ISSN 1990-4789, Journal of Applied and Industrial Mathematics, 2024, Vol. 18, No. 1, pp. 150–166.

Migration Velocity Analysis Using a Ray Method Asymptotics of the Double Square Root Equation

N. N. Shilov[1*] **and A. A. Duchkov**[2**]

[1]*Novosibirsk State University, Novosibirsk, 630090 Russia*

[2]*Trofimuk Institute of Petroleum Geology and Geophysics, Siberian Branch, Russian Academy of Sciences, Novosibirsk, 630090 Russia*

e-mail: *n.shilov@g.nsu.ru, **DuchkovAA@ipgg.sbras.ru

Received April 17, 2023; revised May 2, 2023; accepted June 7, 2023

Abstract—Seismic images of subsurface structures are the most valuable outcome of seismic data processing. The image quality is strongly affected by the accuracy of background velocity model. In this paper, we develop a gradient-descent velocity update algorithm based on our original high-frequency asymptotics of the Double Square Root equation, i.e., a special one-way approximation of the wave equation describing single-scattered wave field only. We propose a loss function consistent with widely adopted imaging condition and derive equations for its gradient computation. We test our method on noise-free synthetic datasets in 2D settings.

Keywords: *seismic inverse problem, velocity analysis, double square root equation, ray method, perturbation theory*

DOI: 10.1134/S1990478924010137

INTRODUCTION

Seismic exploration is widely used to study the internal structure of the Earth using reflected waves data [1]. Historically, there have been two groups of methods for reconstructing the velocity model of the environment: some are aimed at reconstructing smooth velocity distributions (velocity analysis) and others are aimed at reconstructing discontinuities (migration). These methods occur in different positions in the data processing flow, and, in particular, the accuracy of seismic migration is determined by the accuracy of the reconstructed smooth velocity model [2].

Two approaches stand out among the methods of velocity analysis, differing from each other in the choice of optimization criterion: seismic tomography (searching for a velocity model that satisfies the recorded data [3]) and migration velocity analysis (searching for a velocity model that satisfies the physics of reflection: rays of incident and reflected waves converge at the reflection points, and the position of these points does not depend on the angles of incidence [4, 5]). Accordingly, the direct problems that are solved during the selection process differ. In tomographic approaches, modeling problems are solved at each step, and data continuation problems are solved in migration velocity analysis.

Continuation of data into the lower half-space plays a key role in seismic migration. It is implemented as a solution of the wave equation or its approximations in reverse time by a finite-difference [4, 6] or asymptotic [7, 8] method. The present paper develops an algorithm for migration velocity analysis based on the high-frequency asymptotics of one of the approximations to the wave equation.

1. STATEMENT OF THE INVERSE PROBLEM

1.1. Characteristic DSR Equation

Consider the two-dimensional half-space $z \geq 0$, in which a smooth distribution of elastic wave velocity $v(x, z)$ is given. Let the sources and receivers of seismic waves be placed on the upper boundary of this half-space at the coordinates x_s and x_r, respectively. Let also the travel time of a single-reflected wave be specified as a function of the horizontal coordinates of the source and receiver, as well as (formally) of the observation system's depth, $\tau(x_s, x_r, z = 0)$. This function satisfies the reciprocity principle

$$\tau(x_s, x_r, z) = \tau(x_r, x_s, z)$$

and a special form of the eikonal equation,

$$\frac{\partial \tau}{\partial z} = -\sqrt{\frac{1}{v(x_s,z)^2} - \left(\frac{\partial \tau}{\partial x_s}\right)^2} - \sqrt{\frac{1}{v(x_r,z)^2} - \left(\frac{\partial \tau}{\partial x_r}\right)^2}, \tag{1}$$

obtained for the first time in the paper [7]. When deriving this equation, it was assumed that both radical expressions are positive. Physically, this means that the rays of incident and reflected waves do not become horizontal anywhere. In our work we will adhere to the same assumption, but in what follows, instead of the full notation for $v(x_s, z)$ and $v(x_r, z)$, we use the shorthand notation

$$v_s = v(x_s, z),$$
$$v_r = v(x_r, z).$$

Equation (1) describes the kinematics of seismic waves in the data space (x_s, x_r, z). Its solution specifies the travel time of the wave from the point $(x_s, z)^T$ down to the reflecting boundary and back to the point $(x_r, z)^T$ (the superscript "^{T}" indicates transposition). It has found application in the field of seismic migration, where it is associated with a certain pseudodifferential equation (Double Square Root equation), which makes it possible to extrapolate the recorded wave field into the lower half-space and construct images of reflecting boundaries [4]. As for other equations describing wave processes, for the Double Square Root equation one can construct ray asymptotics and use it to solve direct and inverse problems.

1.2. DSR Equation Rays

Following the notation of the ray method, we introduce the coordinate vector and the slowness vector,

$$\vec{x} = (x_s, x_r, z)^T,$$
$$\vec{p} = (p_s, p_r, p_z)^T = \nabla_{\vec{x}} \tau = \left(\frac{\partial \tau}{\partial x_s}, \frac{\partial \tau}{\partial x_r}, \frac{\partial \tau}{\partial z}\right)^T.$$

Here and throughout, the operator ∇ will be used to denote the gradient, and the subscript next to it will indicate the coordinates the partial derivatives are taken with respect to. Unlike the coordinate vector and the slowness vector in the physical space, in the data space the vectors $\vec{x}$ and $\vec{p}$ do not correspond to any one physical point and direction—they specify the position of the "source–receiver" pair at a certain depth and the derivatives of the travel time with respect to independent horizontal movements of the source and receiver or with their simultaneous "sinking" in the vertical direction.

Using the eikonal equation (1) one can construct several different Hamiltonians [8], and different forms of the Hamiltonian will correspond to different parametrizations of the rays. In our paper, we will need to trace rays in reverse time, and therefore we will use the Hamiltonian proposed in [9],

$$H(\vec{x}, \vec{p}) = C_\tau(\vec{x}, \vec{p}) H_0(\vec{x}, \vec{p}), \tag{2}$$

where

$$C_\tau(\vec{x}, \vec{p}) = \frac{1}{\dfrac{1}{v_s^2\sqrt{\dfrac{1}{v_s^2} - p_s^2}} + \dfrac{1}{v_r^2\sqrt{\dfrac{1}{v_r^2} - p_r^2}}},$$
$$H_0(\vec{x}, \vec{p}) = -\left(p_z + \sqrt{\frac{1}{v_s^2} - p_s^2} + \sqrt{\frac{1}{v_r^2} - p_r^2}\right). \tag{3}$$

Such a Hamiltonian specifies a system of equations for a ray parameterized by the reflected wave travel time,

$$\begin{aligned} \frac{d\vec{x}}{d\tau} &= \nabla_{\vec{p}} H, \\ \frac{d\vec{p}}{d\tau} &= -\nabla_{\vec{x}} H. \end{aligned} \tag{4}$$

Expressions for the derivatives of the Hamiltonian are given in the Appendix.

We will call the solutions of this system—the pairs $(\vec{x}(\tau), \vec{p}(\tau))$—the rays. We will set the initial conditions on the measurement surface $z = 0$. Having fixed the source x_s^{obs}, the receiver x_r^{obs}, the travel time $\tau_{\text{obs}} = \tau(x_s^{\text{obs}}, x_r^{\text{obs}}, 0)$, and its derivatives with respect to horizontal coordinates, we can express the unknown vertical component of slowness from the eikonal equation (1) to obtain the initial conditions in the form

$$\vec{x}\,\big|_{\tau=\tau_{\text{obs}}} = \begin{pmatrix} x_s^{\text{obs}} \\ x_r^{\text{obs}} \\ 0 \end{pmatrix}, \quad \vec{p}\,\big|_{\tau=\tau_{\text{obs}}} = \begin{pmatrix} \dfrac{\partial \tau_{\text{obs}}}{\partial x_s} \\[2mm] \dfrac{\partial \tau_{\text{obs}}}{\partial x_r} \\[2mm] -\sqrt{\dfrac{1}{v_s^2} - p_s^2}\,\bigg|_{\tau=\tau_{\text{obs}}} - \sqrt{\dfrac{1}{v_r^2} - p_r^2}\,\bigg|_{\tau=\tau_{\text{obs}}} \end{pmatrix}. \tag{5}$$

By construction, for such initial values of $\vec{x}$ and $\vec{p}$ the eikonal equation is satisfied and the Hamiltonian is equal to zero. These properties will be preserved along the entire ray as along the solution of the Hamiltonian system.

1.3. Statement of the Inverse Problem

We will solve the ray tracing system (4) with the initial conditions (5) in reverse time, from $\tau = \tau_{\text{obs}}$ to $\tau = 0$. The zero travel time for a reflected wave means that the wave travels the entire path from the source to the receiver in no time. Physically, this is realized only when the source and receiver are combined and located directly at the reflection point, on the very interface between the media. Such considerations can be used as a criterion for the accuracy of the velocity model [5]. In terms of the DSR equation rays, we will assume that the velocity model is incorrect if $x_s|_{\tau=0} \neq x_r|_{\tau=0}$. This principle is illustrated in Fig. 1.

