Universal interpolation method for many-dimensional spaces

V. Pohánka
Geophysical Institute of the Slovak Academy of Sciences

January 18, 2005

Abstract

The requirements on the universal interpolation method applicable for the space of any dimension and geometry are formulated; the most important interpolation methods are shown to violate at least some of these conditions. The proposed universal interpolation method is developed and shown to conform the posed requirements. Its main advantage, besides its universality, is that it is numerically safe and efficient. The simplest possible form of the interpolation formula is presented, but the method allows to construct generalized formulae adapted to the particular additional requirements.

Keywords: approximation, smoothness, simplicity, polynomial, minimalization, weighting function.

1 Introduction

In several scientific disciplines we are often facing the problem that some quantity, whose values depend on the position in some space, is known (measured) only at a discrete set of points of this space, but we would like to know the (most probable) value of this quantity at any point within some domain of this space. There are several methods for solving this problem and then we speak about the interpolation, approximation and extrapolation.

For the spaces of dimension 1 the known interpolation methods are quite satisfiable, but for spaces of dimension 2 or higher there arise several difficulties and unpleasant features and there seems to be no widely accepted interpolation method.

Therefore it would be very desirable to have some universal method for both interpolation and extrapolation that would be sufficiently flexible to master a wide range of possible particular requirements on the interpolated values.

2 Requirements on the universal interpolation method

Let us describe the basic requirements on such a method of interpolation (and/or approximation) and extrapolation of the given data that could be used in wide range of situations and spaces.

For convenience we first define for any natural number k the set $I(k)$ as the set of all natural numbers i such that $1 \leq i \leq k$, and, for any natural number i the set $I(k|i)$ as the set of all natural numbers j such that $1 \leq j \leq k$, $j \neq i$.

Let us have some metric space S of dimension n with positively definite metrics; in this space, the distance $d(X, Y)$ of points X and Y can be defined naturally as the minimum of the lengths of geodesic lines connecting these points.

In the space S, let there be some points X_i ($i \in I(N)$) at which we know the values f_i of some quantity which we aim to express by a function $f(X)$ defined at every point of the space S. We shall denote the points X_i as the measurement points and the values f_i as the measured values;

the positions of measurement points and the measured values will be called together the measured data. The function $f(X)$ will be called the interpolating function; we shall denote the point X as the calculation point and the value of the interpolating function at this point as the interpolated value. In some cases it is suitable to define the measurement domain as the smallest convex domain containing all measurement points (it need not to exist for an arbitrary position of measurement points).

The interpolating function $f(X)$ should satisfy the following conditions:

C1: The interpolating function should represent the measured data as exactly as possible.

This can be formulated precisely as follows: the nonnegative quantity E defined as

$$E^2 = \frac{1}{N} \sum\nolimits_{i \in I(N)} (f(X_i) - f_i)^2 \tag{1}$$

should satisfy the inequality

$$E \leq E_0, \tag{2}$$

where E_0 ($E_0 \geq 0$) is some given number (it will be called the approximation limit). This requirement includes the case $E_0 = 0$ when we get the conditions

$$i \in I(N): \quad f(X_i) = f_i. \tag{3}$$

Usually it is spoken about interpolation only in the case of the condition (3). As we prefer the condition (2), we adopt the following terminology: if $E_0 > 0$, we shall speak about the approximate interpolation, and, if $E_0 = 0$, about the exact interpolation. The term approximation will be used only in the case that we represent the measured data by a function from some class and the actual function from this class is obtained by requiring that the quantity E should be as small as possible (without defining any range for the value of E).

Note that our formulation of the condition does not mean that we aim to perform some smoothing of measured values (although the method should also be able to do that) nor that the measured values are considered in a statistical sense (what could be expected in the case we would have several values of the measured quantity at each measurement point). Our formulation simply means that (as we shall see later) the strict requirement (3) could contradict the other conditions on the interpolating function which we shall regard as more important. Our primary goal is not to find an analytic expression of a function $f(X)$ satisfying conditions (3) exactly, but to develop a method suitable for the case that the values f_i are obtained by measurement and thus that they can be expressed as integer multiples of some basic unit. In this case it should be allowed that the values $f(X_i)$ differ from the measured values f_i by an amount smaller than this basic unit.

C2: The interpolating function should be as smooth as possible.

This requirement is perhaps the most fundamental for the interpolation (and approximation): the reason is that the smoother is a function, the less information is needed for its definition. As the only available information for the function $f(X)$ is represented by the measured data (thus the position of the points X_i and the values f_i), the course of this function has to be determined by the requirement that it should be maximally smooth.

This condition can be formulated more exactly as the requirement that the function $f(X)$ should belong to the class C_I (thus it should have in the whole space S continuous partial derivatives of order I) for some $I \geq 0$ and the value of I should be as large as possible. If we define the domain S_N as the space S with exception of all points X_i, $i \in I(N)$, we further require that the function $f(X)$ should belong in this domain to the class C_∞ (thus it should have in the domain S_N continuous partial derivatives of any order).

Note that this condition does not mean that some partial derivatives of the interpolating function would be actually calculated; it only assures that the approximation procedure is maximally simple what would not be the case if the interpolating function would have a singularity at any

points other than the measurement points.

C3: The interpolating function should behave uniformly in the whole space.

This means that the behaviour of the function $f(X)$ in the neighbourhood of any of the points X_i should not be different from its behaviour in the neighbourhood of any other point of the space S. This represents the strenghtening of the condition C2 in the sense that the function $f(X)$ should belong in the whole space S to the class C_∞. Moreover, this function should have no other special anomalous properties as, for example, that all partial derivatives of the first order have to be equal to zero at every measurement point (there will be later given examples of such interpolating functions).

C4: The interpolating function should have a reasonable behaviour far from the measurement points.

This condition enables to use the formula for the interpolating function also for extrapolation of the measured data. Exact formulation of this condition depends on our requirements; for example, we can require that the function $f(X)$ should be bounded everywhere in S. If the space S contains infinitely distant points (thus for any value of distance there are two points whose distance is greater than this value), then we can require that the interpolating function should have the desired asymptotic behaviour at infinity.

C5: The interpolating function should have a reasonable behaviour by any change of (mutual) position of the measurement points.

This means that the function $f(X)$ (which is also a function of all X_i) should be at least continuous with respect to the position of all points X_i. Particularly, the formula for this function should give some value also in the case that any two (or more of the) measurement points coincide.

Once again, this does not mean that the measurement points move in some way, but that the interpolation formula should have no anomalous behaviour if any two measurement points are much closer to each other than the other ones. Then the interpolation formula can be used also for highly irregular nets of measurement points.

C6: The interpolating function should behave uniformly with respect to the number of measurement points.

This condition assures that the calculation of the interpolated value does not change substantially if some new measurement points are added.

C7: The values of the interpolating function should be independent on the chosen coordinate system.

This is a clear requirement: the interpolated value should depend only on the position of the calculation point and the measurement points, and, in the case of nonhomogeneous geometry of the space S, also on the local geometry.

C8: The interpolating function should behave uniformly with respect to the global geometry of the space.

This can be formulated more exactly as that the interpolation method should use for the calculation of the interpolated value only such quantities that can be defined for any geometry of the underlying space. If this condition is satisfied, the interpolation method can be easily adapted from one space to another one.

C9: The interpolating function should behave uniformly with respect to the number of dimensions of the space.

This is a strenghtening of the previous condition in the sense that the interpolation method should not be substantially different for any dimension of the underlying space.

C10: The interpolating function should be as simple as possible.

This requirement means that the analytical formula for the interpolating function should be maximally simple, what reduces the ambiguity of choice of this formula and also enables to perform

the numerical calculation efficiently.

C11: The interpolating function should be numerically well defined in the whole space.

This is a substantial requirement and it means that for any position of the calculation point and any positions of the measurement points, the value of the interpolating function can be numerically calculated without any chance of breaking of the numerical calculation (for example, if we obtain a matrix whose determinant is zero although its exact value is nonzero). In particular, the numerical calculation should be possible without problems for the case that the calculation point is identical with some measurement point or that two (or more) measurement points coincide.

