The problem of identification an unknown substance by the radiographic method

V G Nazarov¹

¹Institute of Applied Mathematics, Far-Eastern Branch of the Russian Academy of Sciences, Radio str., 7, Vladivostok, 690041 Russia.

E-mail: naz@iam.dvo.ru

Abstract. This paper considers the problem of partial identification of the chemical composition of an unknown medium by the multiple X-ray of this medium. A sample of an unknown substance is assumed to be homogeneous in its chemical composition, and the photon flux, collimated both in direction and in energy. A mathematical model for the identification problem is formulated. The approach proposed to solving the problem is based on the method of singular value decomposition of a matrix. At the first stage of the solution the problem is reduced to finding singular numbers and singular vectors for the series systems of algebraic equations linear with respect to products of unknown quantities. Then, based on the data obtained, a special function is built, called an indicator to the distinguishability of substances, which enables the sufficient conditions for the distinguishability of various substances. Based on the tabular data, calculations were made for a number of specific groups of chemical elements.

1. Introduction
The problem of identifying a substance as well as the problem of finding the chemical composition of a substance by radiographic methods, are interesting from the point of view of theory and have an undoubted practical value. Radiographic methods are useful in cases where it is required to perform nondestructive testing of the product or when the direct access to the object of study is difficult or undesirable. The number of scientific publications on this topic both by Russian and foreign researchers remains quite high; note among them [1,2]. At the same time, the authors use different approaches to solving the problem under consideration, and these approaches may significantly differ depending on a specific situation. In this paper, the problem of identifying an unknown substance is formulated, and the results of studying the behavior of some singular numbers associated with it are presented. An estimate of the maximum relative measurement error of radiation, at which the identification problem will be successfully solved for some specific groups of substances is also given.

2. Preliminary remarks and problem statement
Further, we consider that the investigated sample $G_0$ of the unknown substance $X_0$ is homogeneous in its chemical composition and all chemical elements that make up $X_0$ are present in some predetermined list of elements $X_1, X_2, ..., X_N$, which we know. The sample is irradiated with a photon flux collimated both in direction and in energy and going along a certain fixed straight line. The use of collimators before and after the substance under study makes it possible to isolate from the initial radiation flux mainly only those photons that did not interact with the substance. We neglect the interaction of photons
with the air medium and assume that the flux weakening occurs only in the section of the trajectory of the length \( l \) passing through the sample. In the course of each measurement experiment, all photons have the energy \( E_k \) from a fixed (discrete) set of radiation energies

\[
0.1\,\text{MeV} = E_1 < E_2 < \cdots < E_N = 20\,\text{MeV}.
\]

When carrying out calculations for specific substances and energies, we used numerical data taken from [3], where, in particular, all the information we need for 20 energy values is given. Therefore, we restrict ourselves to the case \( 2 \leq N \leq \tilde{N} = 20 \). Let \( h_k = h(E_k) \) and \( H_k = H(E_k) \) be the density of radiation fluxes entering and leaving \( G_0 \), for the energy \( E_k; k = 1, 2, \ldots, \tilde{N} \); \( \mu_{0k} = \mu_0(E_k) \) — radiation attenuation coefficient for the substance \( X_0 \); \( \mu_{xik} = \mu_x(E_k) \), \( i = 1, 2, \ldots, N \) is radiation attenuation coefficients for \( X_1, X_2, \ldots, X_N \); \( \rho_0 \) is the density of the substance \( X_0 \); \( \rho_x \) is the density of \( X_1 \); \( w_i \) is the mass fraction of the element \( X_i \), which is part of the substance \( X_0 \). Taking into account the physics of the radiation transfer process, we assume that always \( h_k, H_k, \mu_{0k}, \mu_{xik}, \rho_0, \rho_x > 0; \ w_i \geq 0; i = 1, 2, \ldots, N \); \( k = 1, 2, \ldots, \tilde{N} \). Having carried out X-rayling the sample \( G_0 \) at \( N \) different energies \( E'_1, E'_2, \ldots, E'_N \), where all \( E'_k \) are included in the list of energies (1), one can obtain the following system of equations and conditions [4].

