Josef Malík; Alexej Kolcun

Determination of the initial stress tensor from deformation of underground opening in excavation process

Applications of Mathematics, Vol. 67 (2022), No. 3, 297–318

Persistent URL: http://dml.cz/dmlcz/150317

67 (2022) APPLICATIONS OF MATHEMATICS No. 3, 297–318

DETERMINATION OF THE INITIAL STRESS TENSOR FROM DEFORMATION OF UNDERGROUND OPENING IN EXCAVATION PROCESS

Josef Malík, Alexej Kolcun, Ostrava

Received September 18, 2020. Published online November 16, 2021.

Abstract. A method for the detection of the initial stress tensor is proposed. The method is based on measuring distances between pairs of points located on the wall of underground opening in the excavation process. This methods is based on solving twelve auxiliary problems in the theory of elasticity with force boundary conditions, which is done using the least squares method. The optimal location of the pairs of points on the wall of underground openings is studied. The pairs must be located so that the condition number of the least square matrix has the minimal value, which guarantees a reliable estimation of initial stress tensor.

Keywords: initial stress tensor; first boundary value problem of the theory of elasticity; least square method; condition number of matrix; continuous dependence of eigenvalues on matrix elements

MSC 2020: 65N30, 74-XX, 93E24

1. Introduction

The knowledge of the initial stress tensor is very important when one evaluates the stability of underground openings like tunnels, compressed gas tanks or radioactive waste deposits. The knowledge of the initial stress tensor enables to optimize the reinforcement of tunnels and to choosing the suitable shape of underground openings and their orientation in the rock environment.

The mathematical modeling of stress fields in the vicinity of underground openings requires precise boundary conditions, which can be derived from the initial stress tensor. Extensive literature is devoted to the determination of the initial

This work was supported by Czech Science Foundation (GAČR) through project No. 19-11441S.

DOI: 10.21136/AM.2021.0266-20

stress tensor. An overview of these methods can be found in papers [3], [5], [8] that describe the development of these methods up to the present. Theoretical and practical aspects of these methods are studied in [11] and [12]. These methods are based on the installation of probes equipped with sensors that measure deformations occurring after removal of a rock, overcoring, in their vicinity. Due to the stress in the rock, the removal of a part of the rock causes deformation of the remaining rock, which is transfered to the sensors. The probes are relatively small, a few centimeters, and the accuracy of such measurements is not high. The complete initial stress tensor can be obtained by applying the conical probe method as described in [10]. However, this method places considerable demands on the equipment needed to install the conical probe and to perform the necessary operations.

In this paper we present a new method, which is based on measuring the distances between pairs of selected points on the walls of the underground opening. When a part of the rock is excavated, the distance between these points changes and the magnitude of these changes depends on the initial stress tensor. A procedure which allows to determine the initial stress tensor from the measured distances is developed. A criterion showing how to select measuring points so that the errors of measurement do not affect the results too much is presented. We assume that the initial stress tensor is constant throughout the studied area before the excavation process.

The article is divided into six sections. The second section gathers some knowledge about the first boundary problem of the theory of elasticity. The third section formulates a method for determining the initial stress tensor based on measuring the distance between selected pairs of points in the process of excavating an underground opening. The fourth section deals with the optimal choice of measuring points, which guarantees a reliable estimate of the initial stress tensor. The procedure is based on the least square method and the criterion of optimal choice of measuring points is based on the conditional number of the matrix of the least square method. Some properties of the least square matrix are proved in relation to the position of measuring points. The fifth chapter is devoted to the numerical solution and part of the chapter is a numerical example, which demonstrates how the selection of measuring points affects the reliability of determining the initial stress tensor.

2. The first boundary problem in linear elasticity

The method described in this section is based on the solution of the first boundary problem of the theory of elasticity, i.e. only the force conditions are prescribed on the boundary of the domain, where the problem is solved. A typical problem solving domain is shown in Fig. 1.

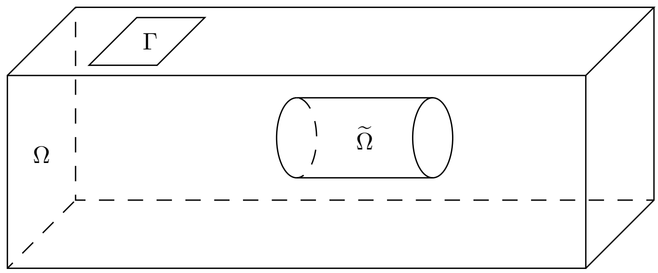

Figure 1. Typical problem solving domain.

Symbol Ω in Fig. 1 is the domain that corresponds to the prism and symbol $\widetilde{\Omega}$ is the domain that represents the excavated space in the domain Ω. Symbol Ω_1 corresponds to domain $\Omega \setminus \widetilde{\Omega}$ and $\Gamma \subset \partial\Omega$ has a nonzero measure.

Let us have a space $V = [H^1(\Omega_1)]^3$, where $H^1(\Omega_1)$ is a Sobolev space of functions having first-oder derivatives that are integrable with the second power. We will continue to apply the Einstein summation convention.

Let us formulate the first variational problem $\mathcal{D}_1$ whose solution is a minimum of the following functional on V:

$$\frac{1}{2}\int_{\Omega_1} c_{ijkl} e_{ij}(u) e_{kl}(u)\,\mathrm{d}x - \int_{\partial\Omega} P_i u_i\,\mathrm{d}S, \tag{1}$$

where $u = (u_1, u_2, u_3)$ is the vector of displacements and belongs to V and

$$e_{ij}(u) = \frac{1}{2}\Big(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\Big)$$

is the tensor of small deformations. Symbol $P = (P_1, P_2, P_3)$ represents the forces on $\partial\Omega$ and $P_i \in L^2(\partial\Omega)$. The coefficients $c_{ijkl} \in L^\infty(\Omega_1)$ meet the following conditions:

$$c_{ijkl} = c_{jikl} = c_{ijlk} = c_{klij}. \tag{2}$$

There is a constant $C > 0$ such that the inequality

$$c_{ijkl} e_{ij} e_{kl} \geqslant C e_{ij} e_{ij} \tag{3}$$

holds for all symmetric tensors e_{ij}.

Problem $\mathcal{D}_1$ is solvable when the conditions

$$\int_{\partial\Omega} P_i \, \mathrm{d}S = 0, \quad \int_{\partial\Omega} (x \times P)_i \, \mathrm{d}S = 0, \quad i = 1, 2, 3, \tag{4}$$

are met. This problem is not uniquely solvable and it has infinite number of solutions. If $u_1(x)$ and $u_2(x)$ are two solutions, then

$$u_2(x) - u_1(x) = Ax + b, \tag{5}$$

where A is an antisymmetric matrix 3×3 and b is a vector from $\mathbb{R}^3$.

This problem can be modified so that it will be uniquely solvable and this solution will be the minimum of the functional (1), i.e. the solution of problem $\mathcal{D}_1$, provided conditions (4) are met.