C12: The interpolating function should be as simple numerically as possible.

This is a strenghtening of the condition C10 and it means that the number of numerical operations needed for the calculation of the interpolated value should not increase with the number of dimensions of the space and with the number of calculation points in a way that would make the numerical calculation impossible.

The most important in this sense is the dependence on the number of the calculation points N. Therefore, we divide the functions containing the variables X_i according to their complexity with respect to the parameter N into three classes:

F_1: functions containing N only in the summation of N terms such that the i-th term ($i \in I(N)$) contains the variable X_i.

F_2: functions containing N in the multiplication of N or $N-1$ terms such that the i-th term ($i \in I(N)$ or $i \in I(N|j)$ for some j) contains the variable X_i.

F_3: functions containing inverse matrix of some $(N \times N)$-matrix such that the (i,j)-th matrix element ($i \in I(N)$, $j \in I(N)$) contains variables X_i or X_j or both.

Then we can require that the interpolating function is constructed in such a way that it contains only functions of the first class (if it is possible). In this case the number of operations needed for the calculation of the interpolated value would increase as a linear function of N.

Now we formulate the main principle for the determination of our interpolation method:

C0: The interpolated value should depend on the measured values at all measurement points. The dependence of the interpolated value on the measured value at some measurement point should decrease with increasing distance between the calculation point and this measurement point.

Of course, we have not defined exactly what is the dependence of the interpolated value on the measured value, but the sense of this principle is clear: the interpolated value should depend first of all on the measured values at the nearby measurement points. The actual dependence on the distance may be formulated differently according to our particular requirements and the nature of the interpolated quantity; in this way the interpolation method should allow a wide range of possibilities.

3 Formulation of the universal interpolation method

Before we start to develop our universal interpolation method, we first consider the main existing interpolation methods with respect to their properties compared with our above requirements.

We can divide the interpolation methods into three broad categories:

M1: The methods that define the interpolating function by a single formula for the whole space (or at least for the whole measurement domain).

These methods mostly construct the interpolating function from some number of simple(r) functions of one or two variables (these variables are the position of points of the space S). According to the way the interpolating function is constructed from these functions, this category can be further divided into several subcategories:

M1a: The interpolating function is a polynomial of functions representing the coordinates.

There are chosen certain functions $x_p(X)$ ($p \in I(n)$) such that any sufficiently smooth function of X can be written as a series of monomials of these functions (in the case of euclidean space these functions can be the euclidean coordinates of the point X). The interpolating function is written as a polynomial of these functions of some order; the coefficients of the polynomial are determined according to the conditions (3).

The simplest example of such an interpolation method is the well known Lagrange interpolation formula in the case of the one-dimensional space. Although this formula is the base for formulation of several other interpolation methods, there arise problems by applying this method for spaces of dimension $n > 1$. It can be easily shown that in an n-dimansional euclidean space, there are $(n+m-1)!/(n-1)!m!$ different monomials of coordinates of order m, and therefore, there are $(n+m)!/n!m!$ different monomials of coordinates of order at most m. For $n = 1$ we have $m+1$ different monomials of order at most m, and therefore, for any number N, if there are N different measurement points, we can construct for them an interpolating polynomial of order m putting $m = N-1$. On the contrary, for any dimension $n > 1$, the interpolating polynomial for N different measurement points can exist only if there is such m that $N = (n+m)!/n!m!$. Even this condition is not sufficient for an arbitrary position of the measurement points: if there exists a polynomial of the order at most m acquiring the zero value at each measurement point, the interpolating polynomial of the order m either does not exist or cannot be chosen uniquely.

In principle it is possible to express the measured data by a polynomial of the smallest order m such that $N \leq (n+m)!/n!m!$, but then the redundant coefficients have to be determined by some additional conditions. Further, the Lagrange formula can be generalized for the case $n > 1$ for certain quasi-regular nets of measurement points (for example, if the n-dimensional net is a cartesian product of n one-dimensional nets); however, this is a very special case.

We see that any interpolation method of this subcategory violates several conditions from the previous section (at least the conditions C5, C6, C11, C12); on the other hand, the conditions C2 and C3 are satified well. Such a method is suitable for extrapolation (condition C4) only for small distances from the measurement domain; at infinity it cannot give bounded results.

M1b: The interpolating function is a linear combination of measured values such that the coefficient of each measured value contains a product of components (which are functions of two position variables) assuring that this coefficient is zero at all other measurement points.

More exactly, we write the interpolating function in the form

$$f(X) = \sum\nolimits_{i \in I(N)} f_i \, w_i(X), \tag{4}$$

where the coefficients $w_i(X)$ (weighting functions) satisfy the following conditions:

$$i \in I(N),\ j \in I(N|i): \quad w_i(X_j) = 0, \tag{5}$$

$$i \in I(N): \quad w_i(X_i) = 1. \tag{6}$$

In analogy with the Lagrange interpolation formula, we choose some function $q_0(X,Y)$ such that $q_0(X,X) = 0$ and $q_0(X,Y)$ is nonzero for X in some neighbourhood of Y, and, for each i we write the function $w_i(X)$ in the form containing the factor $z_i(X)$ defined as

$$i \in I(N): \quad z_i(X) = \prod\nolimits_{j \in I(N|i)} q_0(X, X_j). \tag{7}$$

In this way we satisfy the conditions (5); in order to satisfy also the conditions (6), we have several possibilities of choice of exact form of functions $w_i(X)$. We mention here the two most important:

M1ba: The weighting functions have the form

$$i \in I(N): \quad w_i(X) = \frac{z_i(X)}{z_i(X_i)}. \tag{8}$$

This is an exact analogy of the Lagrange interpolation formula: in the case of one-dimensional euclidean space we obtain this formula by choosing $q_0(X,Y) = x - y$, where x, y are the coordinates of points X, Y. However, for $n > 1$ the function $q_0(X,Y)$ cannot be a polynomial of coordinates of the first order; the simplest possibility is $q_0(X,Y) = d(X,Y)^2$ (recall that $d(X,Y)$ is the distance of points X, Y). If we want to avoid the excessive growing of the interpolated value far from the measurement domain, we can choose this function like $q_0(X,Y) = d(X,Y)^2/(d_0^2 + d(X,Y)^2)$, where d_0 is some suitable constant.

The interpolating function defined by formulae (4), (7) and (8) satisfies well the conditions C2 and C3, but it violates the conditions C5, C11 and C12 (the weights are functions of the class F_2) and it can hardly satisfy the condition C4 (even if the function $q_0(X,Y)$ is chosen to be bounded). Further, it can be shown that the weighting functions (8) (on the contrary to the case of the one-dimensional polynomial Lagrange interpolation formula) do not satisfy the condition

$$\sum\nolimits_{i \in I(N)} w_i(X) = 1. \tag{9}$$

Therefore, in the case that all measured values are equal, the interpolating function is not a constant, what can be considered as an anomaly.

M1bb: The weighting functions have the form

$$i \in I(N): \quad w_i(X) = \frac{z_i(X)}{z(X)}, \tag{10}$$

where

$$z(X) = \sum\nolimits_{i \in I(N)} z_i(X). \tag{11}$$

These functions satisfy the condition (9) and the interpolating function defined by formulae (4), (7), (10) and (11) can be written in a very simple form

$$f(X) = \frac{\sum_{i \in I(N)} f_i \, q_1(X, X_i)}{\sum_{i \in I(N)} q_1(X, X_i)}, \tag{12}$$

where

$$q_1(X,Y) = \frac{1}{q_0(X,Y)}. \tag{13}$$

Here again the simplest choice is $q_0(X,Y) = d(X,Y)^2$; if we choose this fuction as $q_0(X,Y) = d(X,Y)^{2\alpha}$ (where α is some positive constant), the interpolating function will have a singularity at each measurement point if α is not a positive integer. However, even if α is a positive integer, the interpolating function will violate the condition C3, as its first order partial derivatives with respect to the coordinates of X will be equal to zero at each measurement point. On the other hand, such an interpolating function is very simple (it contains only functions of the class F_1 and thus it satisfies the condition C12) and it satisfies well the conditions C2, C5, C6, C8, C9 and C10. Further, as the condition (9) is satisfied, the interpolated values are bounded everywhere (condition C4 is satisfied), and, for equal measured values the interpolating function is constant. However, there is another drawback that the interpolating function can acquire its maximum and minimum only at some measurement point; this is a very unnatural phenomenon. Therefore the interpolating functions of this kind are not suitable for the purpose they were constructed, contrary to the meaning of the author in his earlier work [2].