Let the observation system consist of K pairs of sources and receivers at coordinates x_s^k and x_r^k, respectively, $k = 1, \dots, K$. For each pair, we denote the measured travel time of the wave by τ_{obs}^k, and its derivatives with respect to horizontal coordinates,by $\frac{\partial \tau_{\text{obs}}^k}{\partial x_s}$ and $\frac{\partial \tau_{\text{obs}}^k}{\partial x_r}$. Using these data, we construct the DSR equation rays using system (4) and the initial conditions (5). Let us denote by h_k the distances between $x_s^k\big|_{\tau=0}$ and $x_r^k\big|_{\tau=0}$,

$$h_k = \left(x_r^k - x_s^k\right)\big|_{\tau=0}. \tag{6}$$

The quantities h_k depend on the velocity model $v(x, z)$. Let us compose the residual functional

$$L(v) = \frac{1}{K} \sum_{k=1}^{K} h_k^2(v). \tag{7}$$

This functional is non-negative and takes zero value in the true speed model. Let us formulate the inverse problem.

Inverse DSR problem. Construct a velocity model $\tilde{v}(x, z)$ in which the functional (7) takes a minimum value.

We will seek its solution using the gradient-descent method in a certain class of velocity models $v(x, z; \vec{c})$, parameterized by the set of numbers $\vec{c} = (c_1, c_2, \dots, c_M)^T$. Note that in the case of

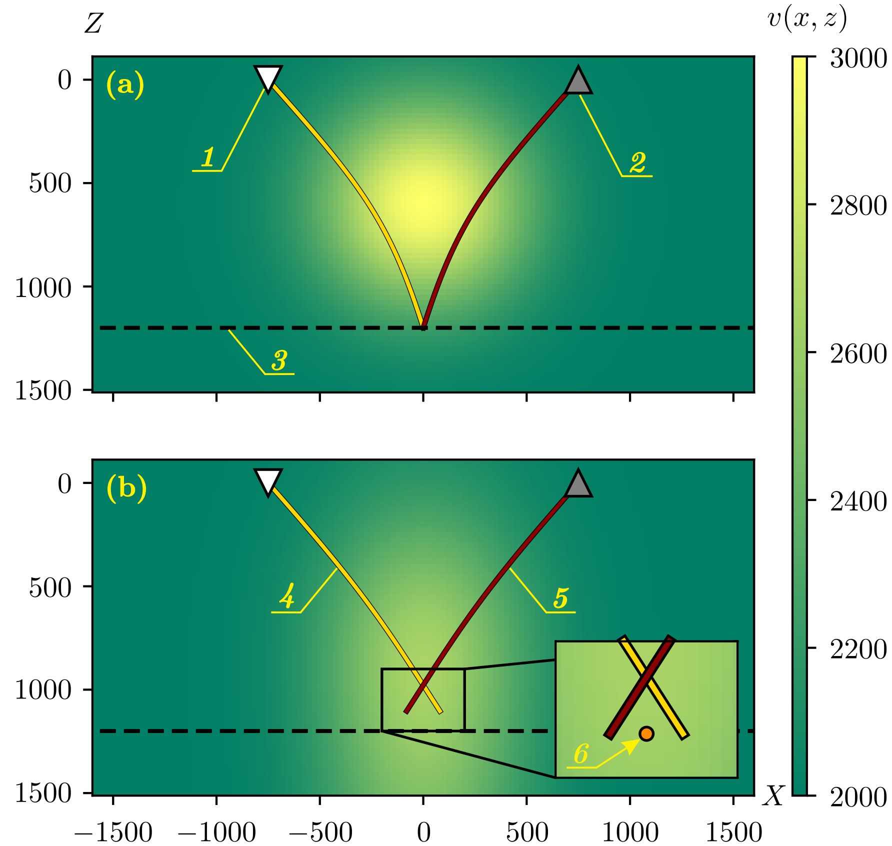

Fig. 1. Optimization principle: (a) correct velocity model ($v(x,z)$ [m/s], Z [m], X [m]), (b) incorrect velocity model. Legend: (*1*) source, (*2*) receiver, (*3*) true position of the reflecting boundary, (*4*) incident ray $(x_s, z)^T$, (*5*) reflected ray $(x_r, z)^T$, (*6*) midpoint—the approximate position of the reflection point.

an arbitrarily complex structure of the medium, the solution may not exist due to the condition on the nowhere-horizontal wave propagation inherent to the eikonal equation (1). The uniqueness of the solution in the case of moderately inhomogeneous media probably depends on the number of reflecting boundaries in the model. In the next section, we will construct a system of equations that allows one to estimate the sensitivity of rays to perturbations of the velocity model, specify the choice of parametrization of the solution, and, finally, create a formula for calculating the gradient of the objective functional.

Remark 1. Even if the velocity model is not completely correct but is close to the true one, the midpoint coordinates

$$m_x^k = \left.\frac{x_r^k + x_s^k}{2}\right|_{\tau=0}, \quad m_z^k = z^k\big|_{\tau=0} \tag{8}$$

will indicate the approximate position of the reflection point. Thus, by constructing all K rays of the double square root equation, we can estimate the shape of the reflecting boundary. However, if the model does not approximate the true velocity distribution well, then the shape of the boundary "visible" to the rays may be distorted. An example of one midpoint is shown at the bottom of Fig. 1.

2. GRADIENT OF OBJECTIVE FUNCTIONAL

2.1. DSR Ray Perturbation System

Consider a DSR equation ray $(\vec{x}(\tau), \vec{p}(\tau))$ constructed in the model $v(x,z)$. Let us introduce a small perturbation $\delta v(x,z)$ into the velocity model. It will cause small perturbations in the

Hamiltonian $H(\vec{x},\vec{p})$, the ray trajectory $\vec{x}(\tau)$, and the slowness $\vec{p}(\tau)$,

$$v \to v + \delta v \implies \begin{array}{c} H \to H + \delta H, \\ \vec{x} \to \vec{x} + \delta\vec{x}, \\ \vec{p} \to \vec{p} + \delta\vec{p}, \end{array}$$

with δH linearly depending on the perturbations of the ray and the velocity model,

$$\delta H(\vec{x},\vec{p};\delta\vec{x},\delta\vec{p},\delta v_s,\delta v_r) = \nabla_{\vec{x}} H \cdot \delta\vec{x} + \nabla_{\vec{p}} H \cdot \delta\vec{p} + \frac{\partial H}{\partial v_s}\delta v_s + \frac{\partial H}{\partial v_r}\delta v_r, \tag{9}$$

where δv_s and δv_r are defined similarly to v_s and v_r and the velocity derivatives of the Hamiltonian are found by symbolic differentiation of (2) with respect to v_s and v_r as by independent variables.

Let us write down the system of equations for the perturbed ray,

$$\begin{aligned} \frac{d}{d\tau}(\vec{x} + \delta\vec{x}) &= \nabla_{\vec{p}}(H + \delta H), \\ \frac{d}{d\tau}(\vec{p} + \delta\vec{p}) &= -\nabla_{\vec{x}}(H + \delta H). \end{aligned}$$

Recall that the system (4) with an unperturbed Hamiltonian on the right-hand side is satisfied on the original ray $(\vec{x},\vec{p})$. By reducing the corresponding terms, we obtain the system of ray perturbations

$$\begin{aligned} \frac{d}{d\tau}\delta\vec{x} &= \nabla_{\vec{p}}\delta H = \nabla_{\vec{p}}\nabla_{\vec{x}} H \cdot \delta\vec{x} + \nabla_{\vec{p}}\nabla_{\vec{p}} H \cdot \delta\vec{p} + \nabla_{\vec{p}}\frac{\partial H}{\partial v_s}\delta v_s + \nabla_{\vec{p}}\frac{\partial H}{\partial v_r}\delta v_r, \\ \frac{d}{d\tau}\delta\vec{p} &= -\nabla_{\vec{x}}\delta H = -\nabla_{\vec{x}}\nabla_{\vec{x}} H \cdot \delta\vec{x} - \nabla_{\vec{x}}\nabla_{\vec{p}} H \cdot \delta\vec{p} - \nabla_{\vec{x}}\frac{\partial H}{\partial v_s}\delta v_s - \nabla_{\vec{x}}\frac{\partial H}{\partial v_r}\delta v_r, \end{aligned} \tag{10}$$

in which the operators $\nabla_{\vec{x}}\nabla_{\vec{x}}$, $\nabla_{\vec{p}}\nabla_{\vec{p}}$, and $\nabla_{\vec{x}}\nabla_{\vec{p}} = \nabla_{\vec{p}}{\nabla_{\vec{x}}}^T$ act as

$$\nabla_{\vec{x}}\nabla_{\vec{x}} = \begin{pmatrix} \dfrac{\partial^2}{\partial {x_s}^2} & \dfrac{\partial^2}{\partial x_s \partial x_r} & \dfrac{\partial^2}{\partial x_s \partial z} \\ \dfrac{\partial^2}{\partial x_r \partial x_s} & \dfrac{\partial^2}{\partial {x_r}^2} & \dfrac{\partial^2}{\partial x_r \partial z} \\ \dfrac{\partial^2}{\partial z \partial x_s} & \dfrac{\partial^2}{\partial z \partial x_r} & \dfrac{\partial^2}{\partial z^2} \end{pmatrix},$$