\[
\sum_{i=1}^{N} \mu_{xik} \rho_x w_i = \ln(h_k/H_k); \quad k = 1, \ldots, N
\]

\[
\sum_{i=1}^{N} w_i = 1; \quad w_i \geq 0; \quad i = 1, \ldots, N
\]

We denote by \( E^{(p)} = (E'_1^{(p)}, E'_2^{(p)}, \ldots, E'_N^{(p)}) \) the vector formed by the energies \( E'_1^{(p)} < E'_2^{(p)} < \cdots < E'_N^{(p)} \) included in set (1). It is clear that there is a total of \( C_N^2 = 20!/N!(20 - N)! \) of such vectors and they are lexicographically ordered. The superscript \( p \) in \( E^{(p)} \) is the ordinal of the vector in the given ordering. By \( P = P(N) \) we denote a set of positive integers \( \{1, 2, \ldots, C_N^2\} \). We rewrite (2) as \( Ax = b \) or

\[
\sum_{i=1}^{N} A_{ki} x_i = b_k; \quad k = 1, \ldots, N
\]

where \( A_{ki} = A_{ki}(E_k) = \mu_{xik}/\rho_{xi}; \ x_i = \ln(p_0 w_i); \ b_k = b_k(E_k) = \ln(h_k/H_k) \). We will consider (4) as a system of linear algebraic equations in which the matrix \( A \) and the vector \( b = (b_1, \ldots, b_N)^T \) are known, and \( x = (x_1, \ldots, x_N)^T = (\ln(p_0 w_1), \ldots, \ln(p_0 w_N))^T \) is an unknown vector (the superscript \( T \) hereinafter means transposition). The specific value of the known vector of \( l \) will be insignificant for us, therefore, to simplify the notation, we further assume \( l = 1 \) cm. The dimension of the components of the vector \( x \) is g/cm². As the previously conducted studies [4] have shown, for almost any choice of a set of energies \( E'_1, E'_2, \ldots, E'_N \) and chemical elements \( X_1, X_2, \ldots, X_N \) the matrix \( A \) is not degenerate, but poorly conditioned. In what follows, we denote the solution space of system (4) by \( \mathcal{X} \) and assume it to be Euclidean with a natural (canonical) orthonormal basis \( e_1, e_2, \ldots, e_N \), so that \( x = \sum_{k=1}^{N} x_k e_k = \sum_{k=1}^{N} \rho_0 w_k e_k \) and the norm \( ||x|| = (\sum_{k=1}^{N} (\rho_0 w_k)^2)^{1/2} \).

Further, we will compare the unknown substance \( X_0 \) with the substances known to us, included in the list of substances \( Z = \{Z_1, Z_2, \ldots, Z_M\}, \ M \geq 1 \). We will assume that \( X_0 \) and each \( Z_k \in Z \) contain only those chemical elements that are included in the list \( X_1, X_2, \ldots, X_N \). Let us explain in which case we will consider substances \( X_0 \) and \( Z_k \in Z \) to be different. Let \( \rho_T, w_T^{(k)} \), \( w_T^{(0)} \), \( w_T^{(k)} \), \( w_T^{(0)} \) be exact (taken from the reference) values of density and mass fractions of elements \( X_1, \ldots, X_N \), which are part of the substance \( Z_k \), and \( x_T^{(k)} = (x_T^{(k)}(k), x_T^{(k)}(k), \ldots, x_T^{(k)}(k)) \) is the point corresponding to \( Z_k \) in the solution space \( \mathcal{X} \). Let \( \rho_T^{(k)} = w_T^{(k)}(0), w_T^{(0)}(0), \ldots, w_T^{(0)}(0) \) be similar values for the (so far unknown to us) substance \( X_0 \), and \( x_T^{(0)} = (x_T^{(0)}(k), \ldots, x_T^{(0)}(k)) \) is the corresponding to \( X_0 \) point in \( \mathcal{X} \). We say that \( X_0 \) and \( Z_k \) are the same \( (X_0 = Z_k) \) if \( \rho_T^{(k)} = \rho_T^{(k)} \) and \( w_T^{(k)}(0) = w_T^{(k)}(0); \ i = 1, \ldots, N \) and
that $X_0$ and $Z_k$ are different ($X_0 \neq Z_k$) otherwise. It was shown in [4] that $X_0 = Z_k$ if and only if $x^{(0)}_T = x^{(k)}_T$. Now let us formulate the identification problem.