Let us define functionals on V as

$$g_\alpha(u) = \begin{cases} \displaystyle\int_\Gamma u_\alpha \, \mathrm{d}S, & \alpha = 1, 2, 3, \\ \displaystyle\int_\Gamma (x \times u)_{\alpha-3} \, \mathrm{d}S, & \alpha = 4, 5, 6. \end{cases} \tag{6}$$

Then there is a constant $C > 0$ such that the inequality

$$C\|u\|_V \leqslant \int_{\Omega_1} c_{ijkl} e_{ij} e_{kl} \, \mathrm{d}x + g_\alpha(u) g_\alpha(u) \tag{7}$$

holds for all $u \in V$.

Let us formulate the second variational problem $\mathcal{D}_2$ whose solution is a minimum of the functional

$$\frac{1}{2} \int_{\Omega_1} c_{ijkl} e_{ij}(u) e_{kl}(u) \, \mathrm{d}x + \frac{1}{2} g_\alpha(u) g_\alpha(u) - \int_{\partial\Omega} P_i u_i \, \mathrm{d}S \tag{8}$$

on V. The minimum of functional (8) is unique. Moreover, the inequality

$$\|u\|_V \leqslant C\|P\|_{[L^2(\partial\Omega)]^3} \tag{9}$$

holds, where C is a positive constant independent of u and P. The last inequality expresses the continuous dependence of the solution of problem $\mathcal{D}_2$ on the force boundary conditions. Note that solving problem $\mathcal{D}_2$ does not require the equilibrium conditions (4) to be met. But if these conditions are satisfied, the solution of $\mathcal{D}_2$ is a solution of $\mathcal{D}_1$. All these results can be found in book [7].

Let τ_{ij} be a symmetric tensor. We say that the force boundary conditions P_i are generated by the tensor τ_{ij} when at every $x \in \partial\Omega$ the equation

$$P_i(x) = \tau_{ij} n_j(x) \tag{10}$$

holds, where $n_j(x)$ is a normal vector to the boundary $\partial\Omega$ at the point x.

Lemma 1. *Let τ_{ij} be a symmetric tensor and let P_i be defined by formula* (10) *on the boundary $\partial\Omega$. Then P_i satisfy the equilibrium conditions* (4).

P r o o f. If we use the Gaussian theorem on the surface integral, then

$$\int_{\partial\Omega} P_i(x)\,\mathrm{d}S = \int_{\partial\Omega} \tau_{ij} n_j(x)\,\mathrm{d}S = \int_{\Omega} \frac{\partial \tau_{ij}}{\partial x_j}\,\mathrm{d}x.$$

Since τ_{ij} is constant, then the last integral is zero. We express the formula $x \times P$ in the individual components. Then

$$\begin{aligned}
(x \times P)_1 &= x_2\tau_{3j}n_j - x_3\tau_{2j}n_j,\\
(x \times P)_2 &= x_3\tau_{1j}n_j - x_1\tau_{3j}n_j,\\
(x \times P)_3 &= x_1\tau_{2j}n_j - x_2\tau_{1j}n_j,
\end{aligned}$$

where $n = (n_1, n_2, n_3)$ is the normal to the boundary $\partial\Omega$ at x. If we use the Gaussian theorem on the surface integral, then

$$\int_{\partial\Omega} (x \times P)_1\,\mathrm{d}S = \int_{\partial\Omega} (x_2\tau_{3j}n_j - x_3\tau_{2j}n_j)\,\mathrm{d}S = \int_{\Omega} \frac{\partial(x_2\tau_{3j} - x_3\tau_{2j})}{\partial x_j}\,\mathrm{d}x.$$

Since τ_{ij} is symmetric and constant, then the last integral is zero. The same equations can be proven for the components $(x \times P)_2$ and $(x \times P)_3$. □

Inequality (9) implies the existence of a continuous mapping

$$K\colon S^{\mathrm{sym}} \to V, \tag{11}$$

where S^{sym} is the set of all symmetric tensors of second order. This mapping assigns a solution to problem $\mathcal{D}_2$ to each second order symmetric tensor. Lemma 1 indicates that the value of this mapping is also a solution to problem $\mathcal{D}_1$.

3. Derivation and analysis of the optimization problem

In this section we will describe the method of obtaining the initial stress tensor from measuring distances between suitably selected pairs of points. The solution to our problem will be based on the first boundary problem of the theory of elasticity and the approach used is shown in Fig. 2a–2c.

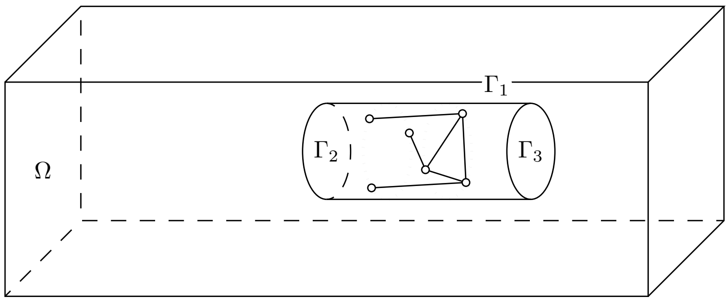

Figure 2a. First step of measuring process.

Suppose we know the initial stress tensor τ. From this tensor we generate the force boundary conditions P. We will divide the solution to our problem into three steps. The first step is shown in Fig. 2a. We solve problem $\mathcal{D}_2$ on the domain $\Omega_1 = \Omega\backslash\widetilde{\Omega}$ with the boundary force conditions P on $\partial\Omega$. The solution $u_1(x)$ of this problem belongs to $[H^1(\Omega_1)]^3$. The boundary $\partial\widetilde{\Omega}$ can be divided into three parts. The parts Γ_2 and Γ_3 are fronts of the tunnel that correspond to the domain $\widetilde{\Omega}$. The remaining part of the boundary is the surface $\Gamma_1 = \partial\widetilde{\Omega} \setminus (\Gamma_2 \cup \Gamma_3)$ on which the measuring points are installed, which are denoted by x_k, $k = 1, \ldots, N$. Then we select the pairs of measuring points x_k, x_l and compute the expression

$$\|u_1(x_k) + x_k - u_1(x_l) - x_l\|, \tag{12}$$

which is the distance between the points x_k, x_l after deformation caused by forces P on $\partial\Omega$. The symbol $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^3$. In Fig. 2a and 2c, the measuring points are located at the ends of the lines that connect the selected pairs of measuring points. The function $u_1(x)$ belongs to $[H^1(\Omega_1)]^3$, so (12) is not defined correctly. For a moment we will assume that solutions are continuous functions defined on the whole domain. We will get rid of this assumption later.