$$\nabla_{\vec{p}}\nabla_{\vec{p}} = \begin{pmatrix} \dfrac{\partial^2}{\partial {p_s}^2} & \dfrac{\partial^2}{\partial p_s \partial p_r} & \dfrac{\partial^2}{\partial p_s \partial p_z} \\ \dfrac{\partial^2}{\partial p_r \partial p_s} & \dfrac{\partial^2}{\partial {p_r}^2} & \dfrac{\partial^2}{\partial p_r \partial p_z} \\ \dfrac{\partial^2}{\partial p_z \partial p_s} & \dfrac{\partial^2}{\partial p_z \partial p_r} & \dfrac{\partial^2}{\partial {p_z}^2} \end{pmatrix},$$

$$\nabla_{\vec{x}}\nabla_{\vec{p}} = \nabla_{\vec{p}}{\nabla_{\vec{x}}}^T = \begin{pmatrix} \dfrac{\partial^2}{\partial x_s \partial p_s} & \dfrac{\partial^2}{\partial x_s \partial p_r} & \dfrac{\partial^2}{\partial x_s \partial p_z} \\ \dfrac{\partial^2}{\partial x_r \partial p_s} & \dfrac{\partial^2}{\partial x_r \partial p_r} & \dfrac{\partial^2}{\partial x_r \partial p_z} \\ \dfrac{\partial^2}{\partial z \partial p_s} & \dfrac{\partial^2}{\partial z \partial p_r} & \dfrac{\partial^2}{\partial z \partial p_z} \end{pmatrix}.$$

Write equations (10) in matrix form as

$$\frac{d}{d\tau}\begin{pmatrix} \delta\vec{x} \\ \delta\vec{p} \end{pmatrix} = \begin{pmatrix} \nabla_{\vec{p}}\nabla_{\vec{x}} H & \nabla_{\vec{p}}\nabla_{\vec{p}} H \\ -\nabla_{\vec{x}}\nabla_{\vec{x}} H & -\nabla_{\vec{x}}\nabla_{\vec{p}} H \end{pmatrix} \cdot \begin{pmatrix} \delta\vec{x} \\ \delta\vec{p} \end{pmatrix} + \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{C} & -\mathbf{D} \end{pmatrix} \cdot \begin{pmatrix} \delta\vec{v}_s \\ \delta\vec{v}_r \end{pmatrix}, \tag{11}$$

where by $\delta\vec{v}_s$ and $\delta\vec{v}_r$ we have denoted the vectors

$$\delta\vec{v}_s = \begin{pmatrix} \delta v_s \\ \dfrac{\partial}{\partial x_s}\delta v_s \\ \dfrac{\partial}{\partial z}\delta v_s \end{pmatrix}, \quad \delta\vec{v}_r = \begin{pmatrix} \delta v_r \\ \dfrac{\partial}{\partial x_r}\delta v_r \\ \dfrac{\partial}{\partial z}\delta v_r \end{pmatrix}$$

and the matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ have the structure

$$\mathbf{A} = \begin{pmatrix} \dfrac{\partial}{\partial p_s}\dfrac{\partial H}{\partial v_s} & 0 & 0 \\ \dfrac{\partial}{\partial p_r}\dfrac{\partial H}{\partial v_s} & 0 & 0 \\ \dfrac{\partial}{\partial p_z}\dfrac{\partial H}{\partial v_s} & 0 & 0 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} \dfrac{\partial}{\partial p_s}\dfrac{\partial H}{\partial v_r} & 0 & 0 \\ \dfrac{\partial}{\partial p_r}\dfrac{\partial H}{\partial v_r} & 0 & 0 \\ \dfrac{\partial}{\partial p_z}\dfrac{\partial H}{\partial v_r} & 0 & 0 \end{pmatrix},$$

$$\mathbf{C} = \begin{pmatrix} \dfrac{\partial}{\partial x_s}\dfrac{\partial H}{\partial v_s} & \dfrac{\partial H}{\partial v_s} & 0 \\ \dfrac{\partial}{\partial x_r}\dfrac{\partial H}{\partial v_s} & 0 & 0 \\ \dfrac{\partial}{\partial z}\dfrac{\partial H}{\partial v_s} & 0 & \dfrac{\partial H}{\partial v_s} \end{pmatrix}, \qquad \mathbf{D} = \begin{pmatrix} \dfrac{\partial}{\partial x_s}\dfrac{\partial H}{\partial v_r} & 0 & 0 \\ \dfrac{\partial}{\partial x_r}\dfrac{\partial H}{\partial v_r} & \dfrac{\partial H}{\partial v_r} & 0 \\ \dfrac{\partial}{\partial z}\dfrac{\partial H}{\partial v_r} & 0 & \dfrac{\partial H}{\partial v_r} \end{pmatrix}.$$

System (11) is a linear inhomogeneous one. Let $\mathbf{P}(\tau)$ denote the fundamental matrix of its homogeneous part. Let us express the perturbation of the ray at the zero travel time in terms of this matrix,

$$\left.\begin{pmatrix} \delta\vec{x} \\ \delta\vec{p} \end{pmatrix}\right|_{\tau=0} = \mathbf{P}^{-1}(\tau_{\text{obs}}) \cdot \left.\begin{pmatrix} \delta\vec{x} \\ \delta\vec{p} \end{pmatrix}\right|_{\tau=\tau_{\text{obs}}} - \int\limits_0^{\tau_{\text{obs}}} \mathbf{P}^{-1}(\tau') \cdot \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{C} & -\mathbf{D} \end{pmatrix} \cdot \begin{pmatrix} \delta\vec{v}_s \\ \delta\vec{v}_r \end{pmatrix} d\tau'. \tag{12}$$

To set conditions at the point $\tau = \tau_{\text{obs}}$, we will assume that the positions of the source and receiver are fixed, and measurements on the surface do not depend on perturbations of the velocity model,

$$\delta\vec{x}|_{\tau=\tau_{\text{obs}}} = \vec{0}, \quad \delta p_s|_{\tau=\tau_{\text{obs}}} = \delta\frac{\partial\tau_{\text{obs}}}{\partial x_s} = 0, \quad \delta p_r|_{\tau=\tau_{\text{obs}}} = \delta\frac{\partial\tau_{\text{obs}}}{\partial x_r} = 0. \tag{13}$$

In addition, we will seek solutions of system (10) on which the perturbed Hamiltonian is equal to zero,

$$H(\vec{x},\vec{p}) + \delta H(\vec{x},\vec{p};\delta\vec{x},\delta\vec{p}) = 0.$$

Here let us substitute δH from (9) and express δp_z, taking into account the fact that $H(\vec{x},\vec{p}) \equiv 0$ on the unperturbed ray,

$$\delta p_z = -\left(\nabla_{\vec{x}} H \cdot \delta\vec{x} + \frac{\partial H}{\partial p_s}\delta p_s + \frac{\partial H}{\partial p_r}\delta p_r + \frac{\partial H}{\partial v_s}\delta v_s + \frac{\partial H}{\partial v_r}\delta v_r\right)\left(\frac{\partial H}{\partial p_z}\right)^{-1}.$$

Finally, substituting the coordinates and slowness perturbations from (13), we obtain the condition on $\delta p_z|_{\tau=\tau_{\text{obs}}}$,

$$\delta p_z|_{\tau=\tau_{\text{obs}}} = -\left.\left(\frac{\partial H}{\partial v_s}\delta v_s + \frac{\partial H}{\partial v_r}\delta v_r\right)\left(\frac{\partial H}{\partial p_z}\right)^{-1}\right|_{\tau=\tau_{\text{obs}}}. \tag{14}$$

System (10) with conditions (13) and (14) allows one to trace ray perturbations from the survey plane to zero reflection time for given perturbations of the velocity model. In particular, having solved this system for the kth ray, we can calculate the perturbation h_k from the residual functional (7) as the difference between the perturbations x_r^k and x_s^k at the zero reflection time.

2.2. Model Parametrization

Let us introduce some constraints on the velocity model. We narrow its domain to a rectangle symmetrical with respect to some point $(\overline{x}, \overline{z})^T$,

$$x \in [\overline{x} - d_x, \overline{x} + d_x],$$
$$z \in [\overline{z} - d_z, \overline{z} + d_z],$$

where the numbers $d_x > 0$ and $d_z > 0$ determine the rectangle sizes. We introduce the change of coordinates

$$\tilde{x} = \frac{x - \overline{x}}{d_x},$$
$$\tilde{z} = \frac{z - \overline{z}}{d_z}.$$

We will use a similar substitution for the coordinates of the source and receiver, $\tilde{x}_s$ and $\tilde{x}_r$. Note that the transformed coordinates are dimensionless, and their values range from -1 to 1.

