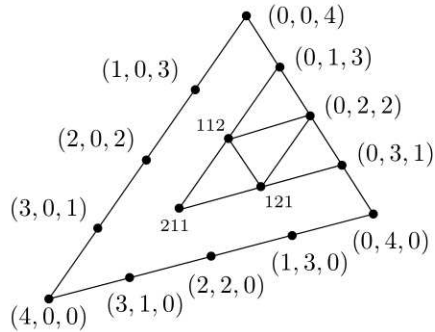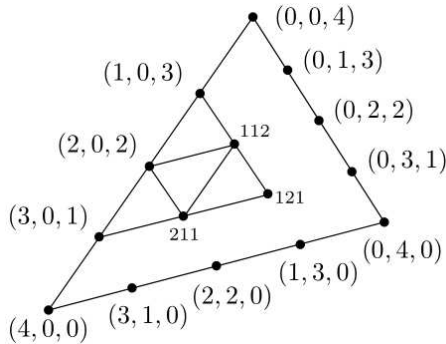
It is noted here that the COMSOL Multiphysics software uses this kind of vector shape functions, however, $n = 0$, $n = 1$ and $n = 2$ are named as linear, quadratic and cubic vector shape functions.

The vector function $\boldsymbol{w}_{ab}$ (associated to the edge pointing from node a to node b) can be multiplied by the Lagrange polynomials as

$$\boldsymbol{W}_{ab}^{IJK} = \alpha_{ab}^{IJK} \, P_i^n(l_1) \, P_j^n(l_2) \, P_k^n(l_3) \, \boldsymbol{w}_{ab}, \tag{115}$$

where n is the order of approximation and the integers i, j and k satisfy $i + j + k = n$ (see the small triangles in Fig. 28-Fig. 33).

If $n = 0$, the basic vector shape functions can be obtained, because $P_m^0(\cdot) = 1$ and α can be selected as the length of the appropriate edge, l_{ab}, since α_{ab}^{IJK} is a normalizing factor. The barycentric coordinates l_1, l_2 and l_3 are imagined in the small triangles. The transformation between local and global numbering is as follows:

$$i = I - 1, \quad j = J - 1, \quad k = K, \quad \text{on the edge} \quad \{1,2\}, \tag{116}$$

$$i = I, \quad j = J - 1, \quad k = K - 1, \quad \text{on the edge} \quad \{2,3\}, \tag{117}$$

$$i = I - 1, \quad j = J, \quad k = K - 1, \quad \text{on the edge} \quad \{3,1\}. \tag{118}$$

The relation between the barycentric coordinates of the small and the big triangles is as follows:

$$l_1 = L_1 \frac{n+2}{n}, \quad l_2 = L_2 \frac{n+2}{n} - \frac{1}{n}, \quad l_3 = L_3 \frac{n+2}{n} - \frac{1}{n}. \tag{119}$$

Using these relations, (115) can be written as (let here $\{ab\} = \{23\}$ for simplicity)

$$\boldsymbol{W}_{23}^{IJK} = \alpha_{23}^{IJK} \, P_I^n \left(L_1 \frac{n+2}{n} \right) P_{J-1}^n \left(L_2 \frac{n+2}{n} - \frac{1}{n} \right)$$
$$P_{K-1}^n \left(L_3 \frac{n+2}{n} - \frac{1}{n} \right) \boldsymbol{w}_{23}. \tag{120}$$