The second step is shown in Fig. 2b. The domain $\widetilde{\widetilde{\Omega}}$ is removed and problem $\mathcal{D}_2$ is resolved on the domain $\Omega_2 = \Omega \setminus (\widetilde{\Omega} \cup \widetilde{\widetilde{\Omega}})$ with the same boundary conditions. The solution to this problem $u_2(x)$ belongs to $[H^1(\Omega_2)]^3$.

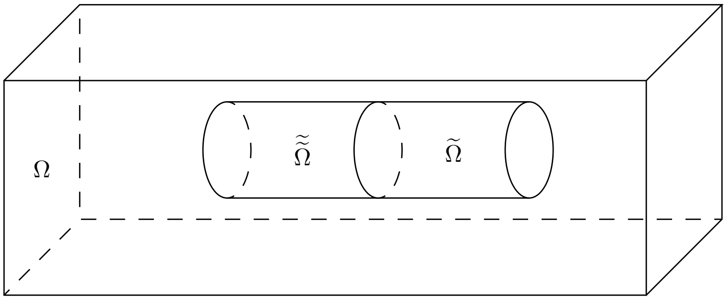

Figure 2b. Second step of measuring process.

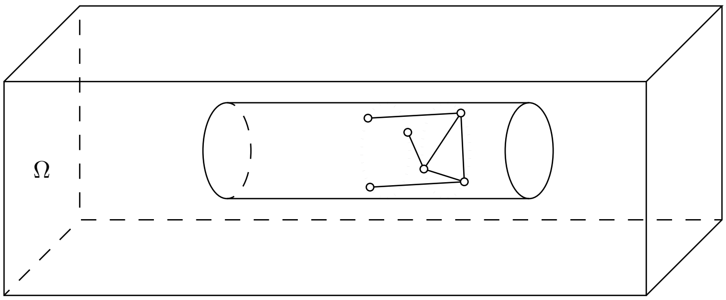

Figure 2c. Third step of measuring process.

The third step is shown in Fig. 2c. We compute the expression

$$\|u_2(x_k) + x_k - u_2(x_l) - x_l\|, \tag{13}$$

which is the distance between the points x_k, x_l after deformation caused by the forces P on $\partial\Omega$. If we subtract expression (12) from (13), then we have

$$\|u_2(x_k) + x_k - u_2(x_l) - x_l\| - \|u_1(x_k) + x_k - u_1(x_l) - x_l\|, \tag{14}$$

which represents the change in distance between points x_k, x_l after the domain $\tilde{\tilde{\Omega}}$ is removed.

Our task is to find the tensor τ such that the boundary conditions P generated by this tensor lead to solutions $u_1(x)$ and $u_2(x)$ for which expressions (14) coincide with the measured changes between the points. Expression (11) implies that the solutions $u_1(x)$ and $u_2(x)$ continuously depend on τ and this dependence is linear. On the other hand, expressions (14) are nonlinear and finding solution to our problem will be difficult. Let us try to linearize expression (14). Suppose that $\|u_1(x_k) - u_1(x_l)\|$ and $\|u_2(x_k) - u_2(x_l)\|$ are very small compared to $\|x_k - x_l\|$. This hypothesis is acceptable, because in practice the measured displacements corresponding to (14) are very small compared to the distance between the points. Then expressions (14) can be linearized and the following lemma shows how to do it.

Lemma 2. *Let* $u_1, u_2, x_1, x_2 \in \mathbb{R}^3$ *and the value*

$$a = \frac{\|u_1 - u_2\|}{\|x_1 - x_2\|} < 1 \tag{15}$$

be so small that a^2 *can be neglected. Then the equality*

$$\|u_1 + x_1 - u_2 - x_2\| - \|x_1 - x_2\| = \frac{\langle u_1 - u_2, x_1 - x_2\rangle}{\|x_1 - x_2\|} \tag{16}$$

holds approximately, where $\langle\cdot,\cdot\rangle$ *is the scalar product in* $\mathbb{R}^3$. *Moreover, if*

$$v_1 = u_1 + Ax_1 + b, \qquad v_2 = u_2 + Ax_2 + b,$$

where A *is an antisymmetric matrix* 3×3 *and* b *is a vector from* $\mathbb{R}^3$, *then*

$$\frac{\langle v_1 - v_2, x_1 - x_2\rangle}{\|x_1 - x_2\|} = \frac{\langle u_1 - u_2, x_1 - x_2\rangle}{\|x_1 - x_2\|}. \tag{17}$$

P r o o f. The expression on the left-hand side of equality (16) can be written in the following form:

$$\Big(\frac{\|u_1 - u_2\|^2}{\|x_1 - x_2\|^2} + \frac{2\langle u_1 - u_2, x_1 - x_2\rangle}{\|x_1 - x_2\|^2} + 1\Big)^{1/2} \|x_1 - x_2\| - \|x_1 - x_2\|. \tag{18}$$

If we consider the assumptions of this lemma, then expression (18) is approximately equal to

$$\Big(1 + \frac{2\langle u_1 - u_2, x_1 - x_2\rangle}{\|x_1 - x_2\|^2}\Big)^{1/2} \|x_1 - x_2\| - \|x_1 - x_2\|$$

and the last expression is approximately equal to

$$\Big(1 + \frac{\langle u_1 - u_2, x_1 - x_2\rangle}{\|x_1 - x_2\|^2}\Big) \|x_1 - x_2\| - \|x_1 - x_2\|.$$

The last expression implies equality (16). Let us proceed to prove the last statement of this lemma. If we consider the relationship between u_1, u_2 and v_1, v_2 in the assumptions of this lemma, we have the system of equations

$$\langle v_1 - u_1 - v_2 + u_2, x_1 - x_2\rangle = \langle A(x_1 - x_2), x_1 - x_2\rangle = 0.$$

The last equality results from matrix A being antisymmetric. □

When tunnels or other underground openings are excavated, displacements between pairs of points on the surface of these structures can be measured. The change of distances after removing some part of rock is usually a few millimeters while the distance between points is a few meters, depending on the dimension of the underground opening. These displacements are always much smaller than the distance between points. It is therefore acceptable to replace (14) with the expression

$$\frac{\langle u(x_k) - u(x_l), x_k - x_l \rangle}{\|x_k - x_l\|}, \tag{19}$$

where $u(x) = u_2(x) - u_1(x)$ and (19) linearly depends on $u(x)$. If we select $\Gamma \subset \partial\Omega$ in the functional (8) in another way, we get different solutions $u_1(x)$, $u_2(x)$, but expression (19) does not change, which results from Lemma 2.

Before we formulate our problem strictly, we have to consider that solution $u_1(x)$ belongs to $[H^1(\Omega_1)]^3$ and solution $u_2(x)$ belongs to $[H^1(\Omega_2)]^3$. The functions $u_1(x)$ and $u_2(x)$ are defined almost everywhere, so expression (19) is not defined correctly. It is necessary to replace it in such a way that it is consistent with weak formulation of the solved problem. In practice, the measuring points are located on small steel bars glued into the rock as shown in Fig. 3.