As noted above, we will minimize the functional (7) in the class of velocity models parameterized by a certain set of numbers. More specifically, we will look for a solution of the optimization problem in the form

$$v(x, z; \vec{c}) = v_0(x, z) + \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} c_{ij} T_i(\tilde{x}) T_j(\tilde{z}), \tag{15}$$

where $v_0(x, z)$ denotes a fixed initial approximation, T_n are Chebyshev polynomials of the first kind of degree n, and c_{ij} are unknowns expansion coefficients (we will denote the entire set of these coefficients by $\vec{c}$, although in this case they can fill a two-dimensional array).

2.3. Gradient of the Objective Functional

Let the perturbation of the velocity model $\delta v(x, z)$ be caused by a small increment of the coefficient c_{ij}. Then

$$\delta v(x, z) = \frac{\partial v}{\partial c_{ij}} \delta c_{ij} = T_i(\tilde{x}) T_j(\tilde{z}) \delta c_{ij}.$$

The initial conditions (13), (14) transform to

$$\delta \vec{x}\,|_{\tau=\tau_{\text{obs}}} = \left.\frac{\partial \vec{x}}{\partial c_{ij}}\right|_{\tau=\tau_{\text{obs}}} \delta c_{ij},$$
$$\delta \vec{p}\,|_{\tau=\tau_{\text{obs}}} = \left.\frac{\partial \vec{p}}{\partial c_{ij}}\right|_{\tau=\tau_{\text{obs}}} \delta c_{ij}$$

with the partial derivatives

$$\left.\frac{\partial \vec{x}}{\partial c_{ij}}\right|_{\tau=\tau_{\text{obs}}} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

$$\left.\frac{\partial \vec{p}}{\partial c_{ij}}\right|_{\tau=\tau_{\text{obs}}} = -\left(\frac{\partial H}{\partial p_z}\right)^{-1} \left.\begin{pmatrix} 0 \\ 0 \\ \dfrac{\partial H}{\partial v_s} T_i(\tilde{x_s}) T_j(\tilde{z}) + \dfrac{\partial H}{\partial v_r} T_i(\tilde{x_r}) T_j(\tilde{z}) \end{pmatrix}\right|_{\tau=\tau_{\text{obs}}},$$

and the vectors $\delta\vec{v}_s$ and $\delta\vec{v}_r$, to the form

$$\delta\vec{v}_s = \frac{\partial\vec{v}_s}{\partial c_{ij}}\delta c_{ij} = \begin{pmatrix} T_i(\tilde{x}_s)T_j(\tilde{z}) \\ \dfrac{1}{d_x}\dfrac{\partial}{\partial\tilde{x}_s}T_i(\tilde{x}_s)T_j(\tilde{z}) \\ \dfrac{1}{d_z}\dfrac{\partial}{\partial\tilde{z}}T_i(\tilde{x}_s)T_j(\tilde{z}) \end{pmatrix}\delta c_{ij},$$

$$\delta\vec{v}_r = \frac{\partial\vec{v}_r}{\partial c_{ij}}\delta c_{ij} = \begin{pmatrix} T_i(\tilde{x}_r)T_j(\tilde{z}) \\ \dfrac{1}{d_x}\dfrac{\partial}{\partial\tilde{x}_r}T_i(\tilde{x}_r)T_j(\tilde{z}) \\ \dfrac{1}{d_z}\dfrac{\partial}{\partial\tilde{z}}T_i(\tilde{x}_r)T_j(\tilde{z}) \end{pmatrix}\delta c_{ij}.$$

By substituting these expressions into (12) and cancelling δc_{ij}, we can calculate the derivatives $\frac{\partial\vec{x}}{\partial c_{ij}}$ and $\frac{\partial\vec{p}}{\partial c_{ij}}$ at time zero,

$$\frac{\partial}{\partial c_{ij}}\begin{pmatrix}\vec{x}\\ \vec{p}\end{pmatrix}\Bigg|_{\tau=0} = \mathbf{P}^{-1}(\tau_{\mathrm{obs}})\cdot\frac{\partial}{\partial c_{ij}}\begin{pmatrix}\vec{x}\\ \vec{p}\end{pmatrix}\Bigg|_{\tau=\tau_{\mathrm{obs}}} - \int_0^{\tau_{\mathrm{obs}}}\mathbf{P}^{-1}(\tau')\cdot\begin{pmatrix}\mathbf{A} & \mathbf{B}\\ -\mathbf{C} & -\mathbf{D}\end{pmatrix}\cdot\frac{\partial}{\partial c_{ij}}\begin{pmatrix}\vec{v}_s\\ \vec{v}_r\end{pmatrix}d\tau'.$$

Finally, based on this, we can easily find the derivative of the objective functional (7) with respect to the perturbation of the coefficient c_{ij},

$$\frac{\partial L}{\partial c_{ij}} = \frac{1}{K}\frac{\partial}{\partial c_{ij}}\sum_{k=1}^{K}h_k^2 = \frac{2}{K}\sum_{k=1}^{K}h_k\frac{\partial h_k}{\partial c_{ij}},$$

where

$$\frac{\partial h_k}{\partial c_{ij}} = \frac{\partial x_r^k}{\partial c_{ij}}\bigg|_{\tau=0} - \frac{\partial x_s^k}{\partial c_{ij}}\bigg|_{\tau=0}.$$

Using these formulas, we can calculate the derivatives for all c_{ij}. Moreover, we note that the matrices **A**, **B**, **C**, **D**, and **P** do not depend on the choice of a specific coefficient and that this calculation can be done in parallel.

Remark 2. Our choice of basis functions is not the result of optimization; we were guided by three simple considerations:

- Chebyshev polynomials form an orthogonal basis.
- Chebyshev polynomials are twice smooth functions.
- Expansion in Chebyshev polynomials is easy to implement in software.

Therefore, instead of polynomial expansion, other parametric models can be used in 15). To apply our formalism, we only need to be able to calculate their derivatives with respect to spatial coordinates and with respect to the parameter vector $\vec{c}$ (in particular, mixed derivatives like $\frac{\partial}{\partial c_i}\frac{\partial v}{\partial x}$).

3. NUMERICAL EXPERIMENTS

3.1. Numerical Algorithms

Three main computational problems can be distinguished within the framework of our work:

1. Representation of the a priori velocity model $v_0(x, z)$,
2. Ray tracing and solution of ray perturbation system.
3. Updating model parameters in the antigradient direction.

To solve them, we used the NumPy [10] and SciPy [11] libraries of the Python programming language. To visualize the results, we used the library Matplotlib [12].

The a priori velocity model was specified on a rectangular grid with a step of 25 m, followed by bicubic interpolation [13]. To solve the ODE systems 4) and (10) at the stage of ray tracing and ray perturbation tracing, we used the fourth-order Runge–Kutta method with a constant step [14]. To implement gradient descent, we used a specialized function from the SciPy library. Among the optimization methods available to this function, we chose the "BFGS" method [15].

3.2. Models, Observation Systems, and Initial Approximations

When testing our method, we limited ourselves to synthetic data with known true velocity distributions. To calculate them, we used our own self-implemented two-point ray tracing [16, 17]. We considered two models: one with a smooth anomaly and a single reflective boundary and the other, divided by three boundaries into contrasting homogeneous layers,

$$v_I(x,z) = 2000 + 1000e^{-\left(\frac{x}{500}\right)^2 - \left(\frac{z-600}{500}\right)^2}[\text{m/s}],$$

$$v_{II}(x,z) = \begin{cases} 2000, & z \le 400 \\ 1500, & 400 < z \le 800 \\ 2500, & 800 < z \le 1200 \\ 2000, & 1200 < z \end{cases} \quad [\text{m/s}].$$

The two velocity models are shown in the upper parts of Figs. 2 and 3. Note that in the first model, a fictitious interface is placed at a depth of 1200 m. The second model is explicitly divided into four layers with flat horizontal boundaries at depths of 400, 800, and 1200 m.

We used the same observation system in both models. The sources were placed at coordinates from -750 to 750 m with a step of 50 m, and the receivers were placed with the same step within a range of one and a half kilometers on both sides of each source. As initial approximations $v_0(x,z)$ we used models of root-mean-square velocities calculated using formulas in [1]; a priori values of the series coefficients (15) were taken to be zero. Initial approximations are shown in the lower parts in Figs. 2 and 3. In addition to the models themselves and the positions of the boundaries, these figures depict clouds of midpoints—the approximate positions of the reflection points calculated in the initial velocity models (see formula (8) and Remark 1). It can be seen that in the first model the shape of the boundary is distorted, and in the second model, one of the boundaries is blurred and shifted in depth. In addition, the absolute values of the distances $h_k(v_0)$ (6) are displayed on separate color scales, from which it is clear that in both models there are rays that diverge significantly at zero time.

3.3. Results of Numerical Experiments

The accuracy of the velocity model reconstruction largely depends on the successful choice of parametrization (15): too low degrees of expansion do not allow a good representation of the desired anomalies, and high ones lead to overfitting—the velocity model is adjusted to individual rays with no single structure emerging despite the low residual values. We were unable to develop any recommendations on the optimal choice of complexity of polynomial expansion, and therefore we limited ourselves to an empirical search.

