Computer Assisted Mechanics and Engineering Sciences, 10: 163–176, 2003.

Topological optimization and inverse problems[†]

Lidia Jackowska-Strumiłło
Computer Engineering Department, Technical University of Łódź
Al. Politechniki 11, 90-924 Łódź, Poland

Jan Sokołowski
Institut Elie Cartan, Laboratoire de Mathématiques, Université Henri Poincaré
Nancy I, B.P. 239, 54506 Vandoeuvre lès Nancy Cedex, France

Antoni Żochowski
Systems Research Institute of the Polish Academy of Sciences
ul. Newelska 6, 01-447 Warszawa, Poland

(Received October 25, 2002)

The topological derivative of an arbitrary shape functional is introduced in [29] for 2D elasticity. The optimality conditions for general shape optimization problems are established in [30] using the shape variations including boundary and topology variations. The topology variations result in the presence of topological derivatives in the necessary conditions for optimality. In the present paper we derive the necessary optimality conditions for a class of shape optimization problems. The topological variations of shape functionals are used for the numerical solution of inverse problems. The numerical method uses neural networks. The results of computations confirm the convergence of the method.

Keywords: topological derivative, shape optimization, optimality conditions, artificial neural network, shape inverse problem, nucleation of openings

1. Introduction

In classical theory of shape optimization the first order necessary optimality conditions account for boundary variations of an optimal domain. On the other hand the relaxed formulation based on homogenization technique is used [1, 2, 19] in the topology optimization of energy functionals, the so called compliances in structural optimization. For such a formulation the coefficients of an elliptic operator are selected in an optimal way and the resulting optimal design takes the form of a composite microstructure rather than any geometrical domain. On the other hand the so-called *bubble method* is used for the topology optimization in structural mechanics [6, 23] which leads to numerical methods. We refer also to [7, 14, 18] for the related results. Further applications in mechanics can be found in [5, 15–17]. It seems that in the literature on the subject there is a lack of general method or technique that can be applied in the process of optimization of an arbitrary shape functional for simultaneous boundary and topology variations. Such an approach would be very useful for numerical solution of e.g. optimum design problems in structural mechanics. In the paper [25] the so called topological derivative (TD) of an arbitrary shape functional is introduced. Such a derivative is evaluated by an application of the asymptotic analysis with respect to geometrical singularities of domains [13], for a class of elliptic equations including 2D elasticity

[†]This is an extended version of a paper presented at the conference *OPTY-2001, Mathematical and Engineering Aspects of Optimal Design of Materials and Structures*, Poznań, Poland, August 27–29, 2001.

system [25] and 3D elasticity system [28]. TD determines whether a change of topology by creation of a small hole, or in similar setting of a small inclusion, at a given point $\mathbf{x} \in \Omega$ would result in improving the value of the given shape functional. In general, the form of topological derivatives is established [25, 28] by using the asymptotic expansions of solutions to elliptic systems obtained by the method of matched (or compound) asymptotics. In the case of cavities in the form of two dimensional circles or three dimensional balls (in the case of Laplace equations in the form of balls in $\mathbb{R}^n$ for an arbitrary space dimension $n \geq 2$) the constructive results are obtained [25, 28] by using the shape calculus combined with the asymptotic expansions of solutions.

In the present paper the approach of [25] is extended to the case of a finite number of circular holes treated by means of TD combined with simultaneous boundary variations by an application of the speed method [24] applied to Frechet differentiable shape functionals. Therefore, the general set of optimality conditions is established for a class of shape optimization problems in more general setting compared to the classical theory [4, 24].

To deal with various types of domain modification we introduce the following general notation for different types of variations of shape functionals and of solutions to partial differential equations:

Shape derivative – is used in order to determine the variations of solutions to boundary value problems resulting from the boundary variations of geometrical domains. In particular, first the Frechet differentiability of shape functionals is established, and then the speed method is applied to determine the shape derivatives. We refer to [24] and the recent book [4] for general description of the speed method and the related results on Frechet differentiability of shape functionals.

Topological derivative – also *topological differential* accounts for variations of shape functionals resulting from the emerging of one or several small holes or cavities in the interior of the geometrical domain. We refer to Section 2.1 for description of topological derivatives of shape functionals in the case of the Laplace equation.

Domain differential – unifies the influence on shape functionals of boundary variations and, at the same time, of the nucleation of internal holes or cavities. We refer to Section 2.3 for the more detailed discussion in the case of the Laplace equation.

In Section 2 optimality conditions are established for general shape modifications, including shape and topological variations. In Section 3 topological derivatives are given for shape functionals associated with an inverse problem. The inverse problem is considered for the identification of small inclusions. The numerical solution of inverse problems uses neural networks. The numerical procedure is described in Sections 4,5. The results for test examples are provided in Section 5.

2. Shape and topology modification

In this section we shall recall certain results concerning the shape and topology derivatives and formulae for their computation. Let us consider a bounded domain $\Omega \subset \mathbb{R}^2$ with the boundary consisting of a finite number of smooth arcs and vertices. This boundary is divided into three parts: Γ^D, Γ^N and Γ^V. The part Γ^V will be subject to variation, and for simplicity we assume on it the homogeneous Neumann condition. As a result we have the model problem

$$\begin{aligned} &\Delta u = f && \text{in } \Omega, \\ &u = g_D && \text{on } \Gamma^D, \\ &\frac{\partial u}{\partial \mathbf{n}} = g_N && \text{on } \Gamma^N, \\ &\frac{\partial u}{\partial \mathbf{n}} = 0 && \text{on } \Gamma^V, \end{aligned} \tag{1}$$

with $f \in H^2(\Omega)$, $g_D \in H^{7/2}(\Gamma_D)$, $g_N \in H^{5/2}(\Gamma_N)$. The goal functional we shall consider in the following general form:

$$J(\Omega) = \int_\Omega [F(u) + G(\nabla u)]\, \mathrm{d}x, \tag{2}$$

where $F(t)$, $t \in \mathbb{R}$, and $G(\mathbf{q})$, $\mathbf{q} \in \mathbb{R}^2$, are differentiable functions satisfying additional condition

$$|G(\mathbf{q})| \leq \Lambda + \Lambda|\mathbf{q}|^4 \qquad (\Lambda - \text{generic constant}),$$

assuring the existence of (2).

2.1. Shape derivative

First we shall define the shape derivative of J. To this goal we consider the transformation

$$T(\tau, \cdot) : \mathbb{R}^2 \mapsto \mathbb{R}^2$$

defined by

$$T(\tau, \mathbf{x}) = \mathbf{x} + \tau\boldsymbol{\Theta}(\mathbf{x}), \tag{3}$$

where $\boldsymbol{\Theta}(\mathbf{x})$ is a C^4-smooth vector field on $\mathbb{R}^2$ with the support concentrated on the part of the tubular neighbourhood of $\Gamma^V + B(\delta)$, namely the set

$$U = (\Gamma^V + B(\delta)) \setminus (\Gamma^N \cup \Gamma^D + B(2\delta)), \tag{4}$$

for some fixed small $\delta > 0$. Thus the vertices of $\partial\Omega$ do not move and the angles between adjacent arcs do not change.