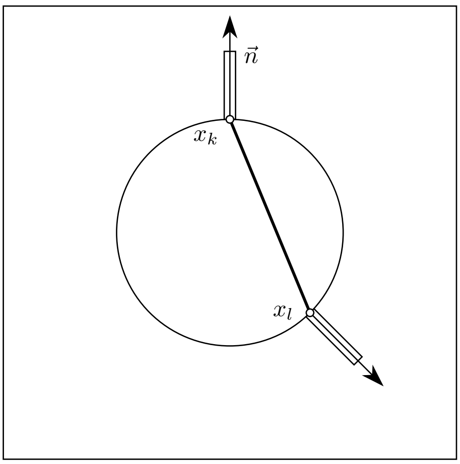

Figure 3. Position of steel bars.

All bars are the same. Suppose that Γ_1, on which the measuring points are located, is smooth enough. Then for every $x \in \Gamma_1$ there is a normal vector that is continuous on Γ_1. We define $\omega(x)$ as the domain corresponding to the steel bar whose longitudinal axis is oriented in the direction of the normal passing through point x as shown in Fig. 3. Let us define the expression

$$\mathcal{E}(x,u) = \frac{1}{\mu(\omega(x))} \int_{\omega(x)} u(z)\,\mathrm{d}z, \tag{20}$$

where $u(x)$ belongs to $[H^1(\Omega_2)]^3$ and μ is the Lebesgue measure in $\mathbb{R}^3$.

The value of the expression does not change if $u(x)$ is replaced by another function that matches the original function up to a set of measure zero. Thus, this expression is correctly defined on Γ_1. The following lemma discusses the continuity of the preceding expression and will be used later.

Lemma 3. *If surface Γ_1 is of class C^1 and $u(x)$ belongs to $[H^1(\Omega_2)]^3$, then $\mathcal{E}(x,u)$ is continuous on $\Gamma_1 \times [H^1(\Omega_2)]^3$.*

Proof. The function $u(x)$ belongs to $[H^1(\Omega_2)]^3$ and thus belongs to $[L(\Omega_2)]^3$. From the properties of measure (see [4]) it follows that for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\left\| \int_{\omega} u(x)\,\mathrm{d}x \right\| < \varepsilon \tag{21}$$

for every $\omega \subset \Omega_2$ satisfying $\mu(\omega) < \delta$.

We shall consider the following inequality:

$$\|\mathcal{E}(x_1,u_1) - \mathcal{E}(x_2,u_2)\| \leqslant \|\mathcal{E}(x_1,u_1) - \mathcal{E}(x_2,u_1)\| + \|\mathcal{E}(x_2,u_1) - \mathcal{E}(x_2,u_2)\|. \tag{22}$$

Let us deal with the right-hand side of (22). For its first term the inequality

$$\|\mathcal{E}(x_1,u_1) - \mathcal{E}(x_2,u_1)\| \leqslant \frac{1}{\mu(\omega(x_1))} \left\| \int_{\omega(x_1)\Delta\omega(x_2)} u_1(x)\,\mathrm{d}x \right\| \tag{23}$$

holds, where $\omega(x_1)\Delta\omega(x_2) = (\omega(x_1)\cup\omega(x_2))\setminus(\omega(x_1)\cap\omega(x_2))$. If $\|x_1 - x_2\|$ converges to zero, then $\mu(\omega(x_1)\Delta\omega(x_2))$ converges to zero. According to (21), the right-hand side of (23) converges to zero too. For the second term in (22), the inequality

$$\|\mathcal{E}(x_2,u_1) - \mathcal{E}(x_2,u_2)\| \leqslant \frac{1}{\mu(\omega(x_2))} \left\| \int_{\omega(x_2)} (u_1(x) - u_2(x))\,\mathrm{d}x \right\| \tag{24}$$

holds. If $u_2(x)$ converges to $u_1(x)$ in $[H^1(\Omega_2)]^3$, then the right-hand side of (24) converges to zero.

Inequalities (23) and (24) imply the continuity of $\mathcal{E}(x,u)$ on $\Gamma_1 \times [H^1(\Omega_2)]^3$. □

Relation (20) gives the average of displacements $u(x)$ on $\omega(x)$. If we consider (20) for identity function $I(z) = z$, we obtain the center of gravity of $\omega(x)$:

$$\mathcal{E}(x, I) = \frac{1}{\mu(\omega(x))} \int_{\omega(x)} z \, \mathrm{d}z.$$

From the last lemma it follows that this function is continuous on Γ_1.

The following lemma corresponds to Lemma 2.

Lemma 4. *Let $u(x)$ belong to $[H^1(\Omega_2)]^3$ and*

$$v(x) = u(x) + Ax + b,$$

where A is an antisymmetric matrix 3×3 and b is a vector in $\mathbb{R}^3$. Then

$$\begin{aligned} \frac{\langle \mathcal{E}(x_k, u) - \mathcal{E}(x_l, u), \mathcal{E}(x_k, I) - \mathcal{E}(x_l, I) \rangle}{\|\mathcal{E}(x_k, I) - \mathcal{E}(x_l, I)\|} \\ = \frac{\langle \mathcal{E}(x_k, v) - \mathcal{E}(x_l, v), \mathcal{E}(x_k, I) - \mathcal{E}(x_l, I) \rangle}{\|\mathcal{E}(x_k, I) - \mathcal{E}(x_l, I)\|}. \end{aligned} \tag{25}$$

P r o o f. The definition of $\mathcal{E}(x, u)$ implies the equality:

$$\mathcal{E}(x, u + v) = \mathcal{E}(x, u) + \mathcal{E}(x, v)$$

that is valid for all $u(x)$ and $v(x)$ from $[H^1(\Omega_2)]^3$. Then the equation

$$\mathcal{E}(x, v) - \mathcal{E}(x, u) = \mathcal{E}(x, Az + b)$$

holds. This equation implies the following equality:

$$\begin{aligned} &\langle \mathcal{E}(x_k, v) - \mathcal{E}(x_l, v), \mathcal{E}(x_k, I) - \mathcal{E}(x_l, I) \rangle - \langle \mathcal{E}(x_k, u) - \mathcal{E}(x_l, u), \mathcal{E}(x_k, I) - \mathcal{E}(x_l, I) \rangle \\ &\quad = \langle \mathcal{E}(x_k, Az + b) - \mathcal{E}(x_l, Az + b), \mathcal{E}(x_k, I) - \mathcal{E}(x_l, I) \rangle \\ &\quad = \langle A(\mathcal{E}(x_k, I) - \mathcal{E}(x_l, I)), \mathcal{E}(x_k, I) - \mathcal{E}(x_l, I) \rangle = 0. \end{aligned}$$

The last equality results from the antisymmetry of the matrix A. □

We have previously concluded that the displacements (14) between points x_k and x_l can be approximated by expression (19). However, this statement is not correct for the functions from $[H^1(\Omega_2)]^3$, but (19) can be approximated by the expression

$$\frac{\langle \mathcal{E}(x_k, u) - \mathcal{E}(x_l, u), \mathcal{E}(x_k, I) - \mathcal{E}(x_l, I) \rangle}{\|\mathcal{E}(x_k, I) - \mathcal{E}(x_l, I)\|}, \tag{26}$$

which is correct in $[H^1(\Omega_2)]^3$. Moreover, Lemma 4 shows that for any solution to problem $\mathcal{D}_1$, expression (26) remains the same.