Identification problem. Let for a matter $X_0$ of unknown composition, as a result of X-raying carried out at a certain set of energies $E'_1, E'_2, \ldots, E'_N$ and measurements, the values $l, h_k, H_k; k = 1, \ldots, N$ be known. It is required to establish whether the substance $X_0$ is included in the list of substances $Z = \{Z_1, Z_2, \ldots, Z_M\}$, of known chemical composition (that is, whether the equality $X_0 = Z_i$ holds for some $i$). Let us give an explanation.

1) The list $Z$ may include, for example, some explosive, poisonous and hazardous substances prohibited for transporting. Lists of such substances, as a rule, are known in advance and are relatively small. If the unknown substance $X_0$ is not included in the list $Z$, then it can be transported, and the chemical composition of $X_0$ is not significant. If the substance $X_0$ cannot be distinguished from some substance $Z_i \in Z$, then it must be studied by other methods. This approach to the problem is quite acceptable, for example, during the customs inspection of cargo.

2) In the problem, it is not required to determine the values $\rho_0, w_1, \ldots, w_N$, i.e. to find the chemical composition of a substance $X_0$.

3) The energies at which the X-raying of the sample $G_0$ can be carried out can be chosen arbitrarily, but all of them must be included in list (1). The amount of these energies can also be arbitrary. In each measurement experiment, there is (with some measurement errors) a pair of values $h_k, H_k$ and for definiteness we assume that only one measurement experiment is carried out at each energy $E'_k$.

If there were no errors during measurements of the quantities $h_k$ and $H_k$, then after solving the (non-degenerate) system (4) using relations (3) it would be possible to easily find all unknowns $\rho_0, w_1, \ldots, w_N$ and to solve not only the identification problem, but also the problem of determining the chemical composition of the substance $X_0$ (the problem of chemistry). However, as a rule, poor conditioning of the matrix $A$ does not allow finding a satisfactory solution to the problem of chemistry, but makes it possible to solve a simpler identification problem with much greater success.

3. Estimation of measurement errors and perturbation of the solution of system (4)

Let us give the necessary explanations regarding measurement errors. Let the sample $G_0$ be X-rayed at some energy $E'_k$ from the list (1), so that $\sum_{i=1}^N A_{ki} x_i = b_k = \ln(\rho_0/h_k)$. Let $h_k = h_{Tk} + \delta h_k, H_k = H_{Tk} + \delta H_k$, where $h_{Tk}$ and $H_{Tk}$ are exact (unknown to us) values of the density of radiation fluxes entering and leaving $G_0$, and $h_k$ and $H_k$ are the same values found as a result of measurements with some unknown errors $\delta h_k$ and $\delta H_k$. Let $b_k$ and $b_{Tk}$ be the corresponding representation for the right-hand side, so that $b_{Tk} = \ln(h_{Tk}/H_{Tk}), \delta b_k = b_k - b_{Tk} = \ln(h_k/H_k) - \ln(h_{Tk}/H_{Tk})$.