Such a transformation has for $|\tau|$ small enough the following properties:

- $T(\tau, \cdot)$ is an identity on $\Omega \setminus U$;
- $T(\tau, \cdot)$ is a bijection of U onto itself.

The geometry of the problem may be seen in Fig. 1.

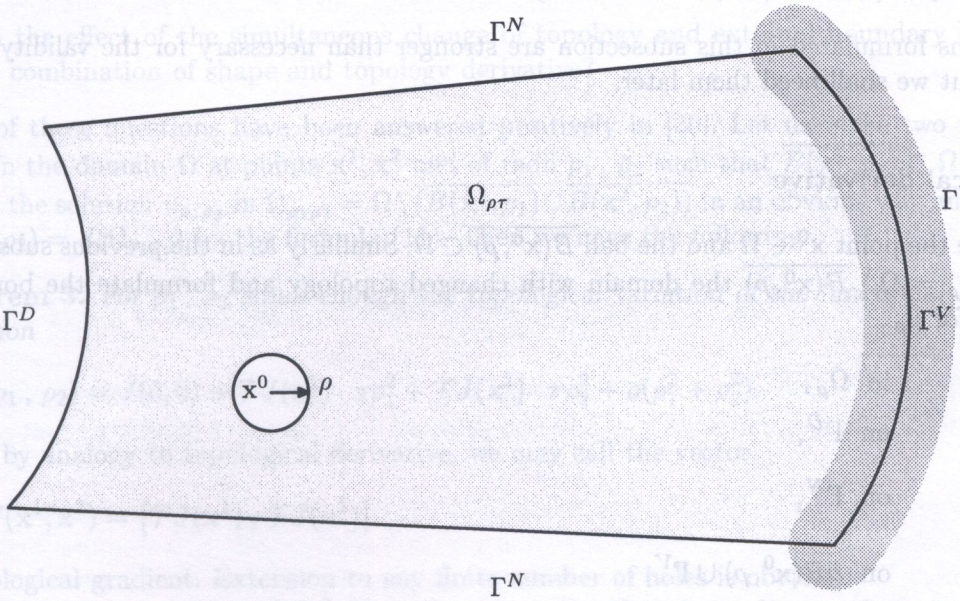


Fig. 1. The configuration of the internal hole and variable boundary

Next we define the boundary value problem for u_τ, similarly to Eq. (1):

$$\begin{aligned} &\Delta u_\tau = f && \text{in } \Omega_\tau, \\ &u_\tau = g_D && \text{on } \Gamma^D, \\ &\frac{\partial u_\tau}{\partial \mathbf{n}} = g_N && \text{on } \Gamma^N, \\ &\frac{\partial u_\tau}{\partial \mathbf{n}} = 0 && \text{on } \Gamma_\tau^V, \end{aligned} \tag{5}$$

where $\Omega_\tau = T(\tau, \Omega)$, $\Gamma_\tau^V = T(\tau, \Gamma^V)$. In accordance with this we take

$$J(\tau) = J(\Omega_\tau) = \int_{\Omega_\tau} [F(u_\tau) + G(\nabla u_\tau)] \, \mathrm{d}x, \tag{6}$$

where $J(0) = J(\Omega)$.

Now we may define the *shape derivative* of $J(\Omega_\tau)$ at Ω ($\tau = 0$) in the direction of the field $\boldsymbol{\Theta}$,

$$\mathcal{S}J(\Omega; \boldsymbol{\Theta}) = \lim_{\tau \to 0} \frac{1}{\tau} [J(\Omega_\tau) - J(\Omega)]. \tag{7}$$

The following result is well known [24]:

Theorem 1. *Under the assumptions stated above the shape derivative exists and is given by the formula*

$$\mathcal{S}J(\Omega; \boldsymbol{\Theta}) = \int_{\Gamma^V} [F(u) + G(\nabla u) + (\nabla u \cdot \nabla v)] \, (\boldsymbol{\Theta} \cdot \mathbf{n}) \, \mathrm{d}s, \tag{8}$$

where v is the adjoint state satisfying the following boundary value problem: find $v \in H_D^1(\Omega)$ such that

$$-\int (\nabla v \cdot \nabla \phi) \, \mathrm{d}x = \int_\Omega [F_u(u)\phi + (\nabla_{\mathbf{q}} G(\nabla u) \cdot \nabla \phi)] \, \mathrm{d}x \tag{9}$$

for all $\phi \in H_D^1(\Omega) = \{ \phi \in H^1(\Omega) \mid \phi = 0 \text{ on } \Gamma^D \}$.

The assumptions formulated in this subsection are stronger than necessary for the validity of the above result, but we shall need them later.

2.2. Topological derivative

Let us now take the point $\mathbf{x}^0 \in \Omega$ and the ball $\overline{B(\mathbf{x}^0, \rho)} \subset \Omega$. Similarly as in the previous subsection, we define by $\Omega_\rho = \Omega \setminus \overline{B(\mathbf{x}^0, \rho)}$ the domain with changed topology and formulate the boundary value problem for u_ρ,

$$\begin{aligned} &\Delta u_\rho = f && \text{in } \Omega_\rho, \\ &u_\rho = g_D && \text{on } \Gamma^D, \\ &\frac{\partial}{\partial \mathbf{n}} u_\rho = g_N && \text{on } \Gamma^N, \\ &\frac{\partial}{\partial \mathbf{n}} u_\rho = 0 && \text{on } \partial B(\mathbf{x}^0, \rho) \cup \Gamma^V, \end{aligned} \tag{10}$$

see again Fig. 1.

The appropriate value of the functional (2) has now the form

$$J(\rho) = J(\Omega_\rho) = \int_{\Omega_\rho} [F(u_\rho) + G(\nabla u_\rho)]\, \mathrm{d}x. \tag{11}$$

The *topological derivative* of the functional J is defined as a limit

$$\mathcal{T}J(\mathbf{x}^0) = \lim_{\rho \to 0+} \frac{J(\Omega_\rho) - J(\Omega)}{\pi \rho^2}. \tag{12}$$

In [30] there has been proved the following theorem.