Let us formulate our task as follows:

- ▷ Suppose that short steel bars are installed on the boundary Γ_1 and the points at the ends of these bars are marked $x_1, x_2, \ldots, x_N$.
- ▷ We measure the distances between points x_k, x_l in the situation shown in Fig. 2a.
- ▷ After removing the domain $\tilde{\tilde{\Omega}}$ (Fig. 2b), we re-measure these distances (Fig. 2c).
- ▷ Let us denote the difference of these distances by $d(x_k, x_l)$, which corresponds to relation (14).
- ▷ Let us define the function $h(\cdot,\cdot,\cdot)$ on the space $S^{\text{sym}} \times \Gamma_1 \times \Gamma_1$ by relation

$$h(\tau, x_k, x_l) = \frac{\langle \mathcal{E}(x_k, u^\tau) - \mathcal{E}(x_l, u^\tau), \mathcal{E}(x_k, I) - \mathcal{E}(x_l, I)\rangle}{\|\mathcal{E}(x_k, I) - \mathcal{E}(x_l, I)\|}, \tag{27}$$

where $u^\tau(x) = u_2^\tau(x) - u_1^\tau(x)$ and $u_1^\tau(x) \in [H^1(\Omega_1)]^3$, $u_2^\tau(x) \in [H^1(\Omega_2)]^3$ are solutions to problem $\mathcal{D}_2$ with the force boundary conditions P generated by the symmetric tensor τ.

Note that the function $h(\cdot,\cdot,\cdot)$ is continuous on $S^{\text{sym}} \times \Gamma_1 \times \Gamma_1$, which results from (11) and Lemma 3. The equation

$$h(z\tau + y\sigma, x_k, x_l) = zh(\tau, x_k, x_l) + yh(\sigma, x_k, x_l) \tag{28}$$

follows from the definition of $h(\cdot,\cdot,\cdot)$, where $\tau, \sigma \in S^{\text{sym}}$ and $z, y \in \mathbb{R}$.

We determine the initial stress tensor τ as the minimum of the following functional

$$\min_{\tau \in S^{\text{sym}}} \sum_{(k,l)\in K} (h(\tau, x_k, x_l) - d(x_k, x_l))^2, \tag{29}$$

where $K \subset \{1, \ldots, N\} \times \{1, \ldots, N\}$. Moreover, if $(k,l) \in K$, then $k < l$. The set K contains indexes of pairs of points at which we measure the distances. A minimum of functional (29) exists, as the following lemma claims.

Lemma 5. *There is $\tau \in S^{\text{sym}}$ which is the minimum of functional* (29).

P r o o f. Let V_1 be a subspace of S^{sym} such that τ belongs to this subspace when

$$h(\tau, x_k, x_l) = 0 \quad \text{for all } (k,l) \in K.$$

Let V_2 be a subspace of S^{sym} such that $V_1 \oplus V_2 = S^{\text{sym}}$. Then

$$\sum_{(k,l)\in K} (h(\tau, x_k, x_l) - d(x_k, x_l))^2 \to \infty,$$

if $\|\tau\| \to \infty$ and $\tau \in V_2$, where $\|\tau\| = \sqrt{\tau_{ij}\tau_{ij}}$. Hence, there is a minimum on V_2 and (28) shows that τ is a minimum of (29) on S^{sym}. □

The proof of the lemma suggests that the minimum need not be determined uniquely, which is contrary to reality. The proof of Lemma 5 shows that the minimum is unique if $V_1 = 0$. It is necessary to select measuring points in an appropriate way, which we will leave to the next section.

Let us reformulate our problem into a form more suitable for further analysis. Consider the following tensors from S^{sym}

$$\tau^1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \tau^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \tau^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$
$$\tau^4 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \tau^5 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \tau^6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Equality (28) implies that we can overwrite expression (29) in the form

$$\min_{z \in \mathbb{R}^6} \sum_{(k,l) \in K} \left(\sum_{m=1}^{6} z_m h(\tau^m, x_k, x_l) - d(x_k, x_l) \right)^2, \tag{30}$$

where $z = (z_1, z_2, z_3, z_4, z_5, z_6)$.

Solving (30) is equivalent to solving the system of linear equations

$$Zz = b, \tag{31}$$

where Z is a matrix 6×6 and b is a vector from $\mathbb{R}^6$ that are defined in the following way:

$$\begin{aligned} Z_{mn} &= \sum_{(k,l) \in K} h(\tau^m, x_k, x_l) h(\tau^n, x_k, x_l), \\ b_m &= \sum_{(k,l) \in K} h(\tau^m, x_k, x_l) d(x_k, x_l). \end{aligned} \tag{32}$$

The definition of the matrix Z implies that Z is symmetric and nonnegative, which means that the inequality

$$z^\top Z z \geqslant 0$$

holds for every $z \in \mathbb{R}^6$. To determine the initial stress tensor, the matrix Z must be nonsingular. The question arises how to select points at which we measure distances in an optimal way. We address this problem in the next section.

4. Optimal selection of measuring points

Our problem is reduced to proper assembly of the matrix Z, i.e. it is necessary to select the optimal number of measuring points and their location on the boundary Γ_1. From the previous analysis it is clear that the set K must contain at least six pairs of measuring points. In geomechanical practice it is a big problem to ensure the accuracy of measurements. The matrix Z must be designed so that small changes of the right-hand side of system (31) do not cause big changes of the solution z. If δb is the change of the right-hand side of system (31), then δz is the change of the solution to that system, which can be expressed as follows:

$$Z(z + \delta z) = (b + \delta b).$$

Then the formula

$$\frac{\|\delta z\|}{\|z\|} \leqslant \|Z\| \|Z^{-1}\| \frac{\|\delta b\|}{\|b\|}$$

holds.

The proof can be found in [6]. The symbol $\|z\|$ is a vector norm in $\mathbb{R}^6$ and $\|Z\|$ is a matrix norm generated by this vector norm. The number $\|Z\| \|Z^{-1}\|$ is the condition number of the matrix Z and is denoted by $\kappa(Z)$. The last relation shows that the small condition number leads to less dependence of the solution on the inaccuracy of measurements. Thus, the initial stress tensor is determined more reliably. If Z is a symmetric positively definite matrix and $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^6$, then the relation

$$\kappa(Z) = \frac{\lambda^{\max}}{\lambda^{\min}} \tag{33}$$

holds, where $\lambda^{\min}$ is the smallest eigenvalue of Z and $\lambda^{\max}$ is the largest eigenvalue of Z. The reader can find the proof of (33) e.g. in the book [6].