M1c: The interpolating function is constructed as a sum of elementary contributions of all measurement points with weights that are determined using the conditions (3); each contribution is some chosen function of two variables: the positions of the calculation point and of the measurement point.

More exactly, there is chosen some function $q_2(X,Y)$ (which is usually nonzero if X is in some neighbourhood of Y), and the interpolating function is written in the form

$$f(X) = p(X) + \sum\nolimits_{i\in I(N)} w_i\, q_2(X, X_i), \tag{14}$$

where $p(X)$ is some simple function (usually a constant or a linear combination of monomials of the coordinates of the point X of some low order); this term is necessary as in the most cases the sum on the rhs of (14) is not able to reproduce such simple function. The coefficients w_i are determined by the equations obtained by inserting the expression (14) in the conditions (3):

$$i \in I(N): \quad p(X_i) + \sum\nolimits_{j\in I(N)} w_j\, q_2(X_i, X_j) = f_i; \tag{15}$$

the unknown coefficients appearing in the function $p(X)$ are determined from certain additional conditions. This method requires that the function $q_2(X,Y)$ has the property that for any positions of mutually different measurement points, the matrix of the resulting system of equations has nonzero determinant.

This category contains several interpolation methods (see, for example, [1]). Perhaps the most important and most widely used method is the interpolation using the thin-plate splines; in this case the function $f(X)$ is defined as a minimum of certain functional subject to the conditions (3), what leads to the equation

$$\Delta_X^{m+1} f(X) = \sum\nolimits_{i\in I(N)} w_i\, \delta(X, X_i),$$

where Δ_X is the n-dimensional Laplacian in the space S (acting on the variable X), $\delta(X,Y)$ is the δ-function in the same space, and m ($m \geq 0$) is the order of the spline (the usual thin-plate spline has the order $m = 1$). The function $q_2(X,Y)$ is then chosen as some simple solution of the equation

$$\Delta_X^{m+1} q_2(X,Y) = \delta(X,Y)$$

and the function $f(X)$ can be written in the form

$$f(X) = h(X) + \sum\nolimits_{i\in I(N)} w_i\, q_2(X, X_i),$$

where $h(X)$ is an arbitrary (m+1)-harmonic function (everywhere in the space S). For $n > 1$ there is an infinite-dimensional base of (m+1)-harmonic functions (in the case of euclidean space this base functions can be chosen as polynomials of coordinates) and thus the function $h(X)$ contains infinitely many free parameters. It would be natural to determine these additional parameters by fitting the general trend of data, but the usual practice is to keep only a polynomial of coordinates of some low order and to determine the unknown coefficients by imposing certain conditions on the asymptotic behaviour of the sum on the rhs.

Contrary to the widespread opinion (expressed also in [1]), the usual ($m = 1$) thin-plate spline (in the 2-dimensional case) cannot be considered as the correct continuous minimum curvature surface; the reason is that the smoothness of the function $f(X)$ is not assured by the form of the minimalization functional alone, but it depends also on the form of the additional constraints. For example, if these constraints would be some boundary conditions on the boundary of the measurement domain, the function $f(X)$ would be maximally smooth within this domain. However, in the case of thin-plate splines the δ-function terms cause that the function $f(X)$ is (m+1)-harmonic (and thus belonging to the class C_∞) only in the domain S_N; at every point X_i (or at almost every such point, if it happens that some of the coefficients w_i are zero), this function has a singularity.

It is instructive to compare the behaviour of usual thin-plate splines in spaces of different dimension: the function $f(X)$ is continuous only for $n \leq 3$; if $n \geq 4$ the function $q_2(X,Y)$ is infinite for $X = Y$ and therefore the conditions (15) can be satisfied only if all coefficients w_i are

zero (thus in this case the function $f(X)$ is reduced to the function $h(X)$ and the spline approach breaks down). Further, the function $\Delta_X f(X)$ is continuous only for $n = 1$: if $n = 2$ and $n = 3$, this function is infinite at (almost) each point X_i. Moreover, if $n = 3$, even the first order partial derivatives of the function $f(X)$ do not exist at (almost) each point X_i. This behaviour is in striking contrast to the requirement C3.

It has to be stressed that it is absolutely necessary that any function which is intended to be used for interpolation should possess everywhere at least all second order partial derivatives. This has the following ground: speaking freely, we can consider the interpolation as some kind of extending the unknown function from the points X_i (at which it is known) to all other points. This extending can be hardly correct, if the partial derivatives of the function we want to construct have incorrect values at the points X_i. In this respect it is first of all necessary that the interpolating function has acceptable values of all first and second order partial derivatives, as then it is possible to approximate this function in the neighbourhood of any point X_i as a second order polynomial of the (relative) coordinates (only the polynomial of at least second order can describe local extrema of the approximated function). With respect to the thin-plate splines this minimal requirement means that we have to use the spline of order $m \geq [n/2] + 1$.

However, even then the thin-plate spline cannot be considered as an acceptable interpolating function, as it gives the measurement points an exclusive status (what can be justified in the case of an n-dimensional elastic object fixed at certain points). On the contrary, in the case of interpolation the measuring of some quantity at some points does not mean that this quantity is somehow influenced by the measurement: the actual course of this quantity in the space S is absolutely independent on the positions of the measurement points and this fact has to be reflected on the form of the interpolating function.

Several other interpolation methods leading to the expression (14) and conditions (15) have smooth function $q_2(X, Y)$ (thus the condition C3 is satisfied), but they all have problems in satisfying the conditions C5 and C11. Further, if the function $q_2(X, Y)$ tends to zero with increasing distance $d(X, Y)$, then it can hardly be universal: if we choose, for example, $q_2(X, Y) = d_0^2/(d_0^2 + d(X, Y)^2)$, the constant d_0 has to be chosen in each case separately according to the mean distance of the measurement points.

All interpolating methods of the category M1c violate the condition C12, as the corresponding interpolating functions belong to the class F_3, and thus the numerical calculation of the parameters of the interpolating function can be difficult if the number of measurement points is too great. In [1] there are presented techniques which allow to avoid this difficulty (the large system of equations is solved approximately in an iterative way), but it is not clear whether this approach does not improve the speed of the numerical calculation at the cost of the accuracy of the solution.

However, the most important drawback of the methods of the category M1c is that they violate the principle C0: the interpolated value depends on all measured values in a way which hides the dependence of the former value on the distance between the calculation point and the particular measurement point. This has the following consequence: as the interpolating function $f(X)$ given by the formula (14) is once for all tied to the actual position of the measurement points X_i and actual measured values f_i, the whole calculation of this function has to be performed again if only a single from these parameters would be changed (for example, if some of them prove to be erroneous or we obtain a more accurate value).

Even more important is the following fact: as the measured data represent usually only a part of the data from some larger data set, we should use for the calculation of the interpolating function this larger set (which would yield a different interpolated values for a large number of the calculation points). Speaking strictly, in the case of a quantity which has a global character, we have to interpolate using the global set of data; but the huge number of data points represents for all methods from the category M1c an insurmountable obstacle.

The second main category of interpolation methods is the following:

M2: The methods that divide the space (or only the measurement domain) into several smaller subdomains, and define the interpolating function by a separate formula for each subdomain in such a way that this function is at least continuous in the whole space (or in the measurement domain).