Let $a_h$ and $a_H$ be the maximum relative measurement errors of the incoming and outgoing radiation that the measuring devices allow for any energy $E'_k$, then $|\delta h_k|/h_{Tk} \leq a_h, |\delta H_k|/H_{Tk} \leq a_H$. We will assume that always $0 \leq a_h + a_H < 1$. Using the results of [5] and denoting the total relative measurement error through $a_S = a_h + a_H$, we can obtain the estimate

$$|\delta b_k| \leq \frac{a_h + a_H}{1 - (a_h + a_H)} = \frac{a_S}{1 - a_S} = a.$$ 

Hence, we obtain $\|\delta b\| = [\sum_{i=1}^N (\delta b_k)^2]^{1/2} \leq \sqrt{Na}$. In what follows, we will assume that the quantity $a_S$ is much less than unity and to simplify the formulas, we assume that for any set of energies $E^{(p)} = (E_1^{(p)}, E_2^{(p)}, \ldots, E_N^{(p)})$, for which system (4) is considered, the (approximate) relation $\|\delta b\| \leq \sqrt{NaS} = \tilde{a}$ holds.

Let us estimate the perturbation of the solution to system (4) in the case when all perturbations of the right-hand side of (4) belong to the unit ball $\Omega = \{\delta b | \delta b \in \mathbb{R}^N, \|\delta b\| \leq 1\}$. Further, unless
otherwise stated, we will consider only the cases of a nondegenerate matrix $A$. Fix $N$, then the vector $E(p), p \in P(N)$ corresponds to the (nondegenerate) matrix $A = A(E(p))$ of the system (4). For $A$ and $A^{-1}$ the decompositions hold [6,7] $A(E(p)) = U(p)S(p) [V(p)]^T$, $[A(E(p))]^{-1} = V(p) [S(p)]^{-1} [U(p)]^T$, where $S(p) = \text{diag} (\sigma_1(p), ..., \sigma_N(p))$ is a diagonal matrix, and $V(p)$ and $U(p)$ are orthogonal. The numbers $\sigma_k(p)$ are the singular values of the matrix $A$, and the column vectors $v_1, ..., v_N$ of the matrix $V(p)$ are the singular vectors of the matrix $A$.

Let $S^{[N-1]} = \{ x | x \in \mathcal{X}, \| x \| = 1 \}$ be the unit sphere in $\mathcal{X}$, $\mathcal{D}(p) = [A(E(p))]^{-1} \Omega = V(p) [S(p)]^{-1} \Omega$ is an ellipsoid in $\mathcal{X}$, $\Pi = \bigcap_{p \in P} \mathcal{D}(p)$ is the intersection of all ellipsoids constructed from all possible $p \in P(N); \Phi_Z: S^{[N-1]} \rightarrow \mathbb{R}$ is the function defined by the formula

$$\Phi_Z(\omega) = \min_{p \in P} \left( \sum_{k=1}^{N} (\omega^T V(p))^2 \sigma_k(p)^2 \right)^{-1/2}$$

(6)

It was shown in [8] that the boundary $\partial \Pi$ of the set $\Pi \subset \mathcal{X}$ is defined by the equality $\partial \Pi = \{ \omega \cdot \Phi_Z(\omega) | \omega \in S^{[N-1]} \}$, that is, it coincides with the graph of the function $\omega \cdot \Phi_Z$ in the polar coordinates. Now let $x_T(0)$ and $x_T(1)$ be points in $\mathcal{X}$ corresponding to an unknown substance $X_0$ and a known substance $Z_1 \in Z$ (moreover, we do not know $x_T(0)$, $x_T(0) \neq x_T(1)$ and $\omega(1,0) = (x_T(1) - x_T(0))/\|x_T(1) - x_T(0)\|$ is the unit vector, $x_T(0,a)$ is the (found with an error) solution of system (4) based on the results of $X_0$ X-raying on some set of energies $E(q), \mathcal{D}(q) = [A(E(q))]^{-1} \Omega, \hat{a} = \sqrt{N}a$. Then, in the above notation, the following is true [8]

**Statement.**

(a) If for some $q \in P(N)$ the point $x_T(0,a) \in x_T(1) + \hat{a} \cdot \mathcal{D}(q)$, then $x_T(0) \neq x_T(1)$ and $X_0 \neq Z_1$.