According to (22), Lagrange polynomials can be reformulated as

$$P_I^n \left(L_1 \frac{n+2}{n} \right) = \frac{1}{I!} \prod_{p=0}^{I-1} \left(n L_1 \frac{n+2}{n} - p \right)$$
$$= \frac{1}{I!} \prod_{p=0}^{I-1} [(n+2)L_1 - p] = P_I^{n+2}(L_1), \quad \text{if} \quad I > 0, \tag{121}$$

and

$$
\begin{aligned}
P_{J-1}^n \left(L_2 \frac{n+2}{n} - \frac{1}{n} \right) &= \frac{1}{(J-1)!} \prod_{p=0}^{J-2} \left[n \left(L_2 \frac{n+2}{n} - \frac{1}{n} \right) - p \right] \\
&= \frac{1}{(J-1)!} \prod_{p=0}^{J-2} \left[(n+2)L_2 - 1 - p \right] \\
&= \frac{1}{(J-1)!} \prod_{p=0}^{J-2} \left[(n+2)\left(L_2 - \frac{1}{n+2} \right) - p \right] \\
&= P_{J-1}^{n+2} \left(L_2 - \frac{1}{n+2} \right), \quad \text{if} \quad J > 1.
\end{aligned}
\tag{122}
$$

The so-called shifted Silvester polynomials can be used to simplify the relations above,

$$
S_J^{n+2}(L_2) = P_{J-1}^{n+2}\left(L_2 - \frac{1}{n+2} \right) = \frac{1}{(J-1)!} \prod_{p=0}^{J-1} \left[(n+2)L_2 - p \right].
\tag{123}
$$

Finally, the higher order vector shape functions can be formulated as follows by using the Lagrange and Silvester polynomials:

$$
\boldsymbol{W}_{12}^{IJK} = \alpha_{12}^{IJK} \, S_I^{n+2}(L_1) \, S_J^{n+2}(L_2) \, P_K^{n+2}(L_3) \, \boldsymbol{w}_{12},
\tag{124}
$$

$$
\boldsymbol{W}_{23}^{IJK} = \alpha_{23}^{IJK} \, P_I^{n+2}(L_1) \, S_J^{n+2}(L_2) \, S_K^{n+2}(L_3) \, \boldsymbol{w}_{23},
\tag{125}
$$

$$
\boldsymbol{W}_{31}^{IJK} = \alpha_{31}^{IJK} \, S_I^{n+2}(L_1) \, P_J^{n+2}(L_2) \, S_K^{n+2}(L_3) \, \boldsymbol{w}_{31}.
\tag{126}
$$

The parameter denoted by α is a normalization factor, which must have the value such that the line integral of vector shape function $\boldsymbol{W}_{ab}^{IJKL}$ is equal to 1 on the edge pointing from node a to node b. Here,

$$
P_0^n(\cdot) = 1, \quad \text{and} \quad S_1^n(\cdot) = 1.
\tag{127}
$$

The number of vector basis functions is

$$
k = (n+1)(n+3).
\tag{128}
$$

There is one shape function associated with the introduced interpolation nodes on the edges. It means $3(n+1)$ basis functions. There are three basis functions for an interior interpolation point, because every interpolation point inside the triangle is used to build all the vector shape functions in the three edges. Since a surface vector has only two degrees of freedom, these three basis functions are not independent and one of them must be discarded. This results in $n(n+1)$ interior basis functions. In total, the number of shape functions is $3(n+1) + n(n+1) = (n+3)(n+1)$.

In the case of first order approximation $n = 0$ and $k = 3$. In the case of second order approximation $n = 1$ and $K = 8$. In the case of third order approximation $n = 2$ and $k = 15$ and so on.

The first order vector shape functions are as follows:

$$\boldsymbol{W}_1 = \boldsymbol{W}_{12}^{120} = \alpha_{12}^{120}\, S_1^3(L_1)\, S_2^3(L_2)\, P_0^3(L_3)\, \boldsymbol{w}_{12} = \alpha_{12}^{120}(3\, L_2 - 1)\, \boldsymbol{w}_{12}, \qquad (129)$$

$$\boldsymbol{W}_2 = \boldsymbol{W}_{12}^{210} = \alpha_{12}^{210}\, S_2^3(L_1)\, S_1^3(L_2)\, P_0^3(L_3)\, \boldsymbol{w}_{12} = \alpha_{12}^{210}(3\, L_1 - 1)\, \boldsymbol{w}_{12}, \qquad (130)$$