Theorem 2. *The topological derivative $\mathcal{T}J(\mathbf{x}^0)$ exists and is given by the formula*

$$\mathcal{T}J(\mathbf{x}^0) = -\left[F(u) + \hat{G}(\nabla u) + fv + 2(\nabla u \cdot \nabla v)\right]_{\mathbf{x}=\mathbf{x}^0}, \tag{13}$$

where v is an adjoint variable defined by Eq. (9) and

$$\hat{G}(\nabla u(\mathbf{x}^0)) = \frac{1}{2\pi} \int_0^{2\pi} G\left(a \sin^2\theta - b \sin\theta \cos\theta\,,\ -a \sin\theta \cos\theta + b \cos^2\theta\right) \mathrm{d}\theta. \tag{14}$$

Here

$$a = \frac{\partial u}{\partial x_1}(\mathbf{x}^0), \qquad b = \frac{\partial u}{\partial x_2}(\mathbf{x}^0),$$

so that $[a, b] = \nabla u(\mathbf{x}^0)$.

2.3. Simultaneous topology and shape modification

Having defined both the shape and topological derivatives it is natural to raise the following questions:

- is the topological derivative additive, i.e. can we use it for approximation of the joint effect of the finite number of holes simultaneously?
- can the effect of the simultaneous change of topology and external boundary be described by the combination of shape and topology derivative?

Both of these questions have been answered positively in [30]. Let us make two non-overlapping holes in the domain Ω at points $\mathbf{x}^1, \mathbf{x}^2$ and of radii ρ_1, ρ_2 such that $\overline{B(\mathbf{x}^1, \rho_1)} \subset \Omega$, $\overline{B(\mathbf{x}^2, \rho_2)} \subset \Omega$. Define the solution $u_{\rho_1\rho_2}$ in $\Omega_{\rho_1\rho_2} = \Omega \setminus (\overline{B(\mathbf{x}^1, \rho_1)} \cup \overline{B(\mathbf{x}^2, \rho_2)})$ in an obvious way and the functional $J(\rho_1, \rho_2) = J(\Omega_{\rho_1\rho_2})$ by the formula (11). Then we have the following.

Theorem 3. *For ρ_1, ρ_2 small enough the topological variation of the functional J has the representation*

$$J(\rho_1, \rho_2) = J(0,0) + \mathcal{T}J(\mathbf{x}^1) \cdot \pi\rho_1^2 + \mathcal{T}J(\mathbf{x}^2) \cdot \pi\rho_2^2 + o(\rho_1^2 + \rho_2^2). \tag{15}$$

Thus, by analogy to topological derivative, we may call the vector

$$\mathcal{T}J(\mathbf{x}^1, \mathbf{x}^2) = [\mathcal{T}J(\mathbf{x}^1), \mathcal{T}J(\mathbf{x}^1)]$$

a topological gradient. Extension to any finite number of holes is obvious.

Now let us make a hole $B(\mathbf{x}^0, \rho)$ at the point $\mathbf{x}^0 \in \Omega$ and simultaneously move the part of the boundary Γ^V by means of the transformation $T(\tau, \cdot)$ defined by the field $\boldsymbol{\Theta}(\mathbf{x})$, as in Section 2.1.

Our assumptions ensure that for ρ, τ and δ small enough the neighbourhood of the hole is not transformed. The domain $\Omega_{\tau\rho}$ is defined in an obvious way (hole at $\mathbf{x}^0$ and Γ^V transformed to Γ^V_τ) and the corresponding $u_{\rho\tau}$ satisfies homogeneous boundary conditions on $\partial B(\mathbf{x}^0, \rho)$ and Γ^V_τ. The functional J is defined as in Eq. (11)

$$J(\rho, \tau) = J(\Omega_{\rho\tau}) = \int_{\Omega_{\rho\tau}} [F(u_{\rho\tau}) + G(\nabla u_{\rho\tau})] \, \mathrm{d}x. \tag{16}$$

Using this notation we are able to formulate the theorem justifying the combined use of both topological and shape derivatives.

Theorem 4. *For $|\tau|, \rho > 0$ small enough the variation of the functional J has the representation*

$$J(\rho, \tau) = J(0, 0) + \mathcal{T}J(\mathbf{x}^0) \cdot \pi\rho^2 + \mathcal{S}J(\Omega; \boldsymbol{\Theta}) \cdot \tau + o(|\tau| + \rho^2). \tag{17}$$

This theorem justifies the name *domain differential* for the expression

$$\mathcal{T}J(\mathbf{x}^0) \cdot \pi\rho^2 + \mathcal{S}J(\Omega; \boldsymbol{\Theta}) \cdot \tau.$$

The validity of the formulae given by Theorems 3 and 4 allow us to formulate new, stronger optimality conditions, see [30], as well as to approximate the values of $J(\rho_1, \rho_2)$ or $J(\tau, \rho)$ for finite ρ, τ. We shall use the last fact in the next sections.

The assumption about homogeneous Neumann condition on Γ^V is used here for simplicity and may be omitted. However, the condition concerning the constancy of angles at the vertices of $\partial\Omega$ and positions of these vertices is important in Theorem 4. The strong regularity of data is also used in proof.

3. Numerical example of shape functionals

We consider the following test examples.

We consider four boundary value problems defined in the same domain $\Omega = (0, 1) \times (0, 1)$. It means, that for $i = 1, 2, 3, 4$

$$\Delta u_i = 0 \quad \text{in } \Omega.$$

These problems differ with respect to the boundary conditions. For $i = 1$ they have the form

$$u_1 = 1 \quad \text{on } \{0\} \times \left(\frac{1}{3}, \frac{2}{3}\right); \qquad u_1 = 0 \quad \text{on } \{1\} \times (0, 1); \qquad \frac{\partial u_1}{\partial n} = 0 \quad \text{otherwise.}$$

For $i = 2, 3, 4$ they are obtained from the above conditions applying the successive rotation by the angle $\pi/2$.

The shape functionals $\mathcal{J}_j = \mathcal{J}_j(\Omega)$ are defined as follows: for $j = 1, \ldots, 12$, $i = 1, 2, 3, 4$.

$$\mathcal{J}_{\{1+3(i-1)\}} = \int_\Omega u_i^2 \, \mathrm{d}\Omega, \qquad \mathcal{J}_{\{2+3(i-1)\}} = \int_\Omega \left(\frac{\partial u_i}{\partial x_1}\right)^2 \mathrm{d}\Omega, \qquad \mathcal{J}_{\{3+3(i-1)\}} = \int_\Omega \left(\frac{\partial u_i}{\partial x_2}\right)^2 \mathrm{d}\Omega$$

In the domain $\Omega_\rho = \Omega \setminus \overline{B(\mathbf{y}, \rho)}$, $\mathbf{y} = (y_1, y_2)$, see Fig. 2, we add the homogeneous Neumann boundary conditions on the boundary Γ_ρ of the ball $B(\mathbf{y}, \rho)$. For clarity of further derivations let us denote

$$J_{ui}(\rho) = \int_{\Omega_\rho} u_{\rho i}^2 \, \mathrm{d}x, \qquad i = 1, 2, 3, 4,$$

$$J_{gik}(\rho) = \int_{\Omega_\rho} \left(\frac{\partial u_{\rho i}}{\partial x_k}\right)^2 \mathrm{d}x, \qquad k = 1, 2.$$