We will now look at optimal selection of measuring points and deal with pairs of sets K, X that satisfy the following relations:

$$\begin{gathered} X = \{x_1, \ldots, x_N\} \subset \Gamma_1, \\ K \subset \{1, \ldots, N\} \times \{1, \ldots, N\}, \quad (i, j) \in K \Rightarrow i < j. \end{gathered} \tag{34}$$

For each such pair of sets we can construct the matrix Z using formula (32) and denote it by $Z(K, X)$. We will assume that $\kappa(Z(K, X)) = \infty$ when the matrix $Z(K, X)$ is singular. Let X be a finite subset of Γ_1 and let us define $M(X)$ as

$$M(X) = \min_K \kappa(Z(K, X)), \tag{35}$$

where K, X satisfy (34). The optimum defined in this way works with pairs of points that are selected from the set X.

A general solution to our problem is associated with the number M defined as

$$M = \inf_{K,X} \kappa(Z(K, X)),$$

where the sets K, X are arbitrary and satisfy (34). The following theorem illustrates the relationship between $M(X)$ and M.

Theorem 1. *For every $\varepsilon > 0$ there exists $\delta > 0$ such that for every set $Y = \{y_1, \dots, y_L\} \subset \Gamma_1$ that satisfies*

$$\Gamma_1 \subset \bigcup_{i=1}^{L} B^\delta(y_i), \tag{36}$$

where

$$B^\delta(y_i) = \{y \in \mathbb{R}^3 \mid \|y - y_i\| < \delta\},$$

the inequality

$$M + \varepsilon > M(Y)$$

holds.

Proof. Let $\varepsilon > 0$, then there exist $X = \{x_1, \dots, x_N\} \subset \Gamma_1$ and K satisfying (34) such that the inequality

$$M + \frac{\varepsilon}{2} > \kappa(Z(K, X)) \tag{37}$$

holds. If we consider the definition $Z(K, X)$, then (32) implies

$$Z_{mn}(K, X) = \sum_{(k,l) \in K} h(\tau^m, x_k, x_l) h(\tau^n, x_k, x_l),$$

where the equality

$$h(\tau^n, x_k, x_l) = \frac{\langle \mathcal{E}(x_k, u^n) - \mathcal{E}(x_l, u^n), \mathcal{E}(x_k, I) - \mathcal{E}(x_l, I) \rangle}{\|\mathcal{E}(x_k, I) - \mathcal{E}(x_l, I)\|}$$

holds, which follows from (27). The function $u^n(x) = u_2^n(x) - u_1^n(x)$, where $u_1^n(x)$ and $u_2^n(x)$ are solutions to problem $\mathcal{D}_2$ with the force boundary conditions generated by τ^n, as described in Section 3. Lemma 3 implies that $h(\tau^n, x_k, x_l)$ is continuous at points x_k and x_l, thus the matrix $Z(K, X)$ continuously depends on x_i, $i = 1, \dots, N$. In [1] it is proved that eigenvalues of any matrix continuously depend on the elements of this matrix. Thus, for $\varepsilon/2 > 0$ there exists $\delta > 0$ such that the inequality

$$|\kappa(Z(K, X)) - \kappa(Z(K, \overline{X}))| < \frac{\varepsilon}{2} \tag{38}$$

holds for any $\overline{X} = \{\bar{x}_1, \dots, \bar{x}_N\}$ satisfying

$$|x_i - \bar{x}_i| < \delta, \quad i = 1, \dots, N. \tag{39}$$

In addition, δ is so small that

$$|x_i - x_j| > \delta, \quad i \neq j \tag{40}$$

holds. Take an arbitrary set $Y = \{y_1, \dots, y_L\} \subset \Gamma_1$ satisfying (36), then from (40) it follows that $L \geqslant N$ and we can arrange the points $y_1, \dots, y_L$ so that

$$|x_i - y_i| < \delta, \quad i = 1, \dots, N,$$

holds. Relationships (38), (39) and the last inequalities give the relation

$$|\kappa(Z(K,X)) - \kappa(Z(K,Y))| < \frac{\varepsilon}{2}.$$

The last relation with (38) give the inequalities

$$M + \varepsilon > \kappa(Z(K,Y)) > M(Y),$$

which hold for any Y satisfying (36). The last inequality implies the statement of this theorem. □

This theorem shows how to find the optimal distribution of measuring points and the pairs of measuring points. Let us describe the procedure how to choose a suitable set X and a set K so that the condition number of the matrix $Z(K,X)$ is close to the optimal value of M.

Let us formulate the procedure:

- ▷ Choose the set $Y = \{y_1, \dots, y_L\} \subset \Gamma_1$ such that it satisfies (36) for a sufficiently small δ.
- ▷ Let us choose the set $K \subset \{1, \dots, N\} \times \{1, \dots, N\}$ such that $\kappa(Z(K,Y))$ is minimal for all K satisfying condition (34).
- ▷ Select the set $X \subset Y$ so that it contains only those points that lie at the ends of the line represented by the pair of numbers belonging to the set K.

It is clear from the construction of the set X that $M(X)$ is equal to $M(Y)$. It follows from Theorem 1 that $M(X)$ converges to M when δ converges to zero. Numerical experiments have shown that it is possible to restrict oneself to sets K, which have only six elements. This corresponds to the fact that the original stress tensor has six independent components. Then the set X also has few elements, which shows that we can install relatively few measuring points, which is advantageous from the measurement point of view. One such example will be presented in the following section. In addition, it should be emphasized that this procedure is completely independent of the measurements.

5. Numerical solution

If we want to determine the original stress tensor, we need to solve twelve auxiliary tasks and get twelve functions u_1^i, u_2^i, $i = 1, \ldots, 6$, such that $u_1^i \in [H^1(\Omega_1)]^3$ and $u_2^i \in [H^1(\Omega_2)]^3$. We assemble the matrix Z using these auxiliary solutions and selecting a sufficiently dense set of points $X \subset \Gamma_1$. Then we select the set K using criterion (35).

In real situations the exact solutions are approximated numerically, which means that the functions u_1^i, u_2^i, $i = 1, \ldots, 6$, are replaced by the sequences of numerical solutions. Let us use the following notation:

- ▷ $u_{1,n}^i$, $u_{2,n}^i$, $i = 1, \ldots, 6$, are approximations of functions u_1^i, u_2^i,
- ▷ $Z(K, X)$ and $Z_n(K, X)$ are matrices constructed using precise solutions u_1^i, u_2^i and approximated solutions $u_{1,n}^i$, $u_{2,n}^i$, where sets K and X represent a selection of measuring points on Γ_1,
- ▷ the symbol τ corresponds to the initial stress tensor obtained by the solution of system (31) and τ_n is the initial stress tensor obtained by the solution of the same system, where the matrix $Z(K, X)$ is replaced by the matrix $Z_n(K, X)$.