$$\boldsymbol{W}_3 = \boldsymbol{W}_{12}^{111} = \alpha_{12}^{111}\, S_1^3(L_1)\, S_1^3(L_2)\, P_1^3(L_3)\, \boldsymbol{w}_{12} = \alpha_{12}^{111}\, 3\, L_3\, \boldsymbol{w}_{12}, \qquad (131)$$

$$\boldsymbol{W}_4 = \boldsymbol{W}_{23}^{012} = \alpha_{23}^{012}\, P_0^3(L_1)\, S_1^3(L_2)\, S_2^3(L_3)\, \boldsymbol{w}_{23} = \alpha_{23}^{012}(3\, L_3 - 1)\, \boldsymbol{w}_{23}, \qquad (132)$$

$$\boldsymbol{W}_5 = \boldsymbol{W}_{23}^{021} = \alpha_{23}^{021}\, P_0^3(L_1)\, S_2^3(L_2)\, S_1^3(L_3)\, \boldsymbol{w}_{23} = \alpha_{23}^{021}(3\, L_2 - 1)\, \boldsymbol{w}_{23}, \qquad (133)$$

$$\boldsymbol{W}_6 = \boldsymbol{W}_{31}^{102} = \alpha_{31}^{102}\, S_1^3(L_1)\, P_0^3(L_2)\, S_2^3(L_3)\, \boldsymbol{w}_{31} = \alpha_{31}^{102}(3\, L_3 - 1)\, \boldsymbol{w}_{31}, \qquad (134)$$

$$\boldsymbol{W}_7 = \boldsymbol{W}_{31}^{201} = \alpha_{31}^{201}\, S_2^3(L_1)\, P_0^3(L_2)\, S_1^3(L_3)\, \boldsymbol{w}_{31} = \alpha_{31}^{201}(3\, L_1 - 1)\, \boldsymbol{w}_{31}, \qquad (135)$$

$$\boldsymbol{W}_8 = \boldsymbol{W}_{31}^{111} = \alpha_{31}^{111}\, S_1^3(L_1)\, P_1^3(L_2)\, S_1^3(L_3)\, \boldsymbol{w}_{31} = \alpha_{31}^{111}\, 3\, L_2\, \boldsymbol{w}_{31}. \qquad (136)$$

The second order vector shape functions are as follows using (124)–(127):

$$\begin{aligned} \boldsymbol{W}_1 = \boldsymbol{W}_{12}^{310} &= \alpha_{12}^{310}\, S_3^4(L_1)\, S_1^4(L_2)\, P_0^4(L_3)\, \boldsymbol{w}_{12} \\ &= \alpha_{12}^{310} \frac{1}{2}(4\, L_1 - 1)(4\, L_1 - 2)\, \boldsymbol{w}_{12}, \end{aligned} \qquad (137)$$

$$\begin{aligned} \boldsymbol{W}_2 = \boldsymbol{W}_{12}^{220} &= \alpha_{12}^{220}\, S_2^4(L_1)\, S_2^4(L_2)\, P_0^4(L_3)\, \boldsymbol{w}_{12} \\ &= \alpha_{12}^{220}(4\, L_1 - 1)(4\, L_2 - 1)\, \boldsymbol{w}_{12}, \end{aligned} \qquad (138)$$

$$\begin{aligned} \boldsymbol{W}_3 = \boldsymbol{W}_{12}^{130} &= \alpha_{12}^{130}\, S_1^4(L_1)\, S_3^4(L_2)\, P_0^4(L_3)\, \boldsymbol{w}_{12} \\ &= \alpha_{12}^{130} \frac{1}{2}(4\, L_2 - 1)(4\, L_2 - 2)\, \boldsymbol{w}_{12}, \end{aligned} \qquad (139)$$