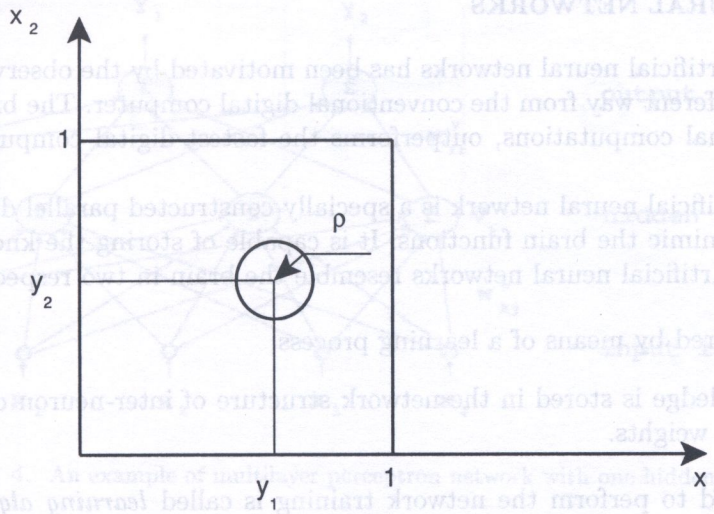


Fig. 2. The parameters describing opening (inclusion)

The topological derivatives of these shape functionals are obtained from Theorem 2 by direct computation of the function $\hat{G}$. To this end we observe that if we take for example $G(\nabla u) = (\partial u/\partial x_1)^2$, then the integrand in Eq. (14) takes on the form

$$G(\nabla u) = \left(\sin^2\theta \frac{\partial u}{\partial x_1}(\mathrm{y}) - \sin\theta\cos\theta \frac{\partial u}{\partial x_2}\right)^2 (\mathrm{y}).$$

As a result, for $i = 1,2,3,4$ and $k = 1,2$ we obtain formulae

$$\begin{aligned}
\mathcal{T}J_{ui}(\mathrm{y}) &= -\frac{1}{2}\left[u_i^2 + 4(\nabla u_i \cdot \nabla w_i)\right](\mathrm{y}),\\
\mathcal{T}J_{gi1}(\mathrm{y}) &= -\frac{1}{2}\left[\frac{3}{2}\left(\frac{\partial u_i}{\partial x_1}\right)^2 + \frac{1}{2}\left(\frac{\partial u_i}{\partial x_2}\right)^2 + 4(\nabla u_i \cdot \nabla v_{i1})\right](\mathrm{y}),\\
\mathcal{T}J_{gi2}(\mathrm{y}) &= -\frac{1}{2}\left[\frac{1}{2}\left(\frac{\partial u_i}{\partial x_1}\right)^2 + \frac{3}{2}\left(\frac{\partial u_i}{\partial x_2}\right)^2 + 4(\nabla u_i \cdot \nabla v_{i2})\right](\mathrm{y}).
\end{aligned} \tag{18}$$

Here w_i and v_{ik} are adjoint states computed according to Eq. (9). They are different for any any $i = 1,2,3,4$ in case of w_i and any pair (i,k), $k = 1,2$ in case of v_{ik}. In other words: in order to compute these topological derivatives we must solve for every of twelve functionals the state and adjoint boundary value problems.

The formulae (18) will be used for approximating the effect of the hole of radius ρ made at the point y on the value of the functionals J_{ui}, J_{gik}, according to the Theorem 3. Namely

$$\begin{aligned}
J_{ui}(\rho) &\approx J_{ui}(0) + \mathcal{T}J_{ui}(\mathrm{y}) \cdot \pi\rho^2, \qquad i = 1,2,3,4,\\
J_{gik}(\rho) &\approx J_{gik}(0) + \mathcal{T}J_{gik}(\mathrm{y}) \cdot \pi\rho^2, \qquad k = 1,2.
\end{aligned}$$

In this way, having computed *once* all the state and adjoint variables, we may very cheaply generate big sets of approximate values of the functionals for the domains with holes. These sets may be then used for constructing identification procedure, which on the basis of measured values of functionals finds the position and size of the hole. In particular, they may constitute training sets for neural networks.

4. ARTIFICIAL NEURAL NETWORKS

Research interest in artificial neural networks has been motivated by the observation that the brain processes data in a different way from the conventional digital computer. The brain, due to "massive parallelism" of neuronal computations, outperforms the fastest digital computer in complexity of processing tasks.

In principle, an artificial neural network is a specially constructed parallel distributed processor, which is designed to mimic the brain functions. It is capable of storing the knowledge and using it for data processing. Artificial neural networks resemble the brain in two respects [10]:

- knowledge is acquired by means of a learning process,
- the acquired knowledge is stored in the network structure of inter-neuron connection strengths known as synaptic weights.

The procedure used to perform the network training is called *learning algorithm*. The task of such an algorithm is to modify the synaptic weights of the network in an orderly fashion so as to attain a desired design objective.

Artificial neurons are the basic elements of artificial neural networks. An example [11] of a model of one artificial neuron from the network is shown in Fig. 3.

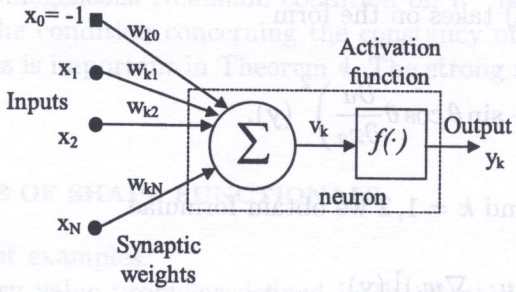


Fig. 3. Model of k-th artificial neuron

It has a set of inputs $x_1, x_2, \ldots, x_N$, denoted as the input vector X and a bias x_0. Each input signal x_j connected to the k-th neuron is multiplied by an associated weight $w_{k1}, w_{k2}, \ldots, w_{kN}$, and bias is multiplied by the weight w_{k0}, before they are applied to the summation block, Σ. Each weight corresponds to the "strength" of a single biological synaptic connection. The summation block adds all of the weighted inputs algebraically and generates a net signal v_k, which is further processed by an activation function $f(\cdot)$ to obtain the neuron's output signal y_k [11],

$$y_k = f(v_k) = f\left(\sum_{j=0}^{N} w_{kj} x_j\right), \tag{19}$$

where N – number of input signals.

