NUMERICAL ANALYSIS OF FUNCTIONALLY GRADED PRESSURE VESSELS USING STRESS FUNCTIONS AND NEURAL NETWORKS

Dávid Gönczi 厄

senior lecturer, Institute of Applied Mechanics, University of Miskolc 3515 Miskolc, Miskolc-Egyetemváros, e-mail: david.gonczi@uni-miskolc.hu

Abstract

The purpose of this paper is to study the displacements and stresses within radially graded spherical pressure vessels. The body of the sphere is made of a functionally graded material, the material properties are arbitrary functions of the radial coordinate. Our aim is to determine the stress state of the sphere using stress functions while the body is subjected to pressure and prescribed temperature field. The results are used to train a neural network to predict the maximum equivalent stress for the design of the pressure vessel.

Keywords: FGM, sphere, thermomechanics, neural network

1. Introduction

As technology advances rapidly, the demand for new, advanced materials with special material properties and behaviour is increasing. Engineers across various fields are exploring applications for these engineered materials instead of using pure metals and more traditional materials due to their material limitations. For example, a lot of problems require materials that are both hard, heat-resistant, or ductile. To improve these properties, metals are combined with other metals or non-metals, often forming composite materials. These advanced materials offer superior performance and lower weight compared to their individual components. However, their use is limited by delamination, particularly in high-temperature environments with different thermal expansion rates. To address the weaknesses of composites, Japanese researchers introduced Functionally Graded Materials (FGMs) in the mid-1980s during a hypersonic spacecraft project. FGMs feature a gradual change in composition and structure, resulting in gradual changes in their material properties and eliminating sharp interfaces, thus reducing failure risks.

Numerous studies have addressed the mechanics of functionally graded materials (FGMs) from various perspectives. Several books provide solutions to linear elastic problems in non-homogeneous bodies, such as those found in (Hetnarski and Eslami, 2010; Lekhnitskii, 1981; Noda et al., 2000). A variety of papers have proposed analytical, semi-analytical, and numerical methods to solve thermomechanical issues in hollow spheres, cylinders, beams, and disks. Noda et al. (Obata and Noda, 1994; Kim and Noda, 2002) examined one-dimensional steady-state thermal stress problems in isotropic functionally graded hollow circular cylinders and spheres, employing the perturbation method, multilayered approach, and Green's functions. Chen and Lin (Chen and Lin, 2008) conducted elastic analyses on thick cylinders and spherical pressure vessels composed of FGMs, where the material properties varied exponentially along the radial direction. Analytical solutions for radial, tangential, and effective stresses in thick spherical pressure vessels made from FGMs, subjected to axisymmetric

mechanical and thermal loads, were developed by Nayak et al. (Nayak et al., 2011) and Bayat et al. (Bayat et al., 2011). In their work, the material properties were modelled as a power-law function of the radial coordinate, while Poisson's ratio remained constant. Pen and Li (Pen and Li, 2009) explored the steady-state thermoelastic behaviour of isotropic radially graded disks with arbitrary radial non-homogeneity, reducing the numerical solution to a Fredholm integral equation. There are a lot of papers, such as (Gönczi and Ecsedi, 2015; Kiss, 2020; Gönczi, 2019; Kiss, 2014), that investigate the stress distribution of heterogeneous structural components.

In this paper radially graded spherical bodies are investigated, in which the material properties are arbitrary functions of the radial coordinate. The temperature field is a given which is the thermal loading of the problem. We have constant pressure exerted on the boundary surfaces of the sphere. Additionally we consider the case, when kinematic boundary conditions are given. Our aim is to present a method to solve these problems using stress functions. Then we would like to train neural networks with the results coming from these calculations. Figure 1 shows the sketch of the problem, the mechanical loads are p_1 and p_2 (constant pressure values) while T(r) denotes the temperature field.



Figure 1. The sketch of the problem.