The advantage of the dividing of the whole space S into smaller subdomains is that in each subdomain, the particular local interpolating function can be substantially simpler that the function defined by a single formula in the whole space (for example, it can be defined as some polynomial of coordinates of some low order). However, for this advantage we pay with difficulties of defining the suitable subdomains and assuring the continuity of the global interpolating function (and its derivatives, if we require higher level of smoothness).

In the one-dimensional case everything is simple, as the subdomains can be defined naturally as the (open) intervals between successive measurement points, and also the continuity conditions at each measurement point can be easily formulated. In the case of cubic splines, the matrix of this system of conditions is a band matrix and it can be easily inverted; in the case of splines using polynomials of fifth or higher order, the parameters of the interpolating polynomial can be directly calculated from the measured values at the ends of the interval and some neighbouring intervals.

However, if the number of dimensions is $n > 1$, things become more complicated: the particular subdomains can be defined in several ways and it cannot be expected that they can be determined uniquely (only by the position of the measurement points). These subdomains are defined usually by the triangulation of the space S with vertices at the measurement points; thus the subdomains are triangles in the case $n = 2$, tetrahedrons in the case $n = 3$, and so forth. The best result can be obtained by the well known Delaunay triangulation, however, also in this case the triangulation need not be unique and there may be present triangles with very small angles. In spite of the fact that the boundaries of the subdomains obtained in this way are planar and relatively simple, they are nevertheless artificial, as they have no counterpart in the course of the interpolated quantity.

Even if we consider these subdomains only as a tool for the calculation of the interpolating function, it is by no means evident that this division of the space S is the most suitable in view of the principle C0. With respect to this principle it would be more natural to divide the space S into the Voronoi domains (for any measurement point, its Voronoi domain is defined as the domain containing all points that are nearer to this measurement point than to any other measurement point): the interpolated value at any calculation point within one such domain should depend first of all on the measured value at the measurement point lying in this domain and in the neighbouring domains. However, in any case we have to determine for each subdomain its neighbouring subdomains (what need not to be simple in spaces of higher dimension), and then to define the partial interpolating function in each subdomain in such a way that the global interpolating function belongs to the class C_I for some $I \geq 0$ (what may be very difficult if $I > 1$). Thus, these methods can hardly satisfy the conditions C2, C3, C5, C6 and C10; they are not suitable for extrapolation (condition C4).

In this category we can also include methods which do not define explicitly any subdomains (but these arise as a by-product): for example, the interpolation methods using a formula similar to (4), but with the sum restricted to some number of the measurement points which are the nearest to the calculation point. These methods have the disadvantage that even the interpolating function itself is not everywhere continuous.

We finally come to the third category of interpolation methods:

M3: The methods which define the interpolating function separately for each point of the space.

Such a definition of interpolating function may seem at first sight absurd, as there is nothing to interpolate at a single point; it may also seem contraproductive to have a definition different for each point of space for a function that should be maximally smooth in the whole space. Nevertheless, we show that this is the simplest and most effective method for interpolation (to the knowledge of the author, there does not exist any example of such a method).

We saw that the methods of the category M2 are numerically simpler than those of the category M1, because the particular local interpolating function has to be used only in a restricted domain; the disadvantage was that the subdomains divide the whole space S in an unnatural manner. Therefore it may be more natural to introduce no subdomains, but to consider each point of the space S as a separate entity for which we define a function describing the measured values at the nearby measurement points as well as possible. Thus, instead of introducing a single interpolating function $f(X)$ from the beginning, we first try to construct a function $f(X,Y)$ of two variables X and Y, which, as a function of Y, fits the measured values at the measurement points in the neighbourhood of the point X as well as possible. After the function $f(X,Y)$ is determined (separately for each point X), we use it to define the most probable value of the interpolated quantity at the point X by the simple expression

$$f(X) = f(X,X). \tag{16}$$

In view of our conditions, we require that the function $f(X,Y)$ is, as a function of Y, as simple as possible. Therefore it will be reasonable to write this function as a polynomial of functions representing the coordinates of the point Y (see the discussion of the category M1a). It is well known that in the metric space S we can introduce a very wide class of coordinate systems; the coordinates of any such system can be expressed as some smooth functions of the coordinates of any other system. However, the conditions C7 and C8 require that the interpolated value should not depend on the particular choice of some coordinate system; moreover, it should depend only on the quantities that can be defined for any geometry of the space S. In a general case of a space S, there is usually no natural coordinate system like in the euclidean space; it can often happen that there exists no coordinate system covering the whole space S. Therefore it is difficult (if not impossible at all) to find the most suitable coordinate system within the space S.

The solution of this problem can be described as follows. We shall assume that the space S can be isometrically embedded into some euclidean space; let S_e be such euclidean space whose dimension n_e is minimal. Then $n_e \geq n$; if the space S itself is euclidean, we have $n_e = n$. Any point X of the space S is a point of the space S_e; the distance $d_e(X,Y)$ of any two points X and Y of the space S_e is the usual distance in an euclidean space (thus, for any points X and Y of the space S, $d_e(X,Y) \leq d(X,Y)$).

Although we could choose a single euclidean coordinate system that would be used in the whole space S_e, it will be more convenient to define for each point X of the space S its own coordinate system. This will be done as follows: at any point X of the space S we choose n_e mutually orthogonal unit base vectors such that the first n of these vectors will be tangential to the space S; the origin of the coordinate system will be at the point X. The coordinates of any point Y of the space S_e in this system will be denoted as $x_p(X,Y)$, where $p \in I(n_e)$; the coordinates with $p \in I(n)$ then describe the position of a projection of the point Y to the hyperplane $S(X)$ tangential to the space S at the point X.

It is clear that for any point X there is a whole class of such coordinate systems: any such system can be transformed to any other such system by suitable orthogonal transformations with respect to the first n and to the remaining $n_e - n$ coordinates. We see that this definition of coordinates is uniform in the whole space S and also uniform with respect to the global geometry of this space. If the space S is euclidean, the coordinates defined in this way are the usual relative coordinates of the euclidean space (thus there are some global coordinates $x_p(X)$, $p \in I(n)$, such that $x_p(X,Y) = x_p(Y) - x_p(X)$).

Consider now some point X of the space S and some geodesic line (in the space S) beginning at the point X; we define $X(t)$ as the point of this geodesic line whose distance $d(X,X(t))$ (thus the length of the segment of the geodesic line between the points X and $X(t)$) is equal to t. If we have chosen some base vectors at the point X, we can perform the parallel transport of these base vectors along the geodesic line to the point $X(t)$. The base vectors then rotate by this transport as

a whole in such a way that the first n of them are always tangential to the space S; thus we obtain at the point $X(t)$ certain vectors which are also allowed base vectors (at this point). Note that the base vectors obtained by this parallel transport can be different from those chosen at the point $X(t)$: the difference can be always removed by suitable orthogonal transformations as described above.

We shall then impose the following condition on the space S: for any point X of the space S, any geodesic line (in the space S) beginning at the point X, any value of the parameter t such that $t \geq 0$ and such that the point $X(t)$ exists, and any point Y of the space S_e, the coordinates $x_p(X(t), Y)$ $(p \in I(n_e))$ defined by this parallel transport are functions of the class C_∞ of the original coordinates $x_p(X, Y)$ $(p \in I(n_e))$ and of the parameter t.