(b) If for some $q \in P(N)$ the point $x_T(0) \in x_T(1) + 2\hat{a} \cdot \mathcal{D}(q)$, then the inequality $X_0 \neq Z_1$ is always established from the results of one series of X-raying at a set of energies $E(q)$.

(c) If $x_T(0) \in \mathcal{X} \setminus (x_T(1) + 2\hat{a} \cdot \Pi)$, then the inequality $X_0 \neq Z_1$ is always established from the results of one series of X-raying at some set of energies $E(p)$.

Taking into account that $\hat{a} = \sqrt{N}a$ from item (c) and formula (6), we can obtain a simple (approximate) sufficient condition

$$a_s < \frac{\| x_T(1) - x_T(0) \|}{2\sqrt{N} \Phi_Z(\omega(1,0))} = a_s$$

(7)

under which the inequality $X_0 \neq Z_1$ is always established from the results of one series of X-raying on the set of energies $E(p_1)$, where $p_1 = p_1(\omega(1,0)) \in P(N)$ is such a number at which the minimum is reached in (6). For example, when $N = 2$ for elements hydrogen and carbon for a pair of substances $X_0$ - divinyl (CH$_2$)$_2$(CH)$_2$ and $Z_1$ - isooctan (CH$_3$)$_3$CCCH(1)2 using the data from [3] and formula (6) from (7) one can get

$$E(p_1) = (0.1 \text{ Mev}, 0.15 \text{ Mev}), \sqrt{N} \approx 1.4142, \| x_T(1) - x_T(0) \| \approx 0.05458, \Phi_Z(\omega(1,0)) \approx 2.3159, a_s < 8.33 \cdot 10^{-5}.$$ 

It is clear from what has been said that the function $\Phi_Z$ turns out to be very useful in the analysis of the fact that how well (or badly) two specific substances can be distinguished, therefore, it can be called an indicator to the distinguishability of substances [8].
Before proceeding to the description of the calculations performed, we will give one more explanation. We will assume that for any \( N \) and \( p \in P(N) \) the singular values of the matrix \( A = A(N, E^{(p)}) \) are always numbered so that \( \sigma_1^{(p)} \leq \sigma_2^{(p)} \leq \cdots \leq \sigma_N^{(p)} \), then
\[
[\sigma_1^{(p)}]^{-1} \geq [\sigma_2^{(p)}]^{-1} \geq \cdots \geq [\sigma_N^{(p)}]^{-1}.
\]
(8)
It is clear that for every \( N \) and \( p \in P(N) \) the numbers \([\sigma_1^{(p)}]^{-1}, \ldots, [\sigma_N^{(p)}]^{-1}\) are the lengths of the semi-axes of the perturbation ellipsoids of the solution of system [4] in the case when the perturbation norm of the right-hand side is \( \| \delta b \| = 1 \). Let us fix \( N \), and let \( p_l \in P(N) \) be a number such that
\[
[\sigma_1^{(p_l)}]^{-1} = \min_{p \in P(N)} \{ [\sigma_1^{(p)}]^{-1} \},
\]
and \( p_T \in P(N) \) is such a number that
\[
[\sigma_N^{(p_T)}]^{-1} = \min_{p \in P(N)} \{ [\sigma_N^{(p)}]^{-1} \}.
\]
For the number \( p = p_l \), the set of numbers (8) is denoted by \( l_{l1} \geq l_{l2} \geq \cdots \geq l_{LN} \), and for the number \( p = p_T \), a similar set of numbers (8) is denoted by \( l_{T1} \geq l_{T2} \geq \cdots \geq l_{TN} \). In [8], it was shown that the number \( l_{t1} \) is used to determine the best ensured accuracy of solving a chemistry problem. When we are dealing with an identification problem, the best result largely depends on the number \( l_{TN} \).