2. The formulation of the problem

Let us consider functionally graded spherical pressure vessels examined in spherical coordinates (r, φ, ϑ) , where the material composition varies only in the radial direction. Consider a two-component, radially graded, axisymmetric structural element. If the material property values at the two boundary surfaces (inner radius *a*, outer radius *b*) are M_1 and M_2 , then the following normalized exponential distribution function can be used:

$$M_{FGM}(r) = [M_1 - M_2] \frac{r-a}{b-a} + M_2.$$
(1)

If we wish to incorporate temperature dependence in the future, the above equation can be extended by the

$$M(T) = P_0(P_1T^{-1} + 1 + P_1T + P_2T^2 + P_3T^3)$$
(2)
168

In this case, *M* represents the elastic modulus (*E*), Poisson's ratio (ν), the coefficient of linear thermal expansion (α), or the thermal conductivity (λ). Due to the axisymmetric nature of the problem and its sole dependence on the radial coordinate, the kinematic equations can be written in the following form

$$\varepsilon_r(r) = \frac{\partial u_r}{\partial r} = \frac{\mathrm{d}u_r(r)}{\mathrm{d}r}, \qquad \varepsilon_\varphi(r) = \varepsilon_\vartheta(r) = \frac{u_r}{r}.$$
⁽³⁾

In the previous equation u_r denotes the radial displacement, $\varepsilon_r, \varepsilon_{\varphi}, \varepsilon_{\vartheta}$ are the normal strains. Considering a linearly elastic, isotropic material, the normal stresses (from the constitutive equations) are

$$\sigma_{r}(r) = \frac{E(r)}{[1+\nu(r)][1-2\nu(r)]} \{ [1-\nu(r)]\varepsilon_{r}(r) + 2\nu(r)\varepsilon_{\varphi}(r) - \alpha(r)[1+\nu(r)]T(r) \},$$

$$\sigma_{\varphi}(r) = \sigma_{\vartheta(r)} = \frac{E(r)}{[1+\nu(r)][1-2\nu(r)]} \{ \nu(r)\varepsilon_{r}(r) + \varepsilon_{\varphi}(r) - \alpha(r)[1+\nu(r)]T(r) \}.$$
(4)

The equilibrium equation can be expressed as

$$\frac{\mathrm{d}\sigma_r}{\mathrm{d}r} + \frac{2(\sigma_r - \sigma_{\varphi})}{r} = 0.$$
⁽⁵⁾

Let's define the stress function F as

$$\sigma_r = F r^{-2}, \qquad \sigma_\varphi = 2r^{-1} \frac{\mathrm{d}F}{\mathrm{d}r}.$$
 (6)

After some manipulation of the previous system of equations we get the following system of differential equations

$$\frac{\mathrm{d}}{\mathrm{d}r} \begin{bmatrix} u\\ F \end{bmatrix} = \begin{bmatrix} -\frac{2\nu}{(1-\nu)}r^{-1} & \frac{(1-2\nu)(1+\nu)}{(1-\nu)E}r^{-2}\\ \frac{2E}{1-\nu} & \frac{2\nu}{1-\nu}r^{-1} \end{bmatrix} \begin{bmatrix} u\\ F \end{bmatrix} + \begin{bmatrix} \frac{1+\nu}{1-\nu}\\ \frac{2E}{1-\nu} \end{bmatrix} \alpha T.$$
(7)

We have essentially reformulated the boundary value problem as an initial value problem in a nontraditional sense. After this, we generate the actual initial values of the problem by combining two cases. Let us consider the problem, where we have two traction boundary conditions. In this case we have to solve the following two initial value problems:

- I. case: $u(r = a) = u_1$: arbitrary value, $F_I(r = a) = F_I = -p_1 a^2$,
- II. case: $u(r = a) = u_2 \neq u_1$: arbitrary value, $F_{II}(r = a) = F_I$.

This can be done, for example, by using the Runge-Kutta method. From these calculations, we need the value of our stress function at the outer boundary $F_I(b)$ and $F_{II}(b)$. Then, for the final (third) initial value problem to be solved, we use the actual F_I value from the previous calculations and the actual displacement value at the inner radius, which can be expressed as

$$u_{1_actual} = u_1 - \frac{u_2 - u_1}{F_{II}(b) - F_I(b)} (b^2 p_2 + F_I(b)).$$
(8)

The previous expression comes from the statement, that the actual displacement value of the problem can be given as the linear combination of the solutions of the first two initial value problems. This means that

$$\begin{bmatrix} u \\ F \end{bmatrix} = C_1 \begin{bmatrix} u \\ F \end{bmatrix}_I + C_2 \begin{bmatrix} u \\ F \end{bmatrix}_{II}$$
⁽⁹⁾

and we have two boundary conditions to calculate the unknown constants C_1, C_2 . Here we note, when there are kinematic boundary conditions, the initial value problems can be formulated as