$$\begin{aligned} \boldsymbol{W}_4 = \boldsymbol{W}_{12}^{211} &= \alpha_{12}^{211}\, S_2^4(L_1)\, S_1^4(L_2)\, P_1^4(L_3)\, \boldsymbol{w}_{12} \\ &= \alpha_{12}^{211}(4\, L_1 - 1)4\, L_3\, \boldsymbol{w}_{12}, \end{aligned} \qquad (140)$$

$$\begin{aligned} \boldsymbol{W}_5 = \boldsymbol{W}_{12}^{121} &= \alpha_{12}^{121}\, S_1^4(L_1)\, S_2^4(L_2)\, P_1^4(L_3)\, \boldsymbol{w}_{12} \\ &= \alpha_{12}^{121}(4\, L_2 - 1)4\, L_3\, \boldsymbol{w}_{12}, \end{aligned} \qquad (141)$$

$$\begin{aligned} \boldsymbol{W}_6 = \boldsymbol{W}_{23}^{031} &= \alpha_{23}^{031}\, P_0^4(L_1)\, S_3^4(L_2)\, S_1^4(L_3)\, \boldsymbol{w}_{23} \\ &= \alpha_{23}^{031} \frac{1}{2}(4\, L_2 - 1)(4\, L_2 - 2)\, \boldsymbol{w}_{23}, \end{aligned} \qquad (142)$$

$$\begin{aligned} \boldsymbol{W}_7 = \boldsymbol{W}_{23}^{022} &= \alpha_{23}^{022}\, P_0^4(L_1)\, S_2^4(L_2)\, S_2^4(L_3)\, \boldsymbol{w}_{23} \\ &= \alpha_{23}^{022}(4\, L_2 - 1)(4\, L_3 - 1)\, \boldsymbol{w}_{23}, \end{aligned} \qquad (143)$$

$$\boldsymbol{W}_8 = \boldsymbol{W}_{23}^{013} = \alpha_{23}^{013} \, P_0^4(L_1) \, S_1^4(L_2) \, S_3^4(L_3) \, \boldsymbol{w}_{23}$$
$$= \alpha_{23}^{013} \frac{1}{2}(4\,L_3 - 1)(4\,L_3 - 2)\,\boldsymbol{w}_{23}, \tag{144}$$

$$\boldsymbol{W}_9 = \boldsymbol{W}_{23}^{121} = \alpha_{23}^{121} \, P_1^4(L_1) \, S_2^4(L_2) \, S_1^4(L_3) \, \boldsymbol{w}_{23}$$
$$= \alpha_{23}^{121}(4\,L_1)(4\,L_2 - 1)\,\boldsymbol{w}_{23}, \tag{145}$$

$$\boldsymbol{W}_{10} = \boldsymbol{W}_{23}^{112} = \alpha_{23}^{112} \, P_1^4(L_1) \, S_1^4(L_2) \, S_2^4(L_3) \, \boldsymbol{w}_{23}$$
$$= \alpha_{23}^{112}(4\,L_1)(4\,L_3 - 1)\,\boldsymbol{w}_{23}, \tag{146}$$

$$\boldsymbol{W}_{11} = \boldsymbol{W}_{31}^{103} = \alpha_{31}^{103} \, S_1^4(L_1) \, P_0^4(L_2) \, S_3^4(L_3) \, \boldsymbol{w}_{31}$$
$$= \alpha_{31}^{103} \frac{1}{2}(4\,L_3 - 1)(4\,L_3 - 2)\,\boldsymbol{w}_{31}, \tag{147}$$

$$\boldsymbol{W}_{12} = \boldsymbol{W}_{31}^{202} = \alpha_{31}^{202} \, S_2^4(L_1) \, P_0^4(L_2) \, S_2^4(L_3) \, \boldsymbol{w}_{31}$$
$$= \alpha_{31}^{202}(4\,L_1 - 1)(4\,L_3 - 1)\,\boldsymbol{w}_{31}, \tag{148}$$