Models $v_I(x,z)$ and $v_{II}(x,z)$ differ in lateral and depth variability. The continuous model $v_I(x,z)$ is symmetric about the center of the anomaly, and to represent it, it is natural to take an equal number of horizontal and vertical expansion terms, corrected for the aspect ratio of the model. On the other hand, the horizontally layered model $v_{II}(x,z)$ does not change at all horizontally, but includes three discontinuities in the vertical direction, so it is logical to take a high degree of series expansion along the z-axis and a low one along the x-axis. In addition, for clarity of the experiment, we wanted to study models of equal complexity. During the search process, we settled on the following parameters:

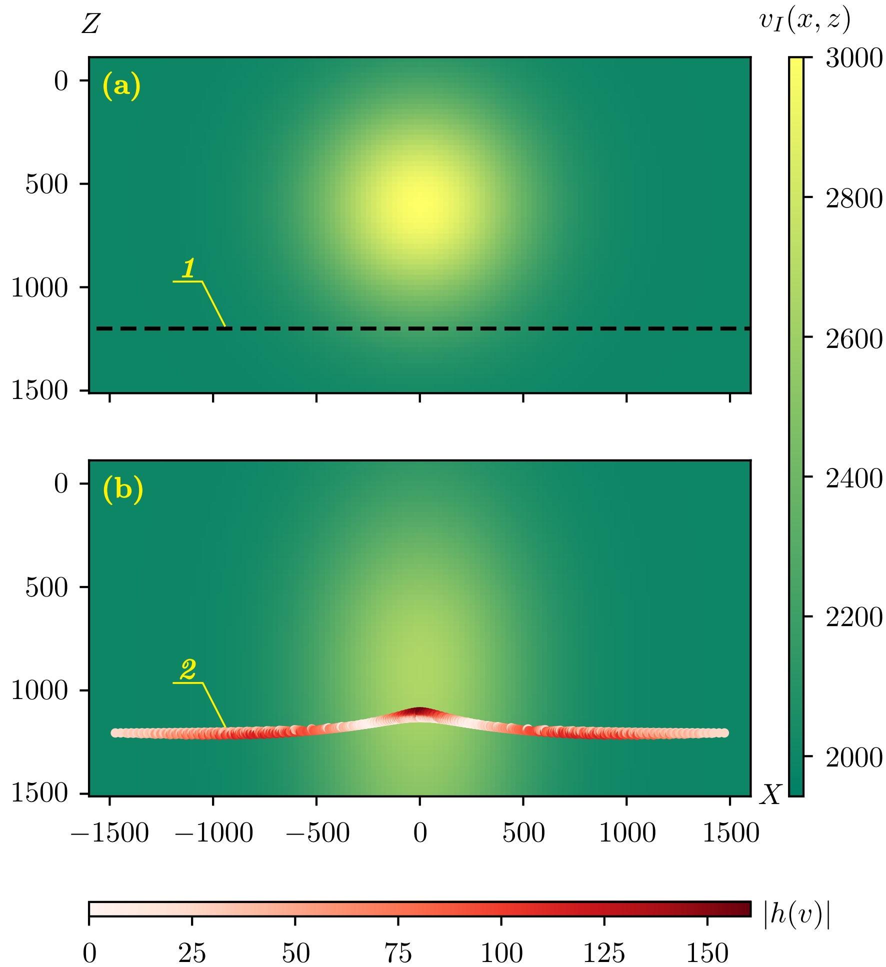

Fig. 2. (a) Model $v_I(x, z)$, [m/s], Z, [m], X, [m] and (b) initial guess ($|h(v)|$, [m] is the distance between source and receiver at zero time); (*1*) the true boundary position, (*2*) the cloud of midpoints—approximate positions of reflection points.

– In model $v_I(x, z)$: 26 terms of expansion horizontally and 9 terms vertically.

– In model $v_{II}(x, z)$: 9 terms of expansion horizontally and 26 terms vertically.

We minimized the functional (7) until the residual became less than 0.1 m^2, which corresponds to the maximum values of $h_k(v)$ at the level of 1 m. This is a heuristic choice, but we consider it justified: the characteristic wavelengths in seismic exploration are tens of meters, and further optimization under more realistic conditions will not add information about the true structure of the medium, but will only adjust the model to errors in the data. The optimization results are presented in Figs. 4 and 5 in the lower parts of the figures; the top parts show the true velocity models for comparison. Similar to Figs. 2 and 3, additional color scales show the distances between sources and receivers at zero reflection time. It can be seen that in both cases the shape and depth of the boundaries were reconstructed, but in the first model the reconstructed velocity anomaly is stretched vertically and its amplitude is underestimated. At the same time, in the second model it was possible to reconstruct three contrast layers without changing their power and quite accurately reconstructing the velocities. We attribute this to the larger amount of input data, including reflections from various depths. Local extrema within homogeneous layers can be explained by the nonuniform convergence of polynomial series to discontinuous functions.

3.4. Behavior of the Residual Functional in a Neighborhood of the Solution

To roughly estimate the uniqueness of the solutions obtained, we examined the behavior of the objective functional (7) in their neighborhood. Using finite-difference formulas, we estimated

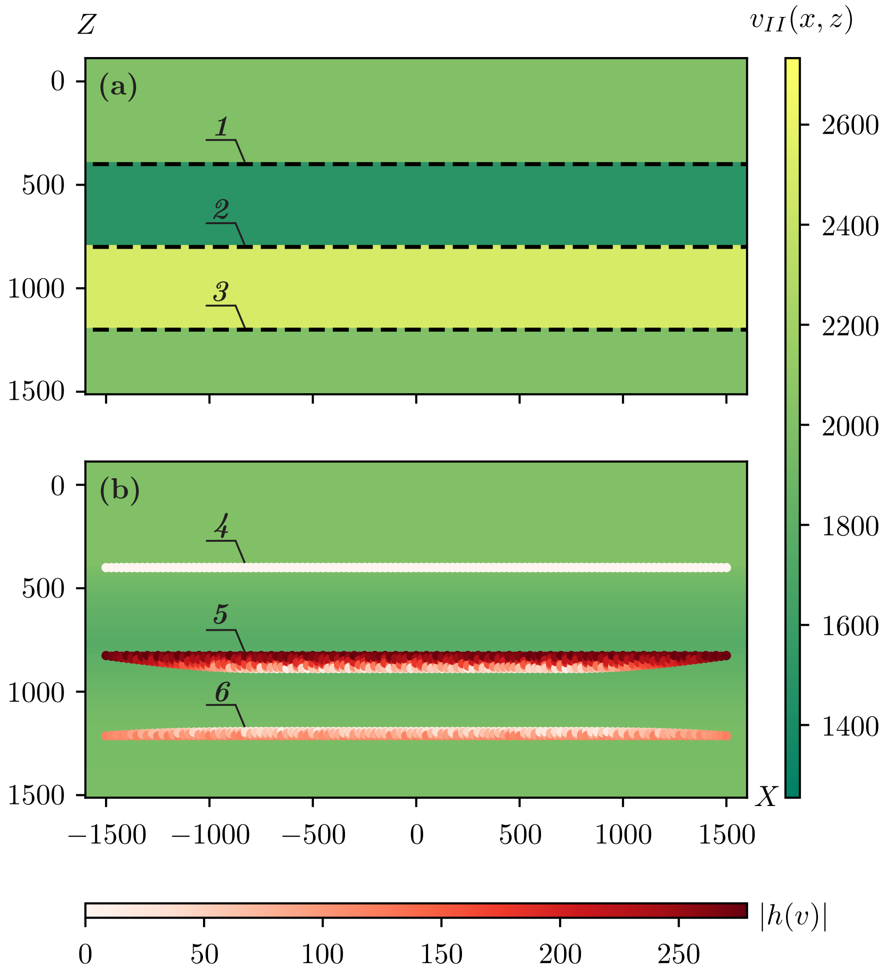

Fig. 3. (a) Model $v_{II}(x, z)$, [m/s], Z, [m], X, [m] and (b) initial guess ($|h(v)|$, [m] is the distance between source and receiver at zero time); (*1*), (*2*), and (*3*) are the true positions of boundaries, (*4*), (*5*), and (*6*) are the clouds of midpoints—approximate positions of reflection points.

the matrices of the second derivatives of the residual at the minimum points and found their eigenvalues and vectors. We then constructed sections of the objective functional in three coordinate axes Q_1, Q_2, and Q_3, defined by the eigenvectors corresponding to the largest eigenvalues. Along these directions, one can expect the fastest growth of the residual functional near the minimum point.

Sections of the residual functional are shown in Fig. 6. The left column shows sections for the first model and the right column, for the second. The displacements from the found solutions along the corresponding eigenvectors are plotted along the axes. The empty areas in the right column mean that some rays became horizontal in the corresponding velocity models, and the residual functional was not defined. However, it is clear that in the three selected directions there is a rapid increase in the residual without local minima and pronounced gully. It should be remembered that the condition numbers of the matrices of second derivatives turned out to be at the level of 10^{13}, and in addition to the selected main directions in the model space (in our case, 234-dimensional), there are others along which the objective functional does not actually change, and if there is noise in the data, the solution may be ambiguous.