I. case: $u(r = a) = u_1$: given value, $F_I(r = a) = F_I$: arbitrary value,

II. case: $u(r = a) = u_1$, $F_{II}(r = a)$: arbitrary value, but $F_{II} \neq F_I$.

Then we can use equations (9) to calculate the initial value of the displacement field in this case. When the displacement field is known, the stress distributions can be calculated according to eqs. (6). It is important to highlight, that the temperature field is arbitrary (and given), furthermore the method works when the material parameters are arbitrary functions of the temperature field.

3. Calculating the critical stresses using neural networks

To train a neural network, we need a large input dataset. We can use the previously presented method to calculate stresses and displacements as an input for the network. Let the maximum equivalent stress of our system be the critical stress. Our input data are the inner radius of the tank, the ratio of the outer to inner radius, the dataset describing the material distribution, the pressure and temperature of the stored internal medium. The output data will be the information relevant to the design of the pressure vessel. This could be, for example, the maximum reduced stress value to be compared with the failure criterion. This means that these problems can be classified as regression problems. We want to use deep neural networks with multiple hidden layers to calculate the chosen data for our problem.

The unit of a neural network is neuron, in which the input values (\mathbf{x}) are combined with weights (\mathbf{w}) and then with a bias value (b) they provide the input $(z = \mathbf{x} \cdot \mathbf{w} + b)$ for an activation function which produces the output of the neuron $\sigma(z)$. A neural network can be built from multiple layers of such neurons. This includes an input layer, which typically contains as many neurons as the amount of input data.



Figure 2. A Deep Neural Network (DNN).

The structure of the output layer is determined by the output data. For example, if we want to predict the maximum reduced stress value, we need 1 neuron in the output layer. Between these layers there are the hidden layers. The structure of these layers are hyperparameters of the system. An example of this is illustrated in Figure 2. Our most commonly used activation functions are the sigmoid function the hyperbolic tangent function (tanh) and the relu (Rectified Linear Unit), which are:

$$\sigma(z) = \frac{1}{1 + e^{-z}}, \ \tanh(z) = \frac{e^{2z} - 1}{e^{2z} + 1}, \ relu(z) = \max\{0, z\}.$$
(10)

The next important question is the determination of the weight functions associated with the neurons, which is referred to as training. Here, we use a training sample through which we calibrate the values of the weight functions according to some strategy. To do this, we need to define a cost/loss function $C(\mathbf{w}, b)$, which quantifies how well the output value approximates the actual value. We start with randomly chosen weight values and then iteratively adjust the weight values based on the training data (and the cost functions) to bring the results closer to the actual values.

One such method is the gradient descent technique. The basis of this method is that we have a cost function that depends on the variables (\mathbf{w}, b) . We calculate its gradient and shift the values in the appropriate direction, aiming to reduce our cost function. An important parameter here is the learning rate η (a small positive value). It needs to be set correctly because if it is too large, we might overshoot the sought minimum, while if it is too small, it will significantly slow down the algorithm. A commonly used variant of this method is the stochastic gradient descent technique, where calculations are performed not on the entire dataset (batch) but rather on a randomly selected smaller subset (mini-batch). The advantage of this variant is that it leads to more stable convergence. We then choose another subset and compute with it until we have gone through the entire training sample. At this point, we have completed what is known as an epoch. Another frequently used variant incorporates a momentum term to adjust the weights. Another efficient optimizer technique is the Adam (Adaptive Moment Estimation) optimization, which dynamically adjusts the rate of change of the parameters. This means that the learning rate adapts dynamically for each parameter. It keeps track of the learning rates for each parameter and remembers the gradients from the previous step, allowing it to make decisions based on the current and previous states. A potent version of Adam is Nadam, which combines Adam with the Nesterov technique, which tries to predict the future outcomes by introducing more parameters.