Then we can conclude that irrespectively of the particular choice of the coordinates $x_p(X, Y)$ $(p \in I(n_e))$ for each point X of the space S, any scalar quantity of the form

$$\sum\nolimits_{p \in I(n_e)} x_p(X, Y)\, x_p(X, Y')$$

or of the form

$$\sum\nolimits_{p \in I(n)} x_p(X, Y)\, x_p(X, Y')$$

(where Y, Y' are arbitrary points of the space S_e) will be a function of the point X belonging to the class C_∞ (evidently, any such quantity will belong to the class C_∞ also as a function of points Y and Y'). In particular, the function $d_e^2(X, Y)$ (where $X \in S$ and $Y \in S_e$) will be such function, as

$$d_e^2(X, Y) = \sum\nolimits_{p \in I(n_e)} x_p(X, Y)^2. \tag{17}$$

Now we express the function $f(X, Y)$ in the form of a polynomial of coordinates $x_p(X, Y)$ $(p \in I(n))$ of the order M whose coefficients will be some functions of X:

$$f(X, Y) = f_0(X) + \sum\nolimits_{m \in I(M)} f_m(X, Y), \tag{18}$$

where

$$f_1(X, Y) = \sum\nolimits_{p \in I(n)} c_p(X)\, x_p(X, Y), \tag{19}$$

$$f_2(X, Y) = \sum\nolimits_{p \in I(n)} \sum\nolimits_{q \in I(n)} c_{pq}(X)\, x_p(X, Y)\, x_q(X, Y), \tag{20}$$

and so on for higher m. This expression of the functions $f_m(X, Y)$ shows explicitly the transformation properties of their coefficients by rotations of the coordinate system at the point X (as the functions $f_m(X, Y)$ themselves should remain unchanged by these rotations). However, because of the natural symmetry of certain coefficients of the functions $f_m(X, Y)$ (for example, $c_{pq}(X) = c_{qp}(X)$), we express the function $f(X, Y)$ also in the form containing only independent coefficients. Let $Q_k(X, Y)$ $(k \geq 1)$ be the mutually different monomials of coordinates $x_p(X, Y)$ $(p \in I(n))$ ordered in some natural order; thus

$$Q_1(X, Y) = 1, \tag{21}$$

and then follow:

for $2 \leq k \leq n + 1$ the n monomials $x_p(X, Y)$ $(p \in I(n))$,

for $n + 2 \leq k \leq 2n + 1$ the n monomials $x_p(X, Y)^2$ $(p \in I(n))$,

for $2n + 2 \leq k \leq (n + 1)(n + 2)/2$ the $n(n - 1)/2$ monomials $x_p(X, Y)x_q(X, Y)$ $(p \in I(n), q \in I(n), p < q)$,

and so forth with increasing k. Then we can write

$$f(X,Y) = \sum\nolimits_{k\in I(K)} a_k(X)\, Q_k(X,Y), \tag{22}$$

where $a_k(X)$ are some (mutually independent) coefficients and

$$K = \frac{(n+M)!}{n!M!} \tag{23}$$

is their total number. From (18) and (22) we easily obtain

$$f(X,X) = f_0(X) = a_1(X); \tag{24}$$

in view of (16) the only coefficient we need to calculate is $a_1(X)$.

Note that the function $f(X,Y)$ is defined for any $X \in S$ and any $Y \in S_e$, but it depends only on the first n coordinates $x_p(X,Y)$ of the point Y in the coordinate system defined at the point X. Thus, if we say that this function acquires (as a function of Y) some value at some point $X_i \in S$, this should be understood as that the point X_i is a point of the space S_e and its last $n_e - n$ coordinates $x_p(X,X_i)$ are ignored. In other words, it should be understood as that this function acquires this value at the point which is the projection of the point X_i to the hyperplane $S(X)$ tangential to the space S at the point X.

Now we define the condition for determination of the coefficients $a_k(X)$: as we want the function $f(X,Y)$ to fit the measured values f_i at the measurement points X_i in the neighbourhood of the point X as well as possible, we introduce the quantity

$$E_N(X) = \sum\nolimits_{i\in I(N)} w(X,X_i)\,(f(X,X_i) - f_i)^2, \tag{25}$$

where $w(X,Y)$ is some suitable nonnegative weighting function, and require that this quantity acquires its minimal value. In order to satisfy the conditions C7 and C8, it is natural to write the function $w(X,Y)$ in the form

$$w(X,Y) = w(d_e(X,Y)), \tag{26}$$

where $w(d)$ is some nonnegative function of d. Note that in principle we could use in this formula also the distance $d(X,Y)$ defined within the space S; but this distance is for a general geometry of the space S more complicated function than the distance $d_e(X,Y)$.

According to the conditions C2, C3 and C10, the function $w(d)$ should be maximally smooth and simple, and the condition C0 requires that $w(d)$ is a nonincreasing function of d for $d \geq 0$. We shall in the next introduce further natural requirements on the form of function $w(d)$, however, in principle it is possible to use any function satisfying the mentioned requirements. From the formulae (22), (25) and (26) we then obtain

$$E_N(X) = \sum\nolimits_{i\in I(N)} w(d_e(X,X_i)) \left(\sum\nolimits_{k\in I(K)} a_k(X)\, Q_k(X,X_i) \;-\; f_i\right)^2. \tag{27}$$

Before we proceed to the equations determining the coefficients $a_k(X)$, we still introduce another term in the quantity that will be minimalized. This regularization term should, first of all, assure the existence of a solution in every possible situation, and it will improve the behaviour of the solution according to our requirements. Such a term can have several different forms; we shall here introduce this term in the form assuring a moderate behaviour of the interpolating function far from the measurement points.

We define the quantity d_1 (which will be called the regularization distance) as such distance (in the space S_e) from the calculation point X beyond which the influence of the measured values on the interpolated value should be negligible. Then, if there are no measurement points X_i such that $d_e(X,X_i) < d_1$, it would be natural to require that the value of $f(X,X)$ should be minimally

different from the mean value of $f(X,Y)$ at all points Y (of the space S) with $d_e(X,Y) = d_1$. However, such a mean value would not be easily calculable; as the function $f(X,Y)$ is a polynomial of the coordinates $x_p(X,Y)$ ($p \in I(n)$), which are coordinates in the tangential space $S(X)$, it will be more convenient to take this mean value with respect to the points Y in the tangential space $S(X)$. Therefore we adopt the following convention: for any point $X \in S$, any distance $d \geq 0$ and any function $g(Y)$ defined in the whole space S_e, the mean value of this function with respect to all points $Y \in S(X)$ such that $d_e(X,Y) = d$ will be denoted as $\langle g(*) \rangle_{X,d}$.

Now we can easily obtain the mean values of every monomial of coordinates $x_p(X,Y)$: this mean value is zero if the degree of monomial is odd, and for the even degrees we have (for $p \in I(n)$, $q \in I(n)$, $r \in I(n)$, $s \in I(n)$)

$$\langle 1 \rangle_{X,d} = 1, \tag{28}$$

$$\langle x_p(X,*)\, x_q(X,*) \rangle_{X,d} = \frac{d^2}{n}\, \delta_{pq}, \tag{29}$$

$$\langle x_p(X,*)\, x_q(X,*)\, x_r(X,*)\, x_s(X,*) \rangle_{X,d} = \frac{d^4}{n(n+2)}\, (\delta_{pq}\, \delta_{rs} + \delta_{pr}\, \delta_{qs} + \delta_{ps}\, \delta_{rq}), \tag{30}$$

and so on. Then we can write the regularization term as

$$E_R(X) = w(d_1)\, \langle (f(X,*) - f(X,X))^2 \rangle_{X,d_1}; \tag{31}$$

using the formulae (22) and (24) we get

$$f(X,Y) - f(X,X) = \sum\nolimits_{k \in I(K|1)} a_k(X)\, Q_k(X,Y) \tag{32}$$

and thus

$$E_R(X) = w(d_1)\, \Big\langle \Big(\sum\nolimits_{k \in I(K|1)} a_k(X)\, Q_k(X,*) \Big)^2 \Big\rangle_{X,d_1}. \tag{33}$$