4. DISCUSSION

As input, the method presented in this paper takes the coordinates of sources and receivers, the travel times of single-reflected waves, and their derivatives, allowing for the simultaneous recon-

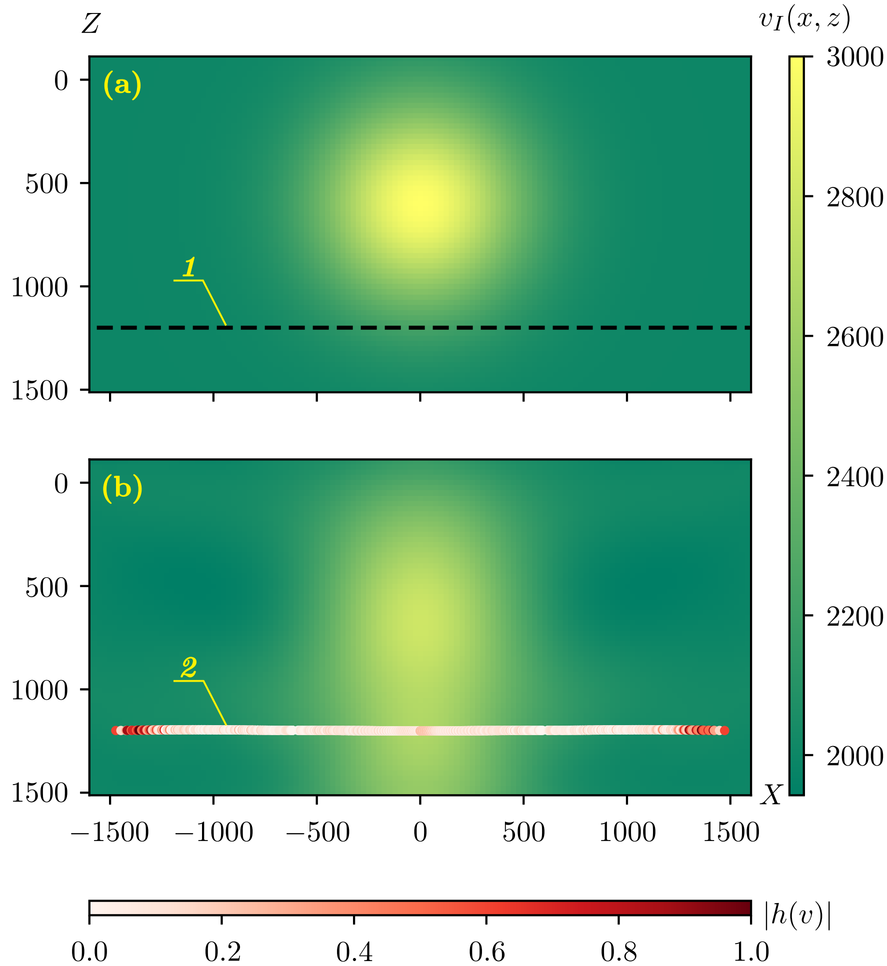

Fig. 4. (a) Model $v_I(x,z)$, [m/s], Z, [m], X, [m] and (b) optimization result ($|h(v)|$, [m] is the distance between source and receiver at zero time); (*1*) the true position of the boundary, (*2*) the position of the boundary reconstructed from the rays of the Double Square Root equation.

struction of a smooth velocity model and the positions of reflection points. Among the methods of velocity analysis known to the authors, it is most similar to stereotomography [18]. The novelty of the proposed approach is as follows:

– We use the characteristic double square root equation (1) rather than the classical eikonal equation [16, 17].
– We use a "physical" optimization criterion (the incident and reflected rays converge at the reflection point) rather than a data-fitting criterion.

The main advantage of our method is the smaller number of unknowns: in stereotomography, reflection points, reflection angles, and dip angles of boundaries are being adjusted in addition to the parameters of the velocity model. In our approach, a special form of the eikonal equation allows one to trace rays in reverse time and unambiguously determine the moment of reflection, and the chosen optimization principle includes only the parameters of the velocity model. The coordinates of the reflection points can be estimated after optimizing the velocity model using formula (8), and the wave reflection angles and the angles of incidence of the boundaries can be easily found from the ray trajectories of the Double Square Root equation [9].

The main disadvantage of the presented approach is the limitation on nowhere-horizontal wave propagation, which does not allow for the reconstruction of steeply dipping sections of reflecting boundaries and high-amplitude anomalies in velocity models.

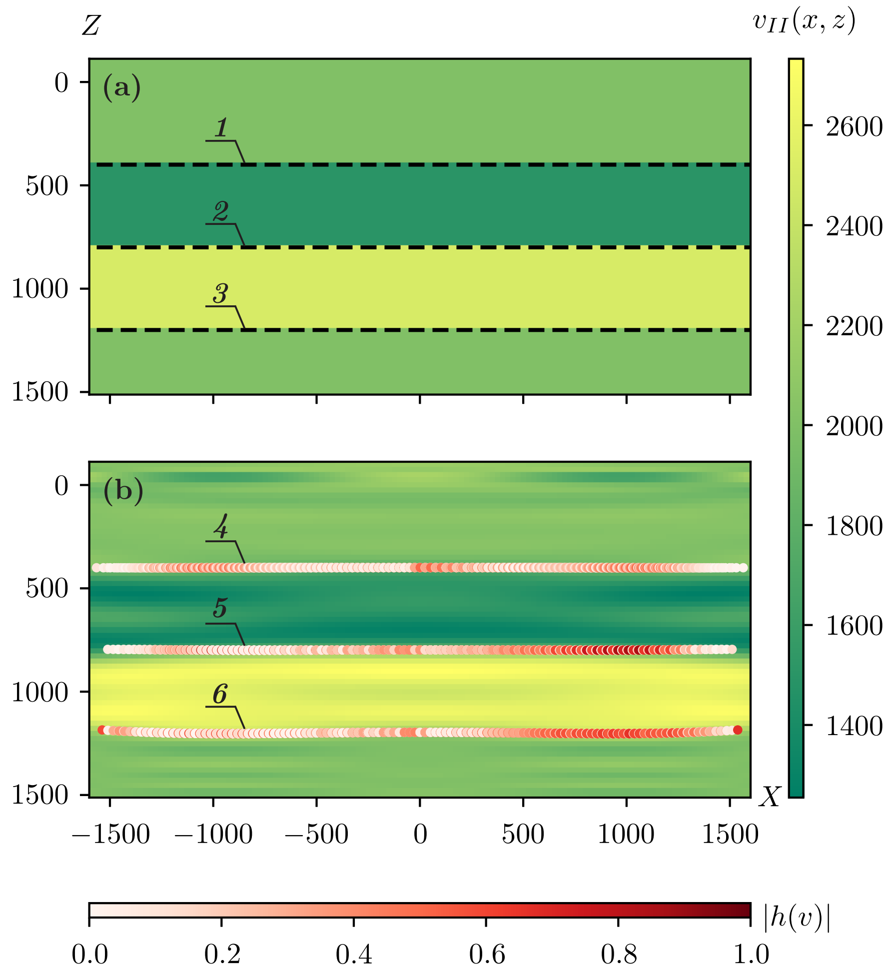

Fig. 5. (a) Model $v_{II}(x,z)$, [m/s], Z, [m], X, [m] and (b) optimization result ($|h(v)|$, [m] is the distance between source and receiver at zero time); (*1*), (*2*), and (*3*) are the true positions of boundaries, (*4*), (*5*), and (*6*) are the positions of boundaries reconstructed along the rays of the Double Square Root equation.

CONCLUSIONS

In our work, we presented a method of velocity analysis based on the simultaneous inversion of data on the travel times of reflected waves and their derivatives with respect to the coordinates of sources and receivers. The proposed method is based on the original high-frequency asymptotics of the Double Square Root equation—a special approximation to the wave equation that describes singly reflected waves in data space. In this work, an inverse problem is formulated and an objective functional is constructed, which we minimize using the gradient descent method in a certain class of parametric velocity models. Equations that allow one to calculate the gradient of the objective functional are given in the text of the paper. The main advantage of the proposed approach is the smaller number of optimized parameters compared to the closest analog that we are aware of, and the main disadvantage is the limitation on the complexity of the reconstructed model. The algorithm was tested on two synthetic datasets. All steps were carried out in a two-dimensional formulation of the problem.

APPENDIX

The equations in the text of the paper often included derivatives of the Hamiltonian (2). This appendix will give explicit formulas for calculating them. When writing these derivatives, we will take into account that all derivatives are calculated on the ray on which the Hamiltonian itself is equal to zero.