$$\boldsymbol{W}_{13} = \boldsymbol{W}_{31}^{301} = \alpha_{31}^{301} \, S_3^4(L_1) \, P_0^4(L_2) \, S_1^4(L_3) \, \boldsymbol{w}_{31}$$
$$= \alpha_{31}^{301} \frac{1}{2}(4\,L_1 - 1)(4\,L_1 - 2)\,\boldsymbol{w}_{31}, \tag{149}$$

$$\boldsymbol{W}_{14} = \boldsymbol{W}_{31}^{112} = \alpha_{31}^{112} \, S_1^4(L_1) \, P_1^4(L_2) \, S_2^4(L_3) \, \boldsymbol{w}_{31}$$
$$= \alpha_{31}^{112}(4\,L_2)(4\,L_3 - 1)\,\boldsymbol{w}_{31}, \tag{150}$$

$$\boldsymbol{W}_{15} = \boldsymbol{W}_{31}^{211} = \alpha_{31}^{211} \, S_2^4(L_1) \, P_1^4(L_2) \, S_1^4(L_3) \, \boldsymbol{w}_{31}$$
$$= \alpha_{31}^{211}(4\,L_2)(4\,L_1 - 1)\,\boldsymbol{w}_{31}. \tag{151}$$

The vector shape functions in 3D can be constructed as the extension of the above presented 2D realization. Three-dimensional zeroth order edge shape functions can be constructed as,

$$\boldsymbol{W}_1 = l_1(N_1\nabla N_2 - N_2\nabla N_1)\delta_1, \tag{152}$$

$$\boldsymbol{W}_2 = l_2(N_2\nabla N_3 - N_3\nabla N_2)\delta_2, \tag{153}$$

$$\boldsymbol{W}_3 = l_3(N_3\nabla N_1 - N_1\nabla N_3)\delta_3, \tag{154}$$

$$\boldsymbol{W}_4 = l_4(N_1\nabla N_4 - N_4\nabla N_1)\delta_4, \tag{155}$$

$$\boldsymbol{W}_5 = l_5(N_2\nabla N_4 - N_4\nabla N_2)\delta_5, \tag{156}$$

$$\boldsymbol{W}_6 = l_6(N_3\nabla N_4 - N_4\nabla N_3)\delta_6. \tag{157}$$

Here l_i (Fig. 26) is the length of the edges and it is used to normalize the edge shape function according to (104). The value of δ_i is also equal to ± 1 depending on whether the local direction of the edge is the same as the global direction or opposite. The edge definition employed in my analysis can be seen in Fig. 26.

If the approximation of the vector function $\boldsymbol{T}$ is known along the edges of the mesh, then (98) can be used to interpolate the function anywhere and in linear case $k = 6$.

To construct higher order vector basis functions, the points of interpolation polynomials are arranged in a pyramid format to build an applicable numbering scheme (I,J,K,L) and $I,J,K,L = 0,1,\cdots,n + 2$, where n is the order of the element. The illustration of numbering scheme in 3D is not easy, but it can be construct as follows. Let us imagine the same numbering scheme on the triangular facets of the tetrahedron as in Fig. 31-Fig. 33 and the integers I, J, K and L can be set up according to the facets.