We finally obtain the expression of the quantity to be minimalized as

$$E(X) = E_N(X) + E_R(X). \tag{34}$$

Performing a variation of the quantity $E(X)$ with respect to all coefficients $a_k(X)$ ($k \in I(K)$) we obtain a system of K equations

$$\sum\nolimits_{l \in I(K)} a_l(X) \Big(A_{kl}(X) + R_{kl}(X) \Big) = F_k(X), \tag{35}$$

where

$$A_{kl}(X) = \sum\nolimits_{i \in I(N)} w(d_e(X,X_i))\, Q_k(X,X_i)\, Q_l(X,X_i), \tag{36}$$

$$R_{kl}(X) = (1 - \delta_{k1})\, (1 - \delta_{l1})\, w(d_1)\, \langle Q_k(X,*)\, Q_l(X,*) \rangle_{X,d_1}, \tag{37}$$

and

$$F_k(X) = \sum\nolimits_{i \in I(N)} w(d_e(X,X_i))\, Q_k(X,X_i)\, f_i. \tag{38}$$

Now we show that the regularization term fulfils its purpose: we first construct the quantities

$$A(X) = \sum\nolimits_{k \in I(K)} \sum\nolimits_{l \in I(K)} a_k(X)\, a_l(X)\, A_{kl}(X), \tag{39}$$

$$R(X) = \sum\nolimits_{k \in I(K)} \sum\nolimits_{l \in I(K)} a_k(X)\, a_l(X)\, R_{kl}(X), \tag{40}$$

and we obtain

$$A(X) = \sum_{i\in I(N)} w(d_e(X,X_i)) \left(\sum_{k\in I(K)} a_k(X)\, Q_k(X,X_i)\right)^2, \tag{41}$$

$$R(X) = E_R(X), \tag{42}$$

thus it holds always $A(X) \geq 0$ and $R(X) \geq 0$. Let us assume that the weighting function $w(d)$ is positive for any value of d ($d \geq 0$); then the equation $A(X) = 0$ implies that the coefficients $a_k(X)$ have to satisfy the conditions

$$i \in I(N): \quad \sum_{k\in I(K)} a_k(X)\, Q_k(X,X_i) = 0. \tag{43}$$

On the other hand, the equation $R(X) = 0$ implies that they have to satisfy the condition

$$\sum_{k\in I(K|1)} a_k(X)\, Q_k(X,Y) = 0 \tag{44}$$

for almost every point $Y \in S(X)$ such that $d_e(X,Y) = d_1$, and this is possible for $d_1 > 0$ only if

$$k \in I(K|1): \quad a_k(X) = 0. \tag{45}$$

Thus, according to (21), (43) and (45) we can have $A(X) = 0$ and $R(X) = 0$ only if

$$k \in I(K): \quad a_k(X) = 0, \tag{46}$$

what implies that for any coefficients $a_k(X)$ which are not identically zero for all $k \in I(K)$ we have

$$A(X) + R(X) > 0. \tag{47}$$

According to (39) and (40) this means that the matrix of the system of equations (35) is positive definite and thus this system has always a unique solution.

4 Properties of the universal interpolating function

In the previous section we have shown the way for determination of the interpolating function $f(X)$: we first construct according to the formulae (36) – (38) the functions $A_{kl}(X)$, $R_{kl}(X)$ and $F_k(X)$ ($k \in I(K)$, $l \in I(K)$), and, solving the linear system of equations (35) we find the function $a_1(X)$; according to (16) and (24) this function is our interpolating function $f(X)$. Now we show that this function satisfies the conditions C1 – C12 and that it conforms to the principle C0.

From the definition of functions $A_{kl}(X)$ and $F_k(X)$ we see that each of them is expressed as a sum of contributions of all measurement points; each term contains the weighting factor depending on the distance of the calculation point and the corresponding measurement point and this factor is a nonincreasing function of this distance. Thus the principle C0 is satisfied if the function $w(d)$ decreases sufficiently with increasing d (this will be discussed below).

In order to investigate whether our solution satisfies the condition C1, we decompose the system of K equations (35) into one equation for $k = 1$ and $K - 1$ equations for $k \in I(K|1)$ (note that according to the formula (23), for $M \geq 0$ we have always $K \geq 1$). Using the fact that $A_{kl}(X)$ and $R_{kl}(X)$ are symmetric matrices and that according to (37) $R_{k1}(X) = 0$ for any $k \in I(K)$, we obtain for $k = 1$ the equation

$$a_1(X)\, A_{11}(X) + \sum_{l\in I(K|1)} a_l(X)\, A_{1l}(X) = F_1(X), \tag{48}$$

and for $k \in I(K|1)$ the system of equations

$$a_1(X)\, A_{k1}(X) + \sum_{l\in I(K|1)} a_l(X) \left(A_{kl}(X) + R_{kl}(X)\right) = F_k(X). \tag{49}$$

For $K = 1$, this system of equations is empty; if $K > 1$, then it is evident that the matrix $A_{kl}(X) + R_{kl}(X)$ ($k \in I(K|1)$, $l \in I(K|1)$) is positive definite, as it is a diagonal block of a positive definite matrix. Therefore this matrix has its inverse matrix which we denote as $B_{kl}(X)$: for $k \in I(K|1)$, $l \in I(K|1)$ it holds

$$\sum\nolimits_{r \in I(K|1)} B_{kr}(X) \Big(A_{rl}(X) + R_{rl}(X) \Big) = \delta_{kl} \tag{50}$$

(note that $B_{kl}(X)$ is also symmetric). Then we obtain from (49) for $k \in I(K|1)$

$$a_k(X) = \sum\nolimits_{l \in I(K|1)} B_{kl}(X) \Big(F_l(X) - a_1(X)\, A_{l1}(X) \Big), \tag{51}$$

and inserting in (48) we get

$$a_1(X)\, B(X) = F(X), \tag{52}$$

where we denoted

$$B(X) = A_{11}(X) - \sum\nolimits_{k \in I(K|1)} \sum\nolimits_{l \in I(K|1)} A_{1k}(X)\, B_{kl}(X)\, A_{l1}(X), \tag{53}$$

$$F(X) = F_1(X) - \sum\nolimits_{k \in I(K|1)} \sum\nolimits_{l \in I(K|1)} A_{1k}(X)\, B_{kl}(X)\, F_l(X). \tag{54}$$

Note that according to (48) the formula (52) holds also for $K = 1$. It can be easily shown that the $(1,1)$-th element of the matrix inverse to the matrix of the system (35) is equal to $1/B(X)$; using the fact that the inverse matrix of a symmetric positive definite matrix is also a symmetric positive definite matrix, and that any diagonal element of a symmetric positive definite matrix is positive, we obtain that the quantity $B(X)$ is always positive.

Now we can consider the case that the calculation point X is closer to some measurement point X_m ($m \in I(N)$) than to all other measurement points (we shall for a moment restrict ourselves to the case that all measurement points are mutually different; the general case will be discussed later). As the function $w(d)$ is a nonincreasing function of d, according to the expression of functions $A_{kl}(X)$ and $F_k(X)$ ((36) and (38)) it is clear that the most important terms will be those corresponding to the measurement points nearest to the calculation point X, first of all the term with $i = m$. If the function $w(d)$ is bounded for $d = 0$, for $X = X_m$ we obtain from the formulae (36) and (38) the expressions

$$A_{kl}(X_m) = w(0)\, \delta_{k1}\, \delta_{l1} + A_{kl|m}, \tag{55}$$

$$F_k(X_m) = w(0)\, \delta_{k1}\, f_m + F_{k|m}, \tag{56}$$

where

$$A_{kl|m} = \sum\nolimits_{i \in I(N|m)} w(d_e(X_m, X_i))\, Q_k(X_m, X_i)\, Q_l(X_m, X_i), \tag{57}$$

$$F_{k|m} = \sum\nolimits_{i \in I(N|m)} w(d_e(X_m, X_i))\, Q_k(X_m, X_i)\, f_i, \tag{58}$$

and we have used the formula

$$Q_k(X, X) = \delta_{k1}, \tag{59}$$

following from the definition of functions $Q_k(X, Y)$. Using (55) and (56) we can write the formulae (53) and (54) for $X = X_m$ in the form

$$B(X_m) = w(0) + A_{11|m} - \sum\nolimits_{k \in I(K|1)} \sum\nolimits_{l \in I(K|1)} A_{1k|m}\, B_{kl|m}\, A_{l1|m}, \tag{60}$$

$$F(X_m) = w(0)\, f_m + F_{1|m} - \sum\nolimits_{k\in I(K|1)} \sum\nolimits_{l\in I(K|1)} A_{1k|m}\, B_{kl|m}\, F_{l|m}, \tag{61}$$

where $B_{kl|m}$ is defined by the formula (holding for $k \in I(K|1)$, $l \in I(K|1)$)

$$\sum\nolimits_{r\in I(K|1)} B_{kr|m} \Big(A_{rl|m} + R_{rl}(X_m) \Big) = \delta_{kl}. \tag{62}$$