4. Some results of the performed calculations

When performing the calculations, two main objectives were set. The first one consisted in elucidating the behavior of the sequence of values of the numbers \( l_{TN} \) and \( \sqrt{N} \cdot l_{TN} \), \( N = 2,3, \ldots \), since the minimal possible error in solving system (4) depends on them. The second objective was to calculate the values of the quantities \( a_c^* \) (see formula (7)) for some specific groups of chemical elements and pairs of substances consisting of these elements. When performing the first part of the work, a computer program was written, which, based on the data from tables [3], for some groups of chemical elements formed all kinds of matrices \( A = A(N, E^{(p)}) \), \( N = 2,3, \ldots, 10 \), \( p \in P(N) \), found singular numbers and singular vectors for them and then performed all the necessary calculations. At \( N = 2 \), the group of chemical elements consisted of hydrogen and carbon. Further, with an increase in \( N \), new elements were added to them, and at \( N = 10 \) a group of elements \{H,C,N,O,F,Na,Mg,P,S,Cl\} was obtained. Some of the results obtained in this case are presented in table 1.

| \( N \)   | \( N=2 \) | \( N=3 \) | \( N=4 \) | \( N=5 \) | \( N=6 \) | \( N=7 \) | \( N=8 \) | \( N=9 \) | \( N=10 \) |
|----------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| \( l_{TN} \) | 2.247   | 1.741   | 1.465   | 1.277   | 1.149   | 1.037   | 0.9558  | 0.8997  | 0.8504  |
| \( \sqrt{N}l_{TN} \) | 3.178 | 3.015 | 2.931 | 2.856 | 2.815 | 2.745 | 2.7035 | 2.6990 | 2.6892 |

The first row of table 1 indicates the dimension of the matrix \( A \), with each value of \( N \) corresponding to a certain set of chemical elements. The second and the third rows of the table indicate the values \( l_{TN} \) and \( \sqrt{N}l_{TN} \) corresponding to the given \( N \). As \( N \) grows, the numbers \( l_{TN} \) decrease. The numbers \( \sqrt{N}l_{TN} \) in the third line also decrease. It is these numbers that are most important for us, since they are included in the denominator of formula (7) at \( \omega^{(10)} = V_N^{(p_T)} \). A decrease in \( \sqrt{N}l_{TN} \) with increasing \( N \) indicates that with an increase in the dimension of the identification problem, the requirements for the accuracy of radiation measurements will probably not increase, and the success of solving the identification problem for a pair of substances \( X_0 \) and \( Z_1 \) will depend mainly on the distance \( \| x_T^{(1)} - x_T^{(0)} \| \) between the points \( x_T^{(1)} \) and \( x_T^{(0)} \) and the relative position of the unit vectors \( \omega^{(10)} \) and \( V_N^{(p_T)} \) (see formula (6)).

The second objective of the calculations was to find the values of \( a_c^* \) (see inequality (7)) for some sets
of specific substances for different $N$. Complete lists of substances are not indicated here due to their large size.

At $N = 2$, a group of 40 hydrocarbons was selected. Figure 1 shows a dot diagram corresponding to the exact position of these 40 substances in the solution space $\mathcal{X}$. If $\rho_T$, $w_{T1}$, $w_{T2}$ are the exact values of density, mass fraction of carbon and hydrogen in (some) substance, then $x_T = (x_{T1}, x_{T2}) = (l\rho_T w_{T1}, l\rho_T w_{T2})$ is the corresponding point in $\mathcal{X}$. Recall that $l = 1$ cm, therefore the dimension of the components of the vector $x_T$ is g/cm$^2$.

![Figure 1. A dot diagram for 40 hydrocarbons in the solution space $\mathcal{X}$, $N = 2$.](image)

A closed curve similar to a strongly oblate ellipse is the boundary of the set $(1, 0.1) + 2\tilde{a} \cdot \Pi$ for the case when the amplitude of the relative error of a single measurement $a_S = 10^{-3}$ (see formula (5)). The number 559 marks the point corresponding to the substance isobutane $(\text{CH}_3)_2\text{CHCH}_3$, and the number 424 – anthracene $\text{C}_{14}\text{H}_{10}$. Point $(1, 0.1)$ does not correspond to any substance and is chosen only for a convenient location of the central point of the set $2\tilde{a} \cdot \Pi$. Figure 2 shows a fragment of figure 1, enlarged approximately 10 times.