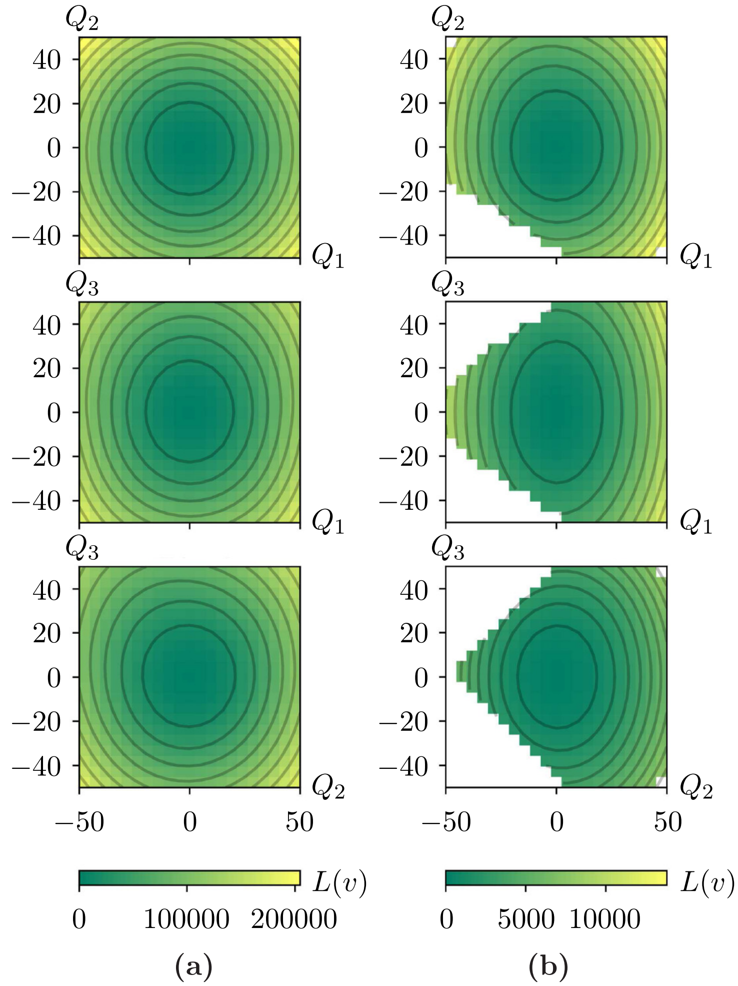

Fig. 6. Sections of the residual functional $L(v)$, [m^2], Q_1, [m/s], Q_2, [m/s], Q_3, [m/s] in a neighborhood of the found solutions of (a) model 1 and (b) model 2. Unfilled areas mean that in the corresponding models the rays became horizontal, and the residual functional was not defined.

Let us introduce some notation,

$$S_m = \frac{1}{v_s^2} - mp_s^2, \quad R_m = \frac{1}{v_r^2} - mp_r^2, \quad m = 1, 2, 3.$$

Let us rewrite the multipliers (3) composing the Hamiltonian (2),

$$C_\tau(\vec{x}, \vec{p}) = \frac{1}{\dfrac{1}{v_s^2\sqrt{S_1}} + \dfrac{1}{v_r^2\sqrt{R_1}}},$$

$$H_0(\vec{x}, \vec{p}) = -\left(p_z + \sqrt{S_1} + \sqrt{R_1}\right).$$

Let us write out their first derivatives

$$\nabla_{\vec{x}} C_\tau = C_\tau^2 \begin{pmatrix} \dfrac{S_2}{v_s^3 S_1^{3/2}} \dfrac{\partial v_s}{\partial x_s}, \\ \dfrac{R_2}{v_r^3 R_1^{3/2}} \dfrac{\partial v_r}{\partial x_r}, \\ \dfrac{S_2}{v_s^3 S_1^{3/2}} \dfrac{\partial v_s}{\partial z} + \dfrac{R_2}{v_r^3 R_1^{3/2}} \dfrac{\partial v_r}{\partial z} \end{pmatrix}, \quad \nabla_{\vec{p}} C_\tau = -C_\tau^2 \begin{pmatrix} \dfrac{p_s}{v_s^2 S_1^{3/2}}, \\ \dfrac{p_r}{v_r^2 R_1^{3/2}}, \\ 0 \end{pmatrix},$$

$$\nabla_{\vec{x}} H_0 = \begin{pmatrix} \dfrac{1}{v_s^3 S_1^{1/2}} \dfrac{\partial v_s}{\partial x_s}, \\ \dfrac{1}{v_r^3 R_1^{1/2}} \dfrac{\partial v_r}{\partial x_r}, \\ \dfrac{1}{v_s^3 S_1^{1/2}} \dfrac{\partial v_s}{\partial z} + \dfrac{1}{v_r^3 R_1^{1/2}} \dfrac{\partial v_r}{\partial z} \end{pmatrix}, \quad \nabla_{\vec{p}} H_0 = \begin{pmatrix} \dfrac{p_s}{S_1^{1/2}}, \\ \dfrac{p_r}{R_1^{1/2}}, \\ -1 \end{pmatrix},$$

$$\frac{\partial C_\tau}{\partial v_s} = C_\tau^2 \frac{S_2}{v_s^3 S_1^{3/2}}, \quad \frac{\partial C_\tau}{\partial v_r} = C_\tau^2 \frac{R_2}{v_r^3 R_1^{3/2}}, \quad \frac{\partial H_0}{\partial v_s} = \frac{1}{v_s^3 S_1^{1/2}}, \quad \frac{\partial H_0}{\partial v_r} = \frac{1}{v_r^3 R_1^{1/2}}.$$

Since $H_0 \equiv 0$ on the rays, the calculation of the second derivatives C_τ is unnecessary: they will appear exclusively in products with H_0. Let us write down the matrices of the second derivatives of the latter,

$$\nabla_{\vec{x}} \nabla_{\vec{x}} H_0 = \begin{pmatrix} \dfrac{1}{v_s^3 S_1^{1/2}} \dfrac{\partial^2 v_s}{\partial {x_s}^2} & 0 & \dfrac{1}{v_s^3 S_1^{1/2}} \dfrac{\partial^2 v_s}{\partial x_s \partial z} \\ 0 & \dfrac{1}{v_r^3 R_1^{1/2}} \dfrac{\partial^2 v_r}{\partial {x_r}^2} & \dfrac{1}{v_r^3 R_1^{1/2}} \dfrac{\partial^2 v_r}{\partial x_r \partial z} \\ \dfrac{1}{v_s^3 S_1^{1/2}} \dfrac{\partial^2 v_s}{\partial x_s \partial z} & \dfrac{1}{v_r^3 R_1^{1/2}} \dfrac{\partial^2 v_r}{\partial x_r \partial z} & \dfrac{1}{v_s^3 S_1^{1/2}} \dfrac{\partial^2 v_s}{\partial z^2} + \dfrac{1}{v_r^3 R_1^{1/2}} \dfrac{\partial^2 v_r}{\partial z^2} \end{pmatrix}$$

$$- \begin{pmatrix} \dfrac{\left(\dfrac{\partial v_s}{\partial x_s}\right)^2}{v_s^4 S_1^{3/2}} \left(\dfrac{1}{v_s^2} + S_3\right) & 0 & \dfrac{\dfrac{\partial v_s}{\partial x_s} \dfrac{\partial v_s}{\partial z}}{v_s^4 S_1^{3/2}} \left(\dfrac{1}{v_s^2} + S_3\right) \\ 0 & \dfrac{\left(\dfrac{\partial v_r}{\partial x_r}\right)^2}{v_r^4 R_1^{3/2}} \left(\dfrac{1}{v_r^2} + R_3\right) & \dfrac{\dfrac{\partial v_r}{\partial x_r} \dfrac{\partial v_r}{\partial z}}{v_r^4 R_1^{3/2}} \left(\dfrac{1}{v_r^2} + R_3\right) \\ \dfrac{\dfrac{\partial v_s}{\partial x_s} \dfrac{\partial v_s}{\partial z}}{v_s^4 S_1^{3/2}} \left(\dfrac{1}{v_s^2} + S_3\right) & \dfrac{\dfrac{\partial v_r}{\partial x_r} \dfrac{\partial v_r}{\partial z}}{v_r^4 R_1^{3/2}} \left(\dfrac{1}{v_r^2} + R_3\right) & \dfrac{\left(\dfrac{\partial v_s}{\partial z}\right)^2}{v_s^4 S_1^{3/2}} \left(\dfrac{1}{v_s^2} + S_3\right) + \dfrac{\left(\dfrac{\partial v_r}{\partial z}\right)^2}{v_r^4 R_1^{3/2}} \left(\dfrac{1}{v_r^2} + R_3\right) \end{pmatrix},$$

$$\nabla_{\vec{p}} \nabla_{\vec{p}} H_0 = \begin{pmatrix} \dfrac{1}{v_s^2 S_1^{3/2}} & 0 & 0 \\ 0 & \dfrac{1}{v_r^2 R_1^{3/2}} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \nabla_{\vec{x}} \nabla_{\vec{p}} H_0 = \nabla_{\vec{p}} {\nabla_{\vec{x}}}^T H_0 = \begin{pmatrix} \dfrac{p_s}{v_s^3 S_1^{3/2}} \dfrac{\partial v_s}{\partial x_s} & 0 & 0 \\ 0 & \dfrac{p_r}{v_r^3 R_1^{3/2}} \dfrac{\partial v_r}{\partial x_r} & 0 \\ \dfrac{p_s}{v_s^3 S_1^{3/2}} \dfrac{\partial v_s}{\partial z} & \dfrac{p_r}{v_r^3 R_1^{3/2}} \dfrac{\partial v_r}{\partial z} & 0 \end{pmatrix},$$