Comparing the formulae (36) and (38) with (57) and (58), we see that the quantities $A_{kl|m}$ and $F_{k|m}$ correspond to the quantities $A_{kl}(X)$ and $F_k(X)$, where the term with $i = m$ is removed from the sum and X is replaced by X_m. Then we obtain by the same argument as above that

$$A_{11|m} - \sum\nolimits_{k\in I(K|1)} \sum\nolimits_{l\in I(K|1)} A_{1k|m}\, B_{kl|m}\, A_{l1|m} > 0, \tag{63}$$

and from the formula (52) we get

$$a_1(X_m) = \frac{F(X_m)}{B(X_m)}. \tag{64}$$

From the formulae (60), (61) and (64) we see that if the term containing $w(0)$ in the quantities $B(X_m)$ and $F(X_m)$ is dominant, the value of $a_1(X_m)$ is close to the value of f_m. In order to examine the dependence of quantities $a_1(X_m)$ and f_m, we choose some particular form of the weighting function $w(d)$. As this function should be maximally simple, we shall write it in the form

$$w(d) = (d_0^2 + d^2)^{-\alpha}, \tag{65}$$

where $d_0 \geq 0$ and $\alpha > 0$ are some constants; the quantity d_0 will be called the smoothing distance. If $d_0 = 0$, the condition C3 requires that α should be an integer; in order to satisfy also the condition C10 we choose α to be always an integer and we shall denote it as L. Inserting this expression in (57) and (58) we obtain

$$A_{kl|m} = \sum\nolimits_{i\in I(N|m)} \frac{1}{(d_0^2 + d_e(X_m, X_i)^2)^L}\, Q_k(X_m, X_i)\, Q_l(X_m, X_i), \tag{66}$$

$$F_{k|m} = \sum\nolimits_{i\in I(N|m)} \frac{1}{(d_0^2 + d_e(X_m, X_i)^2)^L}\, Q_k(X_m, X_i)\, f_i; \tag{67}$$

from the formula (37) we get for $k \in I(K|1)$, $l \in I(K|1)$

$$R_{kl}(X_m) = \frac{1}{(d_0^2 + d_1^2)^L}\, \langle Q_k(X_m, *)\, Q_l(X_m, *) \rangle_{X_m, d_1}, \tag{68}$$

and finally

$$w(0) = \frac{1}{d_0^{2L}}. \tag{69}$$

If $d_0 \to 0$, the values of $A_{kl|m}$, $F_{k|m}$ and $B_{kl|m}$ change slowly to their finite limit values, but the value of $w(0)$ grows to infinity and thus the value of $a_1(X_m)$ tends to f_m. However, it is not reasonable to put $d_0 = 0$, although this choice would be attractive because then the interpolating function would satisfy the conditions (3); this is because in this case the regularity of the system of equations (35) breaks down and there may arise problems by the numerical evaluation of the interpolating function (we shall touch on this problem once more later).

Therefore, we prefer the case $d_0 > 0$; although it is hardly possible to estimate easily the difference of $a_1(X_m)$ and f_m analytically, it is sufficient that this can be done numerically. As all quantities appearing in the expression of the function $a_1(X_m)$ are continuous functions of d_0 for $d_0 > 0$, if the calculated value of $a_1(X_m)$ differs from f_m more than it is required, we can choose

a sufficiently small value of the parameter d_0 to obtain the interpolated value which is sufficiently close to f_m.

The smooth dependence of the interpolating function on the parameter d_0 implies that the quantity E defined by the formula (1) is also a continuous function of d_0. Thus, if for some value of d_0 the condition (2) is not satisfied (for some given value of E_0), it is always possible to find some smaller value of d_0 which will satisfy this condition (if there are no identical measurement points). Of course, the suitable value of d_0 has to be found by numerical evaluation of the quantity E.

From the above discussion it may seem that, in order to obtain a sufficiently satisfying approximate interpolation, the smoothing distance d_0 has to be much smaller than some characteristic minimal distance of the measurement points. The latter distance can be defined (for any position of measurement points) as

$$d_c^2 = \frac{1}{N} \sum\nolimits_{i \in I(N)} \min\nolimits_{j \in I(N|i)} d_e(X_i, X_j)^2, \tag{70}$$

where $d_c \geq 0$; we shall call it the rms-minimal distance (of the set of measurement points X_i, $i \in I(N)$). However, we shall present an indirect argument showing that the smoothing distance d_0 need not to be much smaller than d_c.

We first introduce an useful notion concerning the expressibility of the measured data by the monomials $Q_k(X, Y)$. Let J be some subset of $I(N)$; then we shall say that the set of measurement points X_i, $i \in J$, is M-acceptable at the point X (see the formula (23) for the dependence of M and K), just if there exists for any values of f_i ($i \in J$) such function, which is a linear combination of monomials $Q_k(X, Y)$ ($k \in I(K)$), and which acquires as a function of Y at each point X_i ($i \in J$) the value f_i. Note that in the case of euclidean space, if some set of measurement points is M-acceptable at some point X, then it is M-acceptable at any point of the space.

Consider some given set of measurement points X_i, $i \in J$, and let $n(J)$ be the number of these points. It is clear that this set cannot be M-acceptable (at the point X) if $n(J) > K$ or if it contains (at least) two points whose projections to the hyperplane $S(X)$ are identical. On the other hand, if $n(J) \leq K$ and all points of this set have different projections to $S(X)$, this set will be always M-acceptable (at the point X) if $n(J) \leq M + 1$ (we omit the proof of this fact); in the case that $M + 1 < n(J) \leq K$, this set will be in general M-acceptable (at the point X); the contrary will be true only for a special position of points X_i.

Let the set of measurement points X_i, $i \in J$, be M-acceptable at the point X; in accord with the formula (27) we define the quantity

$$E_{N,J}(X) = \sum\nolimits_{i \in J} w_i(X) \left(\sum\nolimits_{k \in I(K)} a_k(X)\, Q_k(X, X_i) \; - \; f_i \right)^2, \tag{71}$$

where $w_i(X)$ are arbitrary positive weights. Then it is evident that the quantity $E_{N,J}(X)$ is nonnegative and it acquires its minimal value equal to zero for at least one choice of the coefficients $a_k(X)$ ($k \in I(K)$). The most important fact is that $E_{N,J}(X)$ acquires this minimum irrespective of the value of weights $w_i(X)$ (if all of them are positive).

Now we return to the case that all measurement points are mutually different and the calculation point X is identical with the point X_m. Consider any (nonempty) set of measurement points X_i, $i \in J$, which is M-acceptable at the point X_m, and, for any $j \in I(N)$, $j \notin J$, it holds $d_e(X_m, X_j) \geq \max_{i \in J} d_e(X_m, X_i)$. From the above considerations it follows that there exists at least one such set (the set consisting of the single point X_m), and any such set contains the point X_m. Then we can choose from among these sets the set with the greatest number of points (there can be more such sets, but we can choose any of them). Let the chosen set be that containing the points X_i, $i \in J$; then $1 \leq n(J) \leq K$, and the most probable case is that $n(J) = K$.