![Figure 2. A dot diagram for 23 hydrocarbons in the solution space $\mathcal{X}$, $N = 2$.](image)

Dotted lines mark sections of straight lines passing through the origin of the coordinate system; some substances are marked with numbers. Two fragments of solid lines form part of the boundary of the set $x_T^{(579)} + 2\tilde{a} \cdot \Pi$. The point $x_T^{(579)}$ corresponds to the substance $\text{C}_8\text{H}_{16}$ (1-octene). Substances with numbers 563, 564, 565 are isomers $\text{C}_6\text{H}_{14}$. Another group of isomers $\text{C}_6\text{H}_{10}$ is formed by substances with numbers 583, 584, 585. Substances 576 ($\text{C}_6\text{H}_{14}$), 577 ($\text{C}_6\text{H}_{12}$), 578 ($\text{C}_7\text{H}_{14}$), 579 ($\text{C}_8\text{H}_{16}$) are not isomers, but their ratios of mass fractions $w_{T1}/w_{T2}$ are the same and, like the isomers, they lie on one straight line passing through the origin.

According to statement (c) if the point $x_T^{(n)} \notin x_T^{(579)} + 2\tilde{a} \cdot \Pi$, then to establish the difference between the (unknown) substance $x_T^{(n)}$ from the (known) substance $x_T^{(579)}$ one series of measurements...
is always sufficient. In this case we will say that the substances \( x_T^{(n)} \) and \( x_T^{(579)} \) are well distinguishable (for a given \( \alpha_s \)). Otherwise, if the point \( x_T^{(n)} \) $\in$ \( x_T^{(579)} + 2 \hat{a} \cdot \Pi \), we say that the substances \( x_T^{(n)} \) and \( x_T^{(579)} \) are poorly distinguishable. For example, in this case, for \( \alpha_s = 10^{-3} \), the substances \( x_T^{(578)} \) and \( x_T^{(579)} \) are well distinguishable, but \( x_T^{(585)} \) and \( x_T^{(579)} \) are poorly distinguishable. Thus, to solve the question of the possible distinguishability of two specific substances (with known \( x_T^{(n)} \) and \( x_T^{(m)} \) ) for a given \( \alpha_s \), it is necessary to construct a set \( 2 \hat{a} \cdot \Pi \), place its center at the point \( x_T^{(m)} \) and check the validity of the inclusion \( x_T^{(n)} \) $\not\in$ \( x_T^{(m)} + 2 \hat{a} \cdot \Pi \). A computer program was written that automatizes this procedure and a number of calculations were performed for various values of \( N \) and \( \alpha_s^* \) indicated in the first row of table 2.

**Table 2.** Number and percentage of well distinguishable pairs for some values of \( \alpha_s^* \) and \( N \).

| \( \alpha_s^* \) | \( 3 \cdot 10^{-2} \) | \( 10^{-2} \) | \( 5 \cdot 10^{-3} \) | \( 10^{-3} \) | \( 5 \cdot 10^{-4} \) |
|------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| \( N = 2 \) | 40 sub. | 33 | 316 | 533 | 722 | 754 |
| | 780 pairs | 4.23% | 40.5% | 68.3% | 92.6% | 96.7% |
| \( N = 3 \) | 23 sub. | 106 | 202 | 224 | 249 | 250 |
| | 253 pairs | 41.9% | 79.8% | 88.5% | 98.4% | 98.8% |
| \( N = 4 \) | 23 sub. | 22 | 126 | 196 | 233 | 243 |
| | 253 pairs | 8.70% | 49.8% | 77.5% | 92.1% | 96.0% |