Activation function $f(\cdot)$ may be a simple linear or a non-linear function. The most commonly used activation functions are: threshold function, sigmoidal function, hyperbolic tangent function and radial basis function. Sigmoidal function is mathematically expressed as [9, 10]

$$f(z) = \frac{1}{1+e^{-z}}. \tag{20}$$

A single neuron can perform only certain simple functions. The power of neural computations comes from connecting neurons into networks. Structure and size of the designed neural network

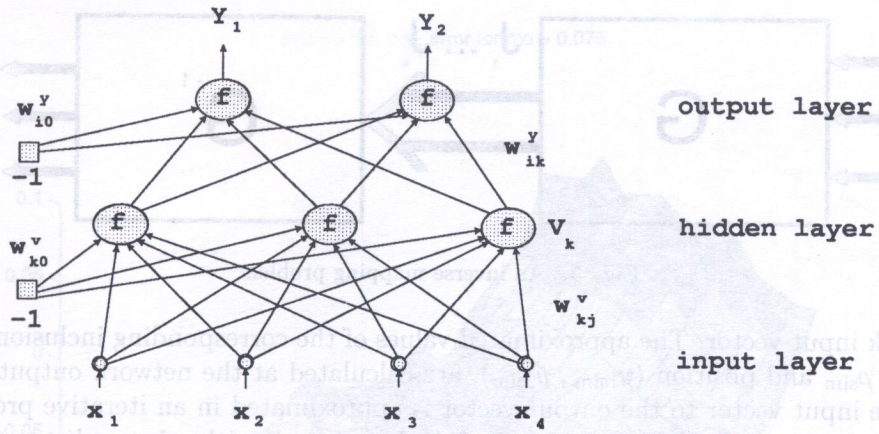


Fig. 4. An example of multilayer perceptron network with one hidden layer

depends on the complexity of the problem, which has to be solved by the network. A great variety of network structures is known [10].

The most commonly used network structure, which has been also applied in the presented approach, is the Multi-Layer Perceptron (MLP). An example of a MLP network is shown in Fig. 4.

MLP network is built from perceptrons grouped in layers [10]. Input signals applied to the network are transmitted in one direction from the network input nodes to the output layer and therefore this type of network is called a feedforward network. The network shown in Fig. 4 is a two-layer network, which has an input layer and two perceptron layers: the hidden layer and the output layer. Additional superscripts v and y in the network synaptic weights in Fig. 4 indicate the layers. Hence, the signal of i-th network output Y_i is given by the equation [10]

$$Y_i = f\left(\sum_{k=0}^{M} w_{ik}^{y} y_k\right) = f\left(\sum_{k=0}^{M} w_{ik}^{y} f\left(\sum_{j=0}^{N} w_{kj}^{v} x_j\right)\right), \tag{21}$$

where N – number of inputs, M – number of neurons in the hidden layer, w_{kj}^{v} – weight of k-th neuron in the hidden layer for the j-th input, w_{ik}^{y} – weight of i-th neuron in the output layer for the signal y_k, which is the output of the k-th neuron in the hidden layer.

It has been proved, that feedforward multilayer perceptron networks are universal approximators [12]. Barron investigated an accuracy of function approximation by the use of neural networks with one-hidden-layer [1]. He characterised sets of functions of many variables, which can be approximated by networks with n hidden units, within an error proportional to $\frac{1}{\sqrt{n}}$.

5. Inverse problem solving by the use of artificial neural networks

Application of Artificial Neural Networks (ANN), instead of analytical calculations, offers a novel and powerful tool for inverse problem solving. The inverse mapping G^{-1}, which allows for identification of inclusion presented in Fig. 2, is difficult to calculate from the mathematical relations and therefore was modelled using artificial neural networks. Similarly as in the classical approach, the inverse mapping G^{-1}, shown in Fig. 5, may be determined unambiguously only when the transformation G has the property, that each input vector y_1, y_2, ρ is transformed into a different values output vector $J_1, \ldots, J_n$ (one to one mapping).

ANN-based inverse model is built on the basis of relations between the network input and output vectors. The knowledge about the inverse mapping is stored within the network structure and network connection weights. Twelve values of functionals $J_1, \ldots, J_{12}$, which ensure unicity of the solution, calculated by the use of topological derivative method for the square with the inclusion

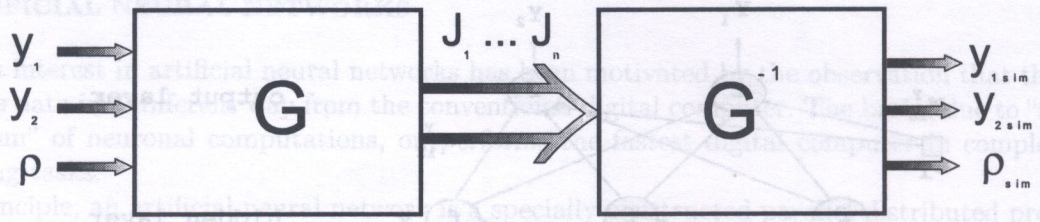


Fig. 5. An inverse mapping problem

are the network input vector. The approximated values of the corresponding inclusion's parameters, such as radius ρ_{sim} and position $(y_{1\text{sim}}, y_{2\text{sim}})$, are calculated at the network output. An unknown mapping of the input vector to the output vector is approximated in an iterative procedure known as neural network training [10]. The objective of the learning algorithm is to adjust network weights on the basis of a given set of input-output pairs for a given cost function to be minimised.

In our particular problem feedforward MLP network, sum square error cost function and back-propagation learning algorithm [10] with Levenberg–Marquardt [9] optimisation method were applied. This algorithm was implemented by the use of MATLAB software package for mathematical computations.

Different MLP networks with a single hidden layer were considered and two of them were tested. The network structure (12-18-3) i.e.: twelve inputs, eighteen processing units with a sigmoidal transfer function in the network hidden layer and three linear units in the output layer, comprising 291 weights; and network (12-24-3) with 387 weights. Numerical computations that were based on the topological derivative have provided data both for network training and testing procedures. The training and testing data were computed for different values of inclusion radius, which were changed from 0.05 to 0.2 and for the corresponding values of the inclusion position. Position coordinates were changed in the range

$$2\rho_i < y_{1i} < 1 - 2\rho_i\,, \qquad 2\rho_i < y_{2i} < 1 - 2\rho_i\,. \tag{22}$$

Then, the corresponding values of functionals J_0, J_1, J_2 for four configurations described earlier were calculated by the use of topological derivative method for each set of inputs. From the available data sets, 1285 of that correspond to the radii $[0.05, 0.088, 0.125, 0.16, 0.2]$ were selected for network training and 205 for radii $[0.075, 0.1, 0.18]$ were selected for network testing. The latter one is required for validation of the network true generalization capabilities. The stopping condition for the learning procedure was the value of sum square error SSE less then 0.02. The network (12-18-3) was trained by the use of Levenberg-Marquardt algorithm in 69 epochs and the network (12-24-3) in 44 epochs.