There are two important (basic) numerical problem during the training. The unstable gradient problem is a common issue that arises during the training of deep neural networks. Within this, the vanishing gradient problem occurs during backpropagation when the gradient of the loss function with respect to the weights becomes very small, causing the weights to be modified negligibly and slowing down convergence. The opposite extreme occurs when the weight values are large and the changes are significant at each step. To address this, we can choose from several techniques. The first is a proper weight initialization paired with a non-saturating activation function. These pairs can be the Glorot initialization with sigmoid, tanh, or softmax activation function compared to sigmoid or hyperbolic tangent functions, in which case it is advisable to set HE initialization. In the case of reduced stress predictions, the values are positive, allowing for the use of the simple relu function. A popular pairing is LeCun initialization functions). Furthermore, batch normalization and gradient clipping are also options to consider. Another commonly used technique is normalizing the data, mapping it to a range between 0 and 1.

To combat the other important problem of training deep neural networks, the overfitting phenomena, we can use regularization (L1 and/or L2), as well as dropout regularization. In simpler cases, we can also use early stopping, but this is generally not recommended, as it means that the training process does not complete.

When it comes to the training of the neural networks, we have a lot of hyperparameters to configure. There are a lot of strategies to do it. We used a simple grid search combined with zooming in methods to tune these parameters.

Let us consider a functionally graded material, which consists of a high-strength steel and a heatresistant ceramic components. The material properties are:

$$E_1 = 190 \text{GPa}, \alpha_1 = 12 \cdot 10^{-6} \text{K}^{-1}, \nu_1 = 0.3, \lambda_1 = 45 \frac{\text{W}}{\text{mK}},$$
$$E_2 = 330 \text{GPa}, \alpha_1 = 5 \cdot 10^{-6} \text{K}^{-1}, \nu_1 = 0.25, \lambda_1 = 6 \frac{\text{W}}{\text{mK}}.$$

The FGM of the sphere has the base parameters specified earlier based on equation (1). For this problem we loaded the tank with internal pressure and a constant temperature field, because the spherical pressure vessel is insulated at the outside surface and the temperature difference is negligible due to the smaller wall thickness range. For the tank failure condition, we chose the stress level associated with yielding. Therefore, we only recorded results that did not exceed 1000 MPa, as the theories used to solve the problem are only valid for small deformations. Additionally, we examined loads only up to a reduced stress level of 600 MPa for the reasons mentioned earlier. The parameter range we investigated was as follows:

 $\begin{aligned} &a = [0.5, 0.7, 0.9, 1.1, 1.3, 1.6, 2] \\ &b_{per} = [1.002, 1.006, 1.01, 1.03, 1.05, 1.75, 1.1, 1.2] \\ &m = [0.001, 0.01, 0.1, 0.5, 0.75, 1, 2, 3, 5, 10, 40, 100, 300] \\ &p_1 = [0.1, 1, 5, 10, 25, 50, 75, 100, 150, 200] \\ &T_1 = [0, 20, 40, 50, 75, 100, 150, 200, 300] \end{aligned}$

We obtain the outer radius in the form of $b = ab_{per}$, meaning that the wall thickness is a maximum of 20% of the inner radius. We kept the loading values within the conventional range. This produced 31,626 results out of a possible 65,500. In the remaining cases, the theory is not applicable, as failure has occurred earlier.

To solve this problem, we used Python programming language with the packages Panda, Sklearn and Tensorflow. The model has 5 input values and 1 output value. For the training metric we used mean average error (mae). We divided the dataset into two parts, a training set and a test set. The training set is further divided into a normal training set and a validation set to track overfitting during training. When it comes to the optimizers, nadam was the most efficient one. Dropout regularization produced inferior results (mostly worst accuracy and slower convergence) compared to normal dense structures. In table 1 a few solutions coming from several investigated neural networks are shown. The best results coming from 3-5 layers with 256-512 neurons. Either relu or selu activations can be used to calculate the maximum stresses. In this case, within the test range, the error magnitude was sufficiently small. The model was able to produce acceptable results even for parameter combinations outside the examined range. However, the farther we moved from the test range, the greater the error became. The results also show that in some regions, more data would have been needed during the training of the network, as the

error in these areas noticeably increased compared to regions where sufficient data was available. In table 1, the errors for the individual examples are given in percentage. For table 1, the investigate parameter combinations are:

$$\begin{split} &1: [5 \cdot 10^{-3}, 5.05 \cdot 10^{-3}, 0.1, 90, 300], 2: [6 \cdot 10^{-3}, 6.1 \cdot 10^{-3}, 0.5, 240, 100], \\ &3: [8 \cdot 10^{-3}, 12 \cdot 10^{-3}, 2, 190, 130], 4: [4 \cdot 10^{-2}, 4.1 \cdot 10^{-2}, 0.2, 200, 157], \\ &5: [5 \cdot 10^{-2}, 7 \cdot 10^{-2}, 50, 50, 150], 6: [0.2, 0.22, 2, 23, 55.0], 7: [0.5, 0.501, 0.5, 0.1, 300], \\ &8: [0.5, 0.505, 50, 50, 150], 9: [0.65, 0.76, 2, 23, 55], 10: [1.76, 2, 0.3, 70, 102], \\ &11: [4, 5, 2.5, 100, 156], 12: [8, 11, 4.4, 200, 188], 13: [8,9,0.6, 100, 111], \\ &14: [10, 12.3, 0.12, 250, 320], 15: [35, 38, 3.6, 110, 130]. \end{split}$$

Hidden	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Layers	1	2	5	4	5	0	'	0		10	11	12	15	14	15	
	50 epoch, mae: 10, test mae: 9.1															
16	25.6	1.5	-9	22.4	1.6	-12.9	-4.4	1.8	-12	19	2.7	8.4	4.5	21.1	20.1	
	100 epoch, mae: 9.65, test mae: 9.64															
	30.0	8.1	-5.8	35	-0.8	-9.4	1.3	0.1	-9.6	25.6	5.6	9.4	- 14.4	21.2	19.2	
32						50 epc	och, mae	e: 6.4, t	est mae	e: 6.28						
	17.6	-6.2	-8.5	5.4	1.6	3.7	-0.7	1.9	3.1	0.4	3.3	12	-8.8	17	-7.9	
	100 epoch, mae: 7.08, test mae: 7.08															
	25.7	1.8	-1.4	18.3	1.3	1.8	1.8	1.4	0.6	8.3	3.8	10.8	- 12.8	15.9	-1.4	
32-32						50 epo	ch, mae	: 7.86, 1	test ma	e: 7.83						
	22.5	7.0	-8	12.7	- 2.1	-1.3	3.5	1.3	-1.4	6.2	3.6	7.2	2.6	18.1	-0.7	
64		100 epoch, mae: 12.2, test mae:12.3														
	34.2	5.1	-5.7	35	- 0.1	-12	-0.3	0.4	9	21	0.2	10	12.9	20.1	4.5	
64-32	100 epoch, mae: 7.05, test mae: 6.6															
	13.9	-7.1	-6.3	-1	0.9	3.2	-0.9	1.1	1.6	-4.1	5.9	7.6	-0.7	12.6	-5.2	
	100 epoch, increased batch size, mae: 4.77, test mae: 4.6															
	17	0	2.9	7.5	-1.1	4.4	-0.2	0-9	2.5	-0.5	1.4	11.4	-7.9	13.1	30	
3 x					rel	u, 200 e	epoch, n	nae: 3.8	32, test	mae: 3.	82					
256	13.1	-1.6	-1.7	8	0.3	-3.1	-0.3	0.6	-2.9	3	4.3	11.7	5.3	14.6	-70	
4 x					se	lu, 200	epoch, i	mae: 3.	6, test 1	nae: 3.5	59					
512	13	1	-7.9	9	0.1	0.8	-1.7	0.5	-0.1	4.1	2.4	5.1	3.1	15.6	23	

Table 1. The results of the neural networks on specific parameter combinations.

Simulations were run to test the accuracy of the previously trained neural networks on new data sets. We picked the last two networks in table 1, the results were similar, although the one with more neurons and selu activation had a bit better accuracy. For example, on a new dataset given by the values

 R_1 : [0.7, 0.96, 1.55, 1.98], R_2 : [1.005, 1.032, 1.09, 1.125, 1.18],

m: [0.006, 0.02, 3.5, 43, 143], p_1 : [7, 54, 96, 145, 182], T_1 : [30, 85, 125, 210]

the average error was 3.3 MPa in case of 4 layers with 512 neurons in each layer. Outside of the training range, on the following data set

 R_1 : [0.03, 0.07, 0.15, 0.3], R_2 : [1.003, 1.042, 1.13, 1.3, 1.4],

m: [0.002, 0.03, 2.5, 33, 216], *p*₁: [3, 34, 75, 175, 300], *T*₁: [30, 125, 225, 350]

the average error was 16 MPa. This is one the advantages of neural networks, that they can provide somewhat useful data even outside its range.