The question arises what happens when we replace the exact solutions with approximate ones in the procedures described above and how the initial stress tensor differs from the tensor obtained by means of precise solutions. The answer is given in the following theorem.

Theorem 2. *Let*

$$u_{1,n}^i \to u_1^i \text{ in } [H^1(\Omega_1)]^3, \quad u_{2,n}^i \to u_2^i \text{ in } [H^1(\Omega_2)]^3,$$
$$\kappa(Z(K, X)) < \infty,$$

for $n \to \infty$. *Then*

$$Z_n(K, X) \to Z(K, X), \quad \kappa(Z_n(K, X)) \to \kappa(Z(K, X)), \quad \tau_n \to \tau.$$

P r o o f. The matrix $Z(K, X)$ is composed by formula (32) using the functions $h(\tau^i, x_k, x_l)$ defined by formula (27). The matrix $Z_n(K, X)$ is composed in the same way, where the functions $h(\tau^i, x_k, x_l)$ are replaced by the functions

$$h_n(\tau^i, x_k, x_l) = \frac{\big\langle \mathcal{E}(x_k, u_n^i) - \mathcal{E}(x_l, u_n^i)), \mathcal{E}(x_k, I) - \mathcal{E}(x_l, I) \big\rangle}{\|\mathcal{E}(x_k, I) - \mathcal{E}(x_l, I)\|},$$

where $u_n^i = u_{2,n}^i - u_{1,n}^i$. The functions u_n^i belong to $[H^1(\Omega_2)]^3$ as well as the functions u^i in formula (27). From Lemma 3 it follows that

$$h_n(\tau^i, x_k, x_l) \to h(\tau^i, x_k, x_l)$$

when $n \to \infty$.

Considering how the matrices $Z_n(K, X)$ and $Z(K, X)$ are assembled, we prove the first statement of this theorem.

Considering that eigenvalues continuously depend on the elements of a matrix, we have

$$\lambda_n^{\min} \to \lambda^{\min}, \quad \lambda_n^{\max} \to \lambda^{\max}$$

when $n \to \infty$, where $\lambda_n^{\min}$, $\lambda_n^{\max}$ are minimum and maximum eigenvalues of the matrix $Z_n(K, X)$.

Considering the assumptions $\kappa(Z(K, X)) < \infty$ we have the proof of the second statement of this theorem.

Considering the same assumptions and the first statement of this theorem, we have

$$(Z_n(K, X))^{-1} \to (Z(K, X))^{-1}$$

when $n \to \infty$. Since

$$\tau_n = \sum_{j=1}^{6} \tau^j z_{j,n},$$

where $z_{j,n}$ is a solution to system (32) in which $Z(K, X)$ is replaced by $Z_n(K, X)$, the last limit proves the last statement of this theorem. □

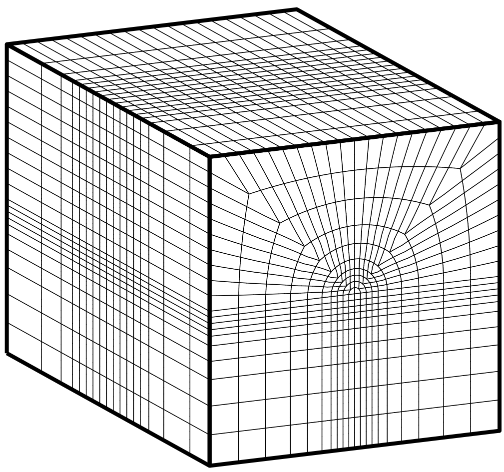

Figure 4. Finite element net of the numerical example.

The choice of measuring points and pairs of these points play an important role. The correct choice can significantly reduce the condition number of the matrix Z and thus affect the reliability of the determination of the initial stress tensor. This fact is demonstrated by a simple numerical experiment shown in Fig. 4. This figure

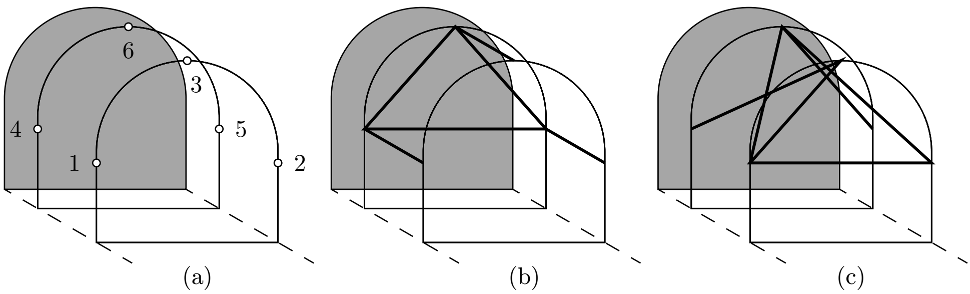

Figure 5. Example of selection of measuring points.

shows a domain of $40 \times 40 \times 90$ m with a tunnel of $4 \times 4 \times 50$ m and discretized by the finite element method. The tunnel is excavated in a homogeneous isotropic rock with Young's modulus $E = 65$ GPa and Poisson's ratio $\nu = 0.25$.

Pair of points	1 4	2 5	3 6	4 5	4 6	5 6
Calculated disp. [mm]	−0.5449	−0.2837	−0.2944	0.0435	0.5025	0.0687
Rounded disp.	−0.5	−0.3	−0.3	0	0.5	0.1

Table 1. Calculated displacements and their rounded values for the set K_1.

Pair of points	1 2	1 5	1 6	2 6	5 4	5 6
Calculated disp. [mm]	0.4807	1.5636	0.3113	−0.2371	0.9346	0.0687
Rounded disp.	0.5	1.6	0.3	−0.2	0.9	0.1

Table 2. Calculated displacements and their rounded values for the set K_2.

Stress tensor	s_x	s_y	s_z	t_{xv}	t_{yz}	t_{zx}
Exact	15.00	3.00	10.00	1.50	2.00	0.00
Rounded for K_1	14.3907	2.8778	10.8788	1.1524	1.8642	−0.4626
Rounded for K_2	14.4631	2.9897	9.9823	1.4758	1.8720	−2.1293

Table 3. Initial stress tensors reconstructed from the exact calculated displacements and their rounded values for the set K_1 and the set K_2.