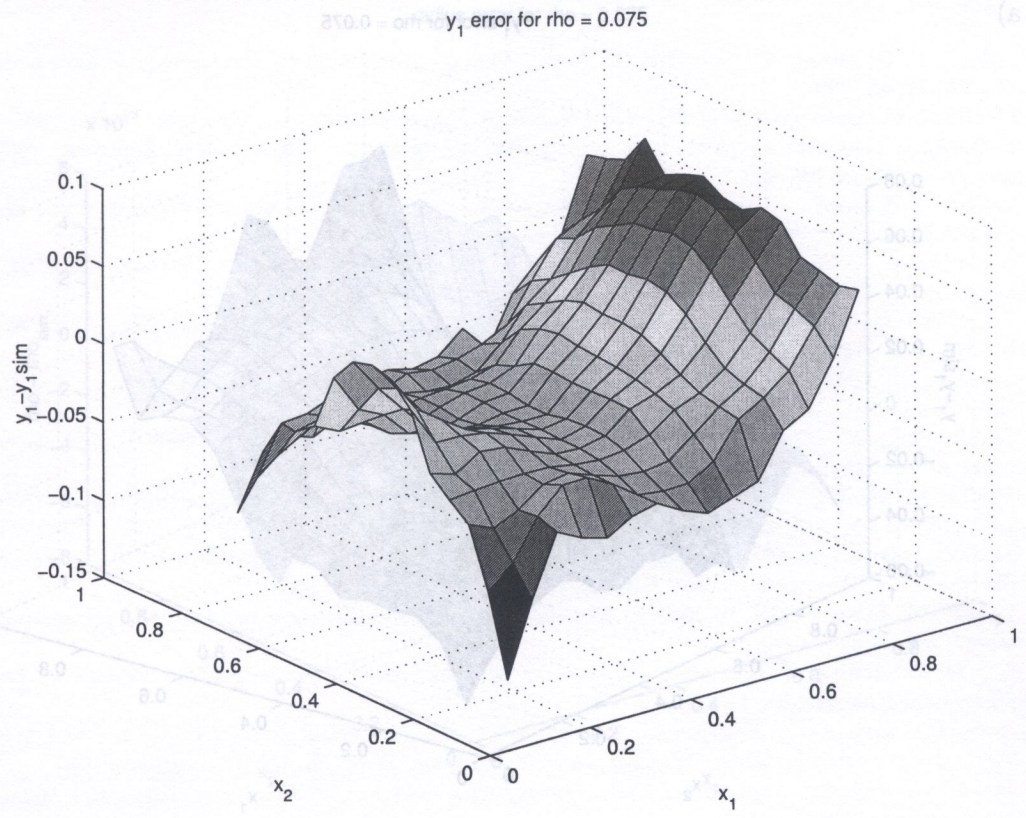
Figures 6, 7 show graphical representation of the results of network (12-18-3) and (12-24-3) testing, namely the examples of error distribution for the inclusion identification calculated for the parameters y_1, y_2, ρ and the radius $\rho = 0.075$. Figures 6, 7 depict absolute errors of the identified parameters. The shapes of error distribution obtained for the networks (12-18-3) and (12-24-3) were very similar, but the values of absolute errors were different. The largest values of the errors in position identification can be observed in Figs. 6, 7 in the corners of the square.

The maximum values of relative errors for both of the tested networks and their training times are given in a Table 1. Increasing the number of neurons in the hidden layer has improved the accuracy by a factor of two, but the learning time has also increased ten times. Computations has been performed at IBM PC computer with Celeron 300 MHz processor and 32 MB RAM.

Table 1. Maximum values of relative errors for $\rho = 0.075$

network structure	δy_1 [%]	δy_2 [%]	$\delta\rho$ [%]	learning time [min]
(12-18-3)	12	12	5	24
(12-24-3)	6	6	3	240

a)



b)

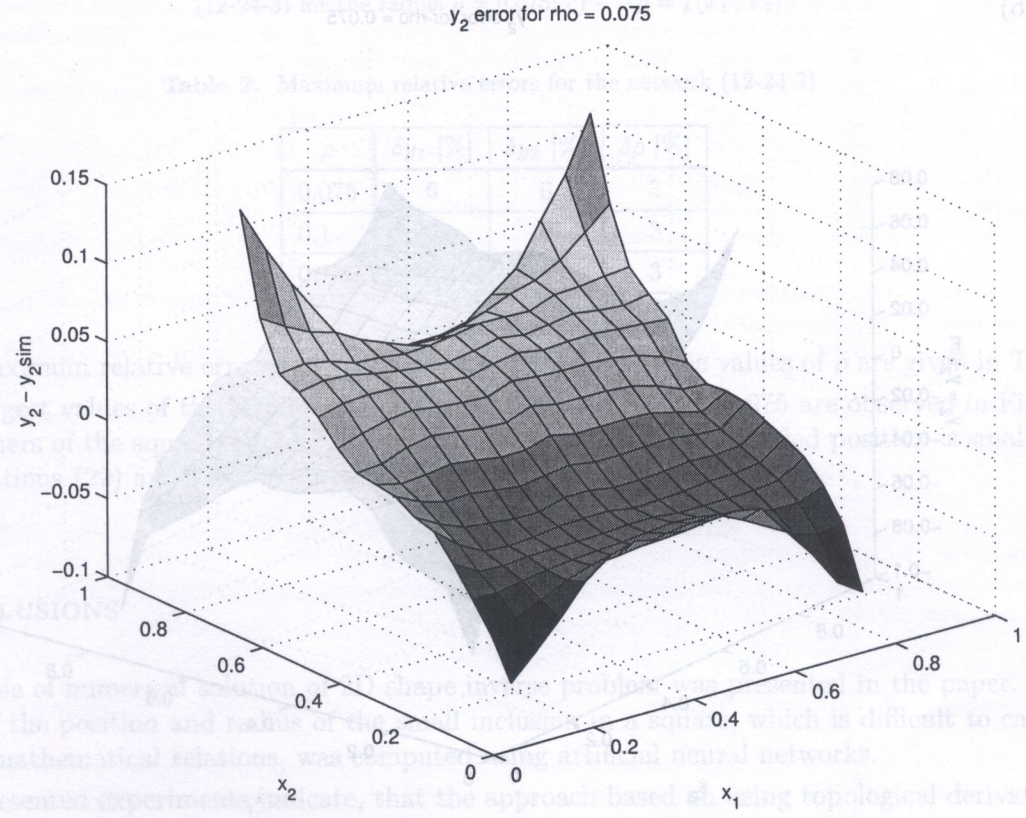


Fig. 6. Distribution of the absolute error of the position parameters determination by the network (12-18-3) for the radius $\rho = 0.075$; a) – $\Delta y_1 = f(x_1, x_2)$; b) – $\Delta y_2 = f(x_1, x_2)$