The vector shape functions of order n are given as,

$$\boldsymbol{W}_{12}^{IJKL} = \alpha_{12}^{IJKL}\, S_I^{n+2}(L_1)\, S_J^{n+2}(L_2)\, P_K^{n+2}(L_3)\, P_L^{n+2}(L_4)\, \boldsymbol{w}_{12}, \tag{158}$$

$$\boldsymbol{W}_{23}^{IJKL} = \alpha_{23}^{IJKL}\, P_I^{n+2}(L_1)\, S_J^{n+2}(L_2)\, S_K^{n+2}(L_3)\, P_L^{n+2}(L_4)\, \boldsymbol{w}_{23}, \tag{159}$$

$$\boldsymbol{W}_{31}^{IJKL} = \alpha_{31}^{IJKL}\, S_I^{n+2}(L_1)\, P_J^{n+2}(L_2)\, S_K^{n+2}(L_3)\, P_L^{n+2}(L_4)\, \boldsymbol{w}_{31}, \tag{160}$$

$$\boldsymbol{W}_{14}^{IJKL} = \alpha_{14}^{IJKL}\, S_I^{n+2}(L_1)\, P_J^{n+2}(L_2)\, P_K^{n+2}(L_3)\, S_L^{n+2}(L_4)\, \boldsymbol{w}_{14}, \tag{161}$$

$$\boldsymbol{W}_{24}^{IJKL} = \alpha_{24}^{IJKL}\, P_I^{n+2}(L_1)\, S_J^{n+2}(L_2)\, P_K^{n+2}(L_3)\, S_L^{n+2}(L_4)\, \boldsymbol{w}_{24}, \tag{162}$$

$$\boldsymbol{W}_{34}^{IJKL} = \alpha_{34}^{IJKL}\, P_I^{n+2}(L_1)\, P_J^{n+2}(L_2)\, S_K^{n+2}(L_3)\, S_L^{n+2}(L_4)\, \boldsymbol{w}_{34}. \tag{163}$$

The parameters denoted by α are normalization factors, which must have the value such that the line integral of the vector shape function $\boldsymbol{W}_{ab}^{IJKL}$ is equal to 1 on the edge pointing from node a to node b.

The number of edge shape functions when defining the nth order family is

$$k = \frac{(n + 1)(n + 3)(n + 4)}{2}. \tag{164}$$

For each interpolation point on the edge, there is one corresponding vector shape function, which means $6(n + 1)$ functions. For each interpolation point on the face of a tetrahedron there are three vector functions, but one of them is depending on the other two and it must be discarded, finally there are $4n(n + 1)$ vector shape functions defined on the four facets. For each interpolation points inside the element there are six basis functions. A 3D vector has only three degree of freedom, that is why three vector basis functions must be discarded resulting in $n(n - 1)(n + 1)/2$ vector basis functions. Totally, there are $6(n + 1) + 4n(n + 1) + n(n - 1)(n + 1)/2 = (n + 1)(n + 3)(n + 4)/2$ vector shape functions. There are $k = 6$, $k = 20$ and $k = 45$ shape functions for $n = 0$, $n = 1$ and $n = 2$, respectively.

As an example, the following vector shape functions can be set up when $n = 1$, $\boldsymbol{W}_{12}^{2100}$, $\boldsymbol{W}_{12}^{1200}$, $\boldsymbol{W}_{23}^{0210}$, $\boldsymbol{W}_{23}^{0120}$, $\boldsymbol{W}_{31}^{1020}$, $\boldsymbol{W}_{31}^{2010}$, $\boldsymbol{W}_{14}^{2001}$, $\boldsymbol{W}_{14}^{1002}$, $\boldsymbol{W}_{24}^{0201}$, $\boldsymbol{W}_{24}^{0102}$, $\boldsymbol{W}_{34}^{0021}$, $\boldsymbol{W}_{34}^{0012}$, $\boldsymbol{W}_{12}^{1110}$, $\boldsymbol{W}_{31}^{1110}$, $\boldsymbol{W}_{14}^{1011}$, $\boldsymbol{W}_{34}^{1011}$, $\boldsymbol{W}_{23}^{0111}$, $\boldsymbol{W}_{24}^{0111}$, $\boldsymbol{W}_{12}^{1101}$, $\boldsymbol{W}_{24}^{1101}$.

References

[1] Kuczmann M, Iványi A: *The Finite Element Method in Magnetics*, Academic Press, Budapest, 2008.