The selection of measuring points and pairs of points is performed according to the procedure described at the end of Section 4. In our experiment, the set Y corresponds to the points of a finite element net located on the wall of the excavated part of the tunnel in Fig. 4. Now choose a subset $X \subset Y$ and six pairs of points represented by the set K_1 so that the number $\kappa(Z(K_1, X))$ is minimal and is equal to 26. The set X is shown in Fig. 5 (a) and the corresponding pair of points in Fig. 5 (b). Consider the other six pairs of points represented by the set K_2 and shown in Fig. 5 (c). In this case, $\kappa(Z(K_2, X))$ is equal to 648. All operations performed so far are independent of measurements. To demonstrate the usefulness of the presented method, let us choose an initial stress tensor and use (10) to generate boundary conditions. The values of this tensor are given in the first row of Table 3 and its visualization in the main directions is given in Fig. 6 (a). Then we calculate the changes in distance between the selected points. These values for the set K_1 are given in the first row of Table 1 and the values for the set K_2 are given in the first row of Table 2. The second rows in Tables 1 and 2 show rounded values, which demonstrate possible inaccuracies in the measurement process. When we reconstruct the initial stress tensor from exact values, we get a tensor that is identical to the originally selected tensor for both sets K_1 and K_2. If we use rounded values for the reconstruction of the tensor, we obtain different tensors, which are listed in the second and third rows of Table 3 and are visualized in the main directions in Fig. 6 (b) and Fig. 6 (c). This numerical experiment shows that the selection of points and their pairs can significantly reduce the effect of measurement errors.

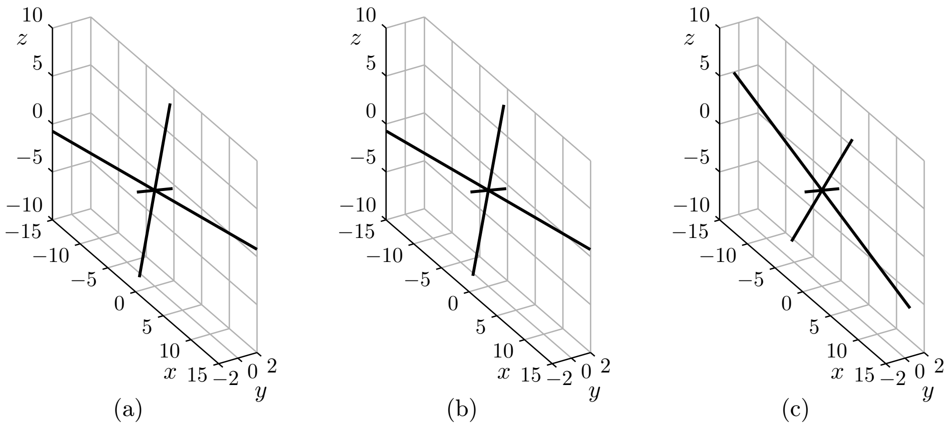

Figure 6. Initial stress tensors visualized in the main directions and their reconstruction from (a) the exact values, (b) rounded values for the set K_1 and (c) rounded values for the set K_2.

The method described above was used to determine the initial stress tensor at Dolní Rožná in the Czech Republic in several cases, but the shapes of underground openings were much more complicated than the simple numerical experiment presented above. There is an underground laboratory there, which serves as a model of radioactive waste repository. Auxiliary problems were solved by the program GEM [2], which was developed at Institute of Geonics and which allows to solve the first problem of the theory of elasticity as described in Section 2. It is possible to use any commercial program that allows to solve the first boundary problem of the theory of elasticity. The construction of the matrix Z and the optimization procedure, which allows to select a suitable set of measuring points, were written within the postprocesing package of SW system GEM.

The initial stress tensor values obtained by the method described above were compared with those obtained by the hydraulic fracture method described in [3], and we reached a good agreement. These results were summarized in the research report [9]. The method suggested in this paper only requires measuring distances between selected pairs of points before and after the excavation of a certain volume of the rock. The authors think that the presented method is faster and easier to use than the methods used so far.

6. Conclusion

A new method of determining the initial stress tensor has been designed and tested. The method is based on the appropriate selection of measuring points on the walls of the underground opening and measuring the distance after the removal of the rock during the excavation of the underground opening the procedure for selection of measuring points guaranteeing maximum accuracy of the determination of the initial stress tensor is part of the solution. In this paper we have focused on mathematical aspects of this problem and limited ourselves to a very simple domain to demonstrate the basic principle of this method. The disadvantage of this method is that it can only be used in the process of building underground openings.

References

[1] *R. Bhatia*: Matrix Analysis. Graduate Texts in Mathematics 169. Springer, New York, 1997. zbl MR doi

[2] *R. Blaheta, O. Jakl, R. Kohut, J. Starý*: GEM – a platform for advanced mathematical geosimulations. Parallel Processing and Applied Mathematics. Part 1. Lecture Notes in Computer Science 6067. Springer, Berlin, 2010, pp. 266–275. doi

[3] *B. C. Haimson, F. H. Cornet*: ISRM suggested methods for rock stress estimation—Part 3: Hydraulic fracturing (HF) and/or hydraulic testing of pre-existing fractures (HTPF). Int. J. Rock Mech. Min. Sci. *40* (2003), 1011–1020. doi

[4] *P. R. Halmos*: Measure Theory. Graduate Texts in Mathematics 18. Springer, Berlin, 1974. zbl MR doi

[5] *J. A. Hudson, F. H. Cornet, R. Christiansson*: ISRM suggested methods for rock stress estimation—Part 1: Strategy for rock stress estimation. Int. J. Rock Mech. Min. Sci. *40* (2003), 991–998. doi

[6] *P. D. Lax*: Linear Algebra and Its Applications. Pure and Applied Mathematics. Wiley, New York, 2007. zbl MR

[7] *J. Nečas, I. Hlaváček*: Mathematical Theory of Elastic and Elasto-Plastic Bodies: An Introduction. Studies in Applied Mechanics 3. Elsevier, Amsterdam, 1981. zbl MR doi

[8] *J. Sjöberg, R. Christiansson, J. A. Hudson*: ISRM suggested methods for rock stress estimation—Part 2: Overcoring methods. Int. J. Rock Mech. Min. Sci. *40* (2003), 999–1010. doi

[9] *K. Souček et al.*: Comprehensive Geological Characterization of URF Bukov—Part II. Geotechnical Characterization. Final report no. 221/2018/ENG. Ostrava, SÚRAO, 2017.

[10] *K. Sugawara, Y. Obara*: Draft ISRM suggested method for in situ stress measurement using the compact conical-ended borehole overcoring (CCBO) technique. Int. J. Rock Mech. Min. Sci. *36* (1999), 307–322. doi

[11] *T. D. Wiles, P. K. Kaiser*: *In situ* stress determination using the under-excavation technique—I. Theory. Int. J. Rock Mech. Min. Sci. & Geomech. Abstr. *31* (1994), 439–446. doi

[12] *T. D. Wiles, P. K. Kaiser*: *In situ* stress determination using the under-excavation technique—II. Applications. Int. J. Rock Mech. Min. Sci. & Geomech. Abstr. *31* (1994), 447–456. doi

Authors' address: *Josef Malík* (corresponding author), *Alexej Kolcun*, Institute of Geonics, Czech Academy of Sciences, Studetská 1768, 708 00 Ostrava-Poruba, Czech Republic, e-mail: `josef.malik@ugn.cas.cz`, `alexej.kolcun@ugn.cas.cz`.