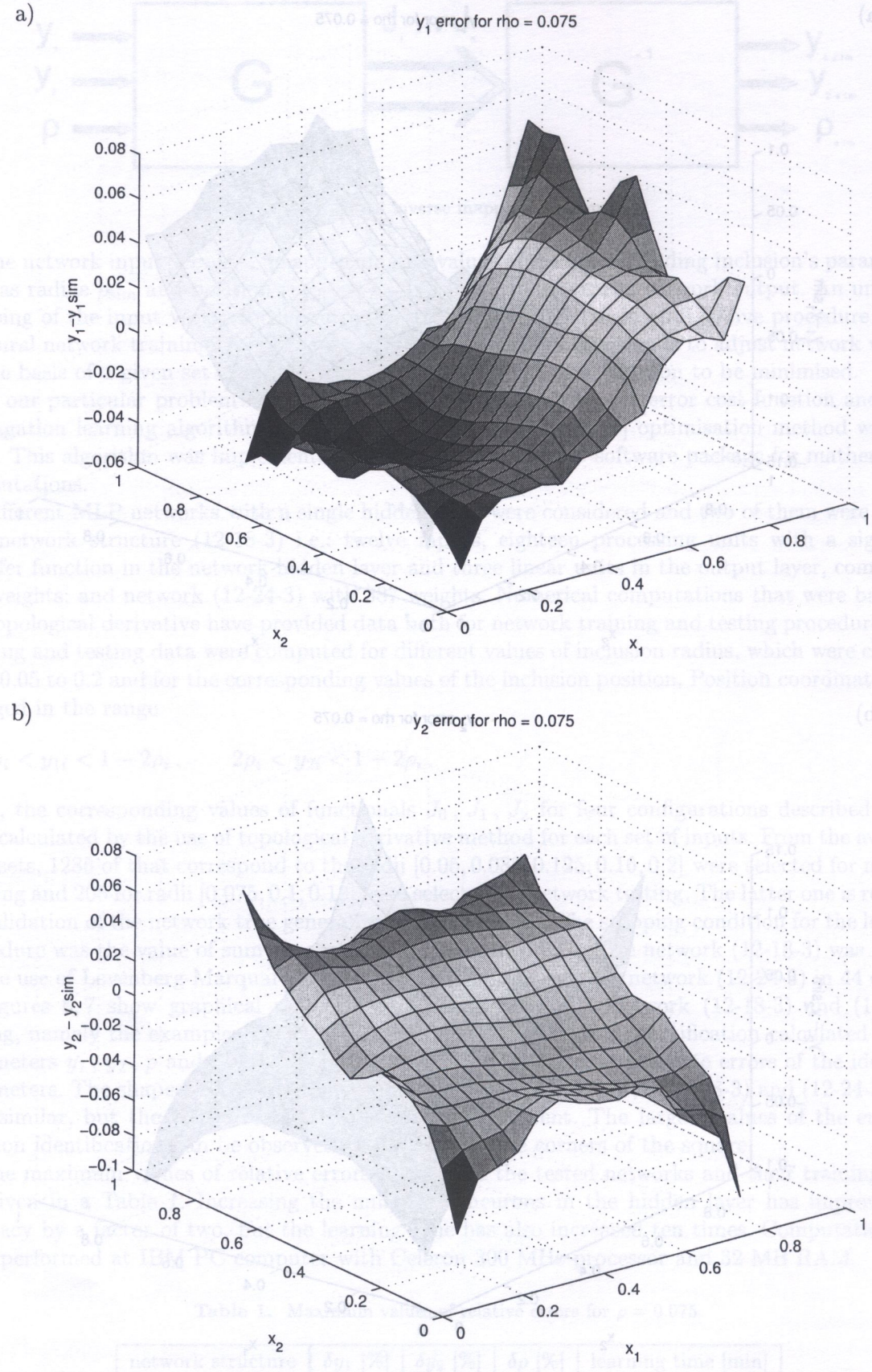


Fig. 7. Distribution of the absolute error of the position parameters determination by the network (12-24-3) for the radius $\rho = 0.075$; a) – $\Delta y_1 = f(x_1, x_2)$; b) – $\Delta y_2 = f(x_1, x_2)$ *(contd. in the next page)*

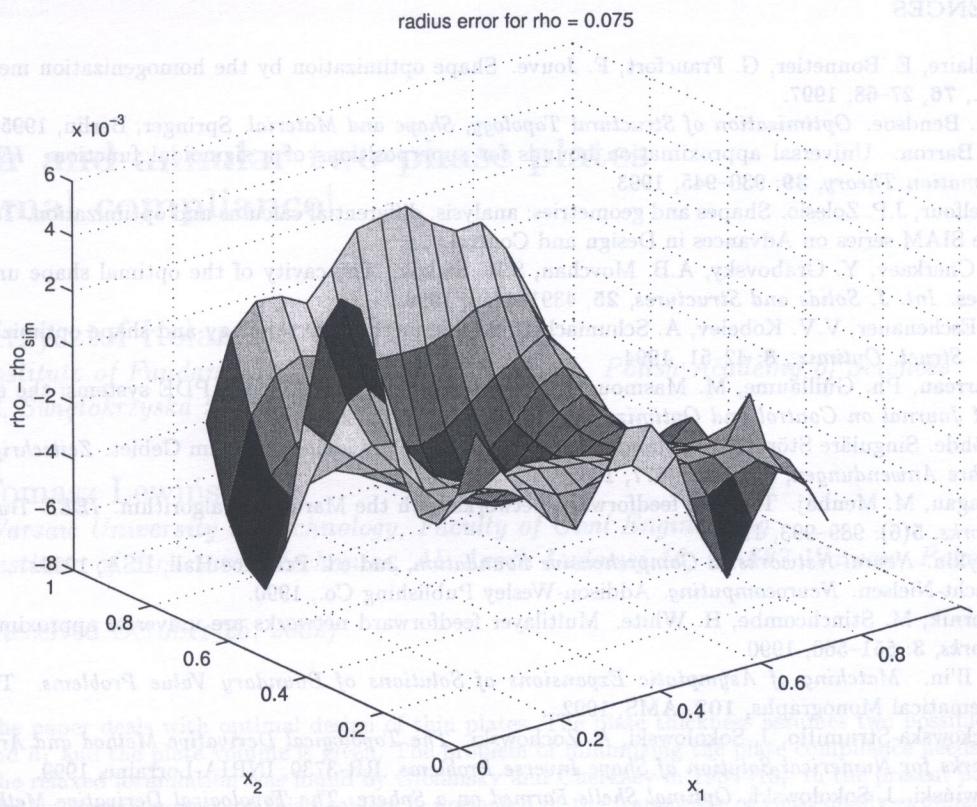


Fig. 7. *(contd.)* Distribution of the absolute error of the position parameters determination by the network (12-24-3) for the radius $\rho = 0.075$; c) – $\Delta\rho = f(x_1, x_2)$

Table 2. Maximum relative errors for the network (12-24-3)

ρ	δy_1 [%]	δy_2 [%]	$\delta\rho$ [%]
0.075	6	6	3
0.1	5	5	3
0.18	2	2	3

The maximum relative errors for the network (12-24-3) for three values of ρ are given in Table 2.