4. Summary

A method was presented to determine the stress distribution and displacement field within a radially graded spherical body subjected to combined mechanical and thermal loads. The material properties and temperature field were arbitrary functions of the radial coordinate. Stress functions were used to formulate an initial value method and to solve the problem. Then a neural network was trained with datasets coming from the previously presented method. A numerical example was shown to investigate the accuracy of the neural networks. The results obtained with the neural network showed good agreement with the numerical solution of the original system of differential equations. Based on the data, it was found that the neural network produced usable outputs not only within the training domain but also beyond it, albeit with a reduction in accuracy.

References

- [1] Hetnarski, R. B., & Eslami, M. R. (2010). *Thermal Stresses Advanced Theory and Applications*. Springer, New York, USA. https://doi.org/10.1007/978-3-030-10436-8
- [2] Lekhnitskii, S. G. (1981). *Theory of Elasticity of an Anisotropic Body*. Mir Publishers, Moscow. https://doi.org/10.12677/ASS.2021.105163
- [3] Noda, N., Hetnarski, R. B., & Tanigawa, Y. (2000). *Thermal Stresses*. Lastran Corporation, Rochester, New York, USA. https://doi.org/10.1115/1.1349549
- [4] Obata, Y., & Noda N. (1994). Steady thermal stress in a hollow circular cylinder and a hollow sphere of a functionally gradient materials. *Thermal stress*, 17, 471–487. https://doi.org/10.1080/01495739408946273
- [5] Kim, K. S., & Noda, N. (2002). Green's function approach to unsteady thermal stresses in an infinite hollow cylinder of functionally graded material. *Acta. Mech.*, 156, 61–145. https://doi.org/10.1007/BF01176753
- [6] Chen, Y. Z., & Lin, X. Y. (2008). Elastic analysis for thick cylinders and spherical pressure vessels made of functionally graded materials. *Computational Materials Science*, 44, 581–587. https://doi.org/10.1016/j.commatsci.2008.04.018
- [7] Nayak, P., Mondal, S. C., & Nandi, A. (2011). Stress, Strain and displacement of a functionally graded thick spherical vessel. *International Journal of Engineering Science and Technology*, 3(4), 2660–2671. https://doi.org/10.12691/ajme-3-5-3
- [8] Bayat, Y., Ghannad, M., & Torabi, H. (2011). Analytical and numerical analysis for the FGM thick sphere under combined pressure and temperature loading. *Arch. Appl. Mech.*, 10, 229–242. https://doi.org/10.1007/s00419-011-0552-x
- [9] Pen, X., Li, X. (2009). Thermoelastic analysis of functionally graded annulus with arbitrary gradient. *Applied Mathematics and Mechanics (English Edition), 30*(10), 1211–1220. https://doi.org/10.1007/s10483-009-1001-7

- [10] Gönczi, D., & Ecsedi, I. (2015). Thermoelastic Analysis of Functionally Graded Hollow Circular Disk. Archive of Mechanical Engineering, 62(1), 5–18. https://doi.org/10.1515/meceng-2015-0001
- [11] Kiss, L. P. (2020). Nonlinear stability analysis of FGM shallow arches under an arbitrary concentrated radial force. *International Journal of Mechanics and Materials in Design*, 16(1), 91–108. https://doi.org/10.1007/s10999-019-09460-2
- [12] Gönczi, D. (2019). Analysis of a curved bimetallic beam. *Journal of Computational and Applied Mechanics*, 14(1/2), 41–51. https://doi.org/10.32973/jcam.2019.003
- [13] Kiss, L. P. (2014). In-plane buckling of rotationally restrained heterogeneous shallow arches subjected to a concentrated force at the crown point. *Journal of Computational and Applied Mechanics*, 9(2), 171–199. https://doi.org/10.32973/jcam.2014.009