$$\nabla_{\vec{x}} \frac{\partial H_0}{\partial v_s} = - \begin{pmatrix} \dfrac{\dfrac{\partial v_s}{\partial x_s}}{v_s^4 S_1^{3/2}} \left(\dfrac{1}{v_s^2} + S_3\right) \\ 0 \\ \dfrac{\dfrac{\partial v_s}{\partial z}}{v_s^4 S_1^{3/2}} \left(\dfrac{1}{v_s^2} + S_3\right) \end{pmatrix}, \qquad \nabla_{\vec{x}} \frac{\partial H_0}{\partial v_r} = - \begin{pmatrix} 0 \\ \dfrac{\dfrac{\partial v_r}{\partial x_r}}{v_r^4 R_1^{3/2}} \left(\dfrac{1}{v_r^2} + R_3\right) \\ \dfrac{\dfrac{\partial v_r}{\partial z}}{v_r^4 R_1^{3/2}} \left(\dfrac{1}{v_r^2} + R_3\right) \end{pmatrix},$$

$$\nabla_{\vec{p}} \frac{\partial H_0}{\partial v_s} = \begin{pmatrix} \dfrac{p_s}{v_s^3 S_1^{3/2}} \\ 0 \\ 0 \end{pmatrix}, \qquad \nabla_{\vec{p}} \frac{\partial H_0}{\partial v_r} = \begin{pmatrix} 0 \\ \dfrac{p_r}{v_r^3 R_1^{3/2}} \\ 0 \end{pmatrix}.$$

Finally, we express the derivatives of the Hamiltonian (2) on the ray as

$$\begin{aligned}
\nabla_{\vec{x}} H &= C_\tau \nabla_{\vec{x}} H_0, \\
\nabla_{\vec{p}} H &= C_\tau \nabla_{\vec{p}} H_0, \\
\nabla_{\vec{x}} \nabla_{\vec{x}} H &= C_\tau \nabla_{\vec{x}} \nabla_{\vec{x}} H_0 + \nabla_{\vec{x}} C_\tau \otimes \nabla_{\vec{x}} H_0 + \nabla_{\vec{x}} H_0 \otimes \nabla_{\vec{x}} C_\tau, \\
\nabla_{\vec{p}} \nabla_{\vec{p}} H &= C_\tau \nabla_{\vec{p}} \nabla_{\vec{p}} H_0 + \nabla_{\vec{p}} C_\tau \otimes \nabla_{\vec{p}} H_0 + \nabla_{\vec{p}} H_0 \otimes \nabla_{\vec{p}} C_\tau, \\
\nabla_{\vec{x}} \nabla_{\vec{p}} H &= \nabla_{\vec{p}} \nabla_{\vec{x}}{}^T H = C_\tau \nabla_{\vec{x}} \nabla_{\vec{p}} H_0 + \nabla_{\vec{x}} C_\tau \otimes \nabla_{\vec{p}} H_0 + \nabla_{\vec{x}} H_0 \otimes \nabla_{\vec{p}} C_\tau, \\
\nabla_{\vec{x}} \frac{\partial H}{\partial v_s} &= C_\tau \nabla_{\vec{x}} \frac{\partial H_0}{\partial v_s} + \frac{\partial H_0}{\partial v_s} \nabla_{\vec{x}} C_\tau + \frac{\partial C_\tau}{\partial v_s} \nabla_{\vec{x}} H_0, \\
\nabla_{\vec{x}} \frac{\partial H}{\partial v_r} &= C_\tau \nabla_{\vec{x}} \frac{\partial H_0}{\partial v_r} + \frac{\partial H_0}{\partial v_r} \nabla_{\vec{x}} C_\tau + \frac{\partial C_\tau}{\partial v_r} \nabla_{\vec{x}} H_0, \\
\nabla_{\vec{p}} \frac{\partial H}{\partial v_s} &= C_\tau \nabla_{\vec{p}} \frac{\partial H_0}{\partial v_s} + \frac{\partial H_0}{\partial v_s} \nabla_{\vec{p}} C_\tau + \frac{\partial C_\tau}{\partial v_s} \nabla_{\vec{p}} H_0, \\
\nabla_{\vec{p}} \frac{\partial H}{\partial v_r} &= C_\tau \nabla_{\vec{p}} \frac{\partial H_0}{\partial v_r} + \frac{\partial H_0}{\partial v_r} \nabla_{\vec{p}} C_\tau + \frac{\partial C_\tau}{\partial v_r} \nabla_{\vec{p}} H_0.
\end{aligned}$$

Here the sign $\otimes$ denotes the Kronecker (tensor) product of vectors. If $\vec{a} = (a_1, a_2, \dots, a_M)^T$ and $\vec{b} = (b_1, b_2, \dots, b_N)^T$, then it is calculated as follows:

$$\vec{a} \otimes \vec{b} = \begin{pmatrix} a_1 b_1 & a_1 b_2 & \cdots & a_1 b_N \\ a_2 b_1 & a_2 b_2 & \cdots & a_2 b_N \\ \vdots & \vdots & \ddots & \vdots \\ a_M b_1 & a_M b_2 & \cdots & a_M b_N \end{pmatrix}.$$

All formulas were tested in the symbolic mathematics system `Wolfram Mathematica 13.2` [19].

FUNDING

The research was carried out within the framework of state assignments nos. FSUS-2022-0019 and FWZZ-2022-0017.

CONFLICT OF INTEREST

The authors of this work declare that they have no conflicts of interest.

REFERENCES

1. I. I. Gurvich and G. N. Boganik, *Seismic Exploration: A Textbook for Universities* (AIS, Tver, 2006) [in Russian].
2. N. Bleistein, J. K. Cohen, and W. Stockwell, Jr., *Mathematics of Multidimensional Seismic Imaging, Migration, and Inversion* (Springer, New York, 2001). https://doi.org/10.1007/978-1-4613-0001-4
3. T. Lo and P. L. Inderwiesen, *Fundamentals of Seismic Tomography* (SEG Books, Tulsa, 1994). https://doi.org/10.1190/1.9781560802334
4. J. F. Claerbout, *Imaging the Earth's Interior* (Blackwell, Oxford, 1985).

5. P. C. Sava and S. Fomel, "Angle-domain common-image gathers by wavefield continuation methods," Geophysics **68** (3), 1065–1074 (2003). https://doi.org/10.1190/1.1581078
6. Y. Zhang, S. Xu, N. Bleistein, and G. Zhang, "True-amplitude, angle-domain, common-image gathers from one-way wave-equation migrations," Geophysics **72** (1), S49–S58 (2007). https://doi.org/10.1190/1.2399371
7. A. V. Belonosova and A. S. Alekseev, "On one formulation of the inverse kinematic seismic problem for a two-dimensional inhomogeneous medium," in *Some Methods and Algorithms for Interpreting Geophysical Data* (Nauka, Moscow, 1967), 137–154 [in Russian].
8. A. A. Duchkov and M. V. de Hoop, "Extended isochron rays in prestack depth migration," Geophysics **75** (4), S139–S150 (2010). https://doi.org/10.1190/1.3459955
9. N. N. Shilov, "Dynamic ray tracing system for the double square root equation," *Master's Final Qualifying Thesis* (MMF NGU, Novosibirsk, 2021) [in Russian].
10. C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Rio, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, Ch. Gohlke, and T. E. Oliphant, "Array programming with NumPy," Nature **585** (7825), 357–362 (2020). https://doi.org/10.1038/s41586-020-2649-2
11. P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, I. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, Ch. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, "SciPy 1.0: fundamental algorithms for scientific computing in Python," Nature Methods **17** (3), 261–272 (2020). https://doi.org/10.1038/s41592-019-0686-2
12. J. D. Hunter, "Matplotlib: A 2D graphics environment," Comput. Sci. Eng. **9** (3), 90–95 (2007). https://doi.org/10.1109/MCSE.2007.55
13. G. I. Marchuk, *Methods of Computational Mathematics* (Nauka, Moscow, 1980) [in Russian].
14. V. P. Il'in, *Methods for Solving Ordinary Differential Equations* (IPTs NGU, Novosibirsk, 2017) [in Russian].
15. J. Nocedal and S. J. Wright, *Numerical Optimization* (Springer, New York, 2006).
16. V. Červený, *Seismic Ray Theory* (Cambridge Univ. Press, Cambridge, 2001).
17. A. S. Alekseev, V. M. Babich, and B. Ya. Gelchinskii, "Ray method for calculating the intensity of wave fronts," in *Questions of the Dynamic Theory of Seismic Wave Propagation* (1961), **5**, 3–24 [in Russian].
18. L. Gilles, "Stereotomography," Geophysics **73** (5), VE25–VE34 (2008). https://doi.org/10.1190/1.2952039
19. Wolfram Research, Inc., *Mathematica. Version 13.2* (Champaign, IL, 2022).