The table shows the following. For \( N = 2 \), 40 substances were used for calculations; they form 780 different pairs \{ \( x_T^{(n)} \), \( x_T^{(m)} \) \}, \( n < m \). With the maximum relative error \( \alpha_s^* = 3 \cdot 10^{-2} \), 33 pairs are well distinguishable, they make up 4.23% of all pairs. With \( \alpha_s^* = 10^{-2} \), 316 pairs are clearly distinguishable, they already account for 40.5% of the number of all pairs. Further, with a decrease in \( \alpha_s^* \), the proportion of well-distinguishable pairs increases, however, even with \( \alpha_s^* = 5 \cdot 10^{-4} \), a certain number of poorly distinguishable pairs remains. They include, for example, a pair of isomers \( C_6 H_{10} \), which is formed by substances with numbers 565 and 563. To achieve a good distinguishability of the substances of this pair, the (limiting) value \( \alpha_s^* = 6.628 \cdot 10^{-5} \) is required. Both of these substances are marked in figure 2. For \( N = 3 \) and \( N = 4 \), we used 23 substances for the calculations; they form 253 pairs. Similarly, to the case \( N = 2 \) the columns of table 2 for these cases indicate the number of well distinguishable pairs for different values of \( \alpha_s^* \). At \( N = 5 \), only one pair of substances was found in the database compiled by the author – these are the poisonous substances sarin \( (CH_3)_3O_2PF \) and soman \( (CH_3)_5C_2O_2PFH \). This pair becomes well distinguishable when \( \alpha_s^* < 5.87 \cdot 10^{-3} \).

In conclusion, we note that the obtained numerical results refer only to the set of substances consisting of the groups of chemical elements considered here. First of all, these are numerous organic compounds. In practice, situations are often possible when a group of chemical elements should be different and calculations for it may show otherwise. These results can also noticeably differ from those obtained here when changing the energy interval at which the problem is considered. Preliminary studies carried out for the same groups of chemical elements indicate that the expansion of the energy range considered here [0.1 – 20] MeV to [0.001 – 20] MeV should significantly improve the results obtained in this study.

**Acknowledgments**

The work was performed as part of the state assignment № 075-01095-20-00 and funded by RFBR, project number 20-01-00173.
References

[1] Jumanazarov D, Koo Ja, Busi M, Poulsen H F, Olsen U L and Iovea M 2020 System-independent material classification through X-ray attenuation decomposition from spectral X-ray CT *NDT and E International* **116** 102336

[2] Osipov S P, Udod V A and Wang Ya 2017 Identification of Materials in X-Ray Inspections of Objects by the Dual-Energy Method *Russ. J. of Nondestructive Testing* **53**(8) pp 568-587

[3] Berger M J, Hubbell J H, Seltzer S M, Chang J, Coursey J S, Sukumar R and Zucker D S 2005 *XCOM: Photon Cross Section Database*, National Institute of Standards and Technology, Gaithersburg, URL: http://www.physics.nist.gov/xcom

[4] Nazarov V G 2020 Estimation of the Calculation Accuracy in the Problem of Partial Identification of a Substance *J. of Applied and Industrial Mathematics* **14**(3) pp 555-565. DOI: 10.1134/S199047892003014X Available as http://link.springer.com/article/10.1134/S199047892003014X

[5] Nazarov V G 2020 Some estimations of errors at problem of continuous medium radiography *Dalnevostochnyj matematicheskij zhurnal* **20**(1) pp 82-89 (in Russian)

[6] Godunov S K, Antonov A G, Kiriluyk O P and Kostin V I 1988 Guaranteed accuracy of solving systems of linear equations in Euclidean spaces Novosibirsk Nauka (in Russian)

[7] Forsythe G E, Malcolm M A and Moler C B 1977 *Computer methods for mathematical computations* Prentice-Hall Inc. Englewood Cliffs N.J. 07632

[8] Nazarov V G 2019 Problem of partial identification of unknown medium *Dalnevostochnyj matematicheskij zhurnal* **19**(1) pp 43-62 (in Russian)