The largest values of the errors in position identification for $\rho = 0.075$ are observed in Figs. 6, 7 in the corners of the square. For the other radii values, the area of identified position is smaller due to the relations (22) and the corresponding values of errors are also smaller.

6. Conclusions

An example of numerical solution of 2D shape inverse problem was presented in the paper. Identification of the position and radius of the small inclusion in a square, which is difficult to calculate from the mathematical relations, was computed using artificial neural networks.

The presented experiments indicate, that the approach based on using topological derivative for producing training data for neural networks, gives promising results.

Application of Artificial Neural Networks (ANN), instead of analytical calculations, offers a novel and powerful tool for inverse problem solving.

References

[1] G. Allaire, E. Bonnetier, G. Francfort, F. Jouve. Shape optimization by the homogenization method. *Numer. Math.*, **76**, 27–68, 1997.

[2] M.Ph. Bendsoe. *Optimization of Structural Topology, Shape and Material.* Springer, Berlin, 1995.

[3] A.R. Barron. Universal approximation bounds for superpositions of a sigmoidal function. *IEEE Trans. on Information Theory*, **39**: 930–945, 1993.

[4] M. Delfour, J.P. Zolesio. Shapes and geometries: analysis, differential calculus and optimization. To be published in the SIAM series on Advances in Design and Control.

[5] A.V. Cherkaev, Y. Grabovsky, A.B. Movchan, S.K. Serkov. The cavity of the optimal shape under the shear stresses. *Int. J. Solids and Structures*, **25**, 4391–4410, 1999.

[6] H.A. Eschenauer, V.V. Kobelev, A. Schumacher. Bubble method for topology and shape optimization of structures. *Struct. Optimiz.*, **8**: 42–51, 1994.

[7] S. Garreau, Ph. Guillaume, M. Masmoudi. The topological asymptotic for PDE systems: the elasticity case. *SIAM Journal on Control and Optimization*, **39**(6): 1756–1778, 2001.

[8] D. Göhde. Singuläre Störung von Randvertproblemen durch ein kleines Loch im Gebiet. *Zeitschrift für Analysis und ihre Anwendungen*, **4**(5): 467–477, 1985.

[9] M. Hagan, M. Menhaj. Training feedforward networks with the Marquardt algorithm. *IEEE Trans. on Neural Networks*, **5**(6): 989–993, 1994.

[10] S. Haykin. *Neural Networks: a Comprehensive Foundation*, 2nd ed. Prentice-Hall, USA, 1999.

[11] R. Hecht-Nielsen. *Neurocomputing.* Addison-Wesley Publishing Co., 1990.

[12] K. Hornik, M. Stinchcombe, H. White. Multilayer feedforward networks are universal approximators. *Neural Networks*, **3**: 551–560, 1990.

[13] A.M. Il'in. *Matching of Asymptotic Expansions of Solutions of Boundary Value Problems.* Translations of Mathematical Monographs, **102**, AMS, 1992.

[14] L. Jackowska-Strumiłło, J. Sokołowski, A. Żochowski. *The Topological Derivative Method and Artificial Neural Networks for Numerical Solution of Shape Inverse Problems.* RR-3739, INRIA-Lorraine, 1999.

[15] T. Lewiński, J. Sokołowski. *Optimal Shells Formed on a Sphere. The Topological Derivative Method.* RR-3495, INRIA-Lorraine, 1998.

[16] T. Lewiński, J. Sokołowski. Topological derivative for nucleation of non-circular voids. In: R. Gulliver, W. Littman, R. Triggiani, eds., *Differential Geometric Methods in the Control of Partial Differential Equations*, 1999 AMS-IMS-SIAM Joint Summer Research Conference, Univ. of Colorado, Boulder June 27–July 1, 1999. *Contemporary Mathematics*, American Math. Soc. **268**: 341–361, 2000.

[17] T. Lewiński, J. Sokołowski. *Energy Change due to Appearing of Cavities in Elastic Solids.* Les prépublications de l'Institut Élie Cartan 23/2001.

[18] T. Lewiński, J. Sokołowski, A. Żochowski. *Justification of the bubble method for the compliance minimization problems of plates and spherical shells.* CD-ROM, *3rd World Congress of Structural and Multidisciplinary Optimization (WCSMO-3)* Buffalo/Niagara Falls, New York, May 17–21, 1999.

[19] T. Lewiński, J.J. Telega. *Plates, Laminates and Shells.* Series on Advances in Applied Sciences, World Scientific, Singapore, 2000.

[20] S.A. Nazarov, B.A. Plamenevsky. *Elliptic Problems in Domains with Piecewise Smooth Boundaries*,. De Gruyter Exposition in Mathematics, **13**, Walter de Gruyter, 1994.

[21] S.A. Nazarov, J. Sokołowski. Asymptotic analysis of shape functionals. *Journal de Mathématiques Pures et Apliquées*, **82**(2): 125–196, 2003.

[22] J.R. Roche, J. Sokołowski. Numerical methods for shape identification problems. Special issue of *Control and Cybernetics: Shape Optimization and Scientific Computations*, **5**, 867–894, 1996.

[23] A. Shumacher. *Topologieoptimierung von Bauteilstrukturen unter Verwendung von Lochpositionierungkriterien*, Ph.D. Thesis, Universität-Gesamthochschule-Siegen, Siegen, 1995.

[24] J. Sokołowski, J-P. Zolesio. *Introduction to Shape Optimization. Shape Sensitivity Analysis.* Springer-Verlag, 1992.

[25] J. Sokołowski, A. Żochowski. On topological derivative in shape optimization. *SIAM Journal on Control and Optimization*, **37**(4): 1251–1272, 1999.

[26] J. Sokołowski, A. Żochowski. Topological derivative for optimal control problems. *Control and Cybernetics*, **28**(3): 611–626, 1999.

[27] J. Sokołowski, A. Żochowski. Topological derivatives for elliptic problems. *Inverse Problems*, **15**(1): 123–134, 1999.

[28] J. Sokołowski, A. Żochowski. *Topological Derivatives of Shape Functionals for Elasticity Systems.* Les prepublications de l'Institut Elie Cartan 1999/no. 35, *Mechanics of Structures and Machines*, **29**(3): 333–351, 2001.

[29] J. Sokołowski, A. Żochowski. *On Topological Derivative in Shape Optimisation.* INRIA-Lorraine, Rapport de Recherche No. 3170, 1997.

[30] J. Sokołowski, A. Żochowski. Optimality conditions for simultaneous topology and shape design. Les prepublications de l'Institut Elie Cartan 8/2001; to appear in *SIAM Journal on Control and Optimization*, 2003.