Now, the system of equations (35) represents the conditions for a minimum of the quantity $E(X)$; we shall consider the formulae (27), (34) – (36), and (71) for the case $X = X_m$. The quantity $E(X_m)$ can be written as a sum of three terms: the part of the sum (27) with $i \in J$ (this

is equal to the quantity $E_{N,J}(X_m)$, where $w_i(X_m) = w(d_e(X_m, X_i)))$, the part of the sum (27) with $i \notin J$, and the quantity $E_R(X_m)$. Consequently, the matrix of the system (35) consists of three parts: the part of $A_{kl}(X_m)$ with the sum restricted to $i \in J$, the part of $A_{kl}(X_m)$ with the sum restricted to $i \notin J$, and $R_{kl}(X_m)$. If the second and third part would be neglected, the system (35) would have at least one solution that acquires at the point X_m the value f_m irrespectively of the values of parameters of the weighting function $w(d)$ (provided the weight remains bounded everywhere). Thus, if the second and third part are much smaller than the first one, the value of the solution will be at the point X_m close to f_m. The third part (coming from the regularization term) should be always very small in order not to disturb the solution at points within the measurement domain; this is achieved by the suitable choice of the regularization distance d_1. The second term comes from the measurement points which are farther from X_m than those contributing to the first term; thus the second term can be much smaller than the first one if the weighting function $w(d)$ decreases with the distance sufficiently fast. This means according to (65) that it is not so important to have the smoothing distance d_0 much smaller than the rms-minimal distance d_c, but that the value of the exponent L should be sufficiently large.

The question is what is the suitable value of the exponent L in order to conform to the principle C0. According to the previous paragraph we can estimate this value as follows: for any point X within the measurement domain, the number of measurement points within the hypersphere (in the space S_e) with the centre X and the diameter λd_c is roughly of the order λ^n. According to the formula (23), the value of K is roughly equal to $(M+1)^n$; thus the number of measurement points within the hypersphere with the diameter $(M+1)\,d_c$ will be roughly of the same order as K. As the nearest measurement point can be expected at the distance $d_c/2$ from the point X, from the expression (65) we obtain the inequality

$$\frac{1}{(d_0^2 + (d_c/2)^2)^L} \gg \frac{1}{(d_0^2 + ((M+1)d_c/2)^2)^L}, \tag{72}$$

and for $d_0 = d_c/2$ we have

$$(1 + (M+1)^2)^L \gg 2^L. \tag{73}$$

Another condition on the exponent L we obtain as follows. Consider the formulae (36) and (38): both quantities $A_{kl}(X)$ and $F_k(X)$ are expressed as a sum of contributions of all measurement points. These points belong to the measurement domain which is, of course, bounded. However, we can imagine that additional measurements will be performed in the future; let us assume that these additional measurement points will all lie outside the present measurement domain and their mean density will be the same as for the present measurement points (thus the value of the rms-minimal distance d_c will remain the same). Consider a calculation point X lying within the present measurement domain in a distance from its boundary greater than $(M+1)\,d_c$; then all additional measurement points will lie farther from X than some number of the present measurement points (this number will be of the same order as K). We may require that the presence of the additional measurement points will not change substantially the interpolated value at the point X. The number of the additional measurement points will be in any case finite, but we may in principle imagine that the whole space S will be filled with new measurement points (with the same constant mean density); if the space S contains infinitely distant points, there will be an infinite number of the new measurement points. Therefore we should require that the sums in (36) and (38) converge if the number of measurement points grows to infinity (but the mean density of these points remains the same) and the measured values remain bounded.

From the definition of functions $Q_k(X, Y)$ and coordinates $x_p(X, Y)$ we obtain that the function $Q_k(X, Y)$ ($k \in I(K)$) grows with the distance $d_e(X, Y)$ at most as $d_e(X, Y)^{h(k)}$, where $h(k)$ is a nondecreasing function of k: we have $h(1) = 0$ and the maximum of $h(k)$ (for $k \in I(K)$) is M (see (18) – (23)). For the hyperspherical layer with the centre X, the inner radius d and thickness

d_c, the number of the (additional) measurement points in this layer will be proportional to d^{n-1}. Then the contribution of the measurement points in this layer to the quantity $A_{kl}(X)$ ($F_k(X)$) will be proportional to $d^{n-1-2L+h(k)+h(l)}$ ($d^{n-1-2L+h(k)}$). We put $d = \lambda d_c$; then the sum in (36) and (38) can be converted to the sum with respect to λ from some fixed integer value to infinity or to an integral with respect to λ from some fixed value to infinity. In order that all these integrals converge it is necessary that $n - 1 - 2L + 2M < -1$; thus we obtain the condition

$$2L > n + 2M. \tag{74}$$

Up to now we have considered only the case that all measurement points are mutually different; now we shall assume that the position of the measurement points is arbitrary (but that they are fixed for a moment). Let $\bar{N}$ be the number of different measurement points (thus $\bar{N} \leq N$); we define the function $u(i)$ ($i \in I(N)$) which acquires values from $I(\bar{N})$ and for any $i \in I(N)$, $j \in I(N)$, the equality $u(i) = u(j)$ holds if and only if $X_i = X_j$. Further we define for each $u \in I(\bar{N})$: the set of indices $J(u)$ such that $i \in J(u)$ just if $u(i) = u$, the number of members of $J(u)$ as $m_u = n(J(u))$, the measurement point with the new index $\bar{X}_u$ as $\bar{X}_u = X_i$ for any i such that $u(i) = u$, and finally, the modified measured value $\bar{f}_u$ as

$$\bar{f}_u = \frac{1}{m_u} \sum\nolimits_{i \in J(u)} f_i. \tag{75}$$

In other words, the points $\bar{X}_u$ ($u \in I(\bar{N})$) are mutually different; for each $u \in I(\bar{N})$, m_u is the number of measurement points identical with the point $\bar{X}_u$ and $\bar{f}_u$ is the mean value of the measured values at the points identical with the point $\bar{X}_u$.

Using the above definitions, we can rewrite the formulae (36) and (38) as follows:

$$A_{kl}(X) = \sum\nolimits_{i \in I(\bar{N})} m_i \, w(d_e(X, \bar{X}_i)) \, Q_k(X, \bar{X}_i) \, Q_l(X, \bar{X}_i), \tag{76}$$

$$F_k(X) = \sum\nolimits_{i \in I(\bar{N})} m_i \, w(d_e(X, \bar{X}_i)) \, Q_k(X, \bar{X}_i) \, \bar{f}_i. \tag{77}$$

We see that the only difference with respect to the original formulae is the presence of the multiplicity factors m_i. Therefore the whole contents of the present section can be adapted for the case that some measurement points are identical: we have only to multiply the weighting function by the corresponding multiplicity factor and replace the measured values by the corresponding mean values. Then we can conclude that the interpolated value at some point $\bar{X}_m$ ($m \in I(\bar{N})$) will be close to the mean measured value $\bar{f}_m$. It is now clear why we have adopted the condition (2) instead of the stricter condition (3): in the case that there are some identical measurement points, the condition (2) cannot be satisfied at all.

It may be argued that the normal situation is such that the measurement points are mutually different; however, if there are any two of them that are mutually much nearer than the rms-minimal distance d_c, the condition (3) can be satisfied only at a cost of an enormous absolute value of the gradient of the interpolating function in the domain between these two points. This may be accepted if we are sure that the measured values at these two points are absolutely correct; but otherwise it is better to assume that the absolute value of the gradient of the interpolating function is not extremely high. If we adopt the restriction $d_0 > 0$, the interpolating function $f(X)$ is according to the formulae (36), (38) and (65) continuous with respect to all variables X_i ($i \in I(N)$). Therefore, if we move any two measurement points to some common position (without changing the measured values at these two points), the interpolated value at any calculation point will change continuously.

In view of the presented facts we adopt definitively the restriction $d_0 > 0$ and we can change the definition (65) of the weighting function $w(d)$ to satisfy the condition $w(0) = 1$:

$$w(d) = \left(\frac{d_0^2}{d_0^2 + d^2} \right)^L. \tag{78}$$

We still have to consider the question what is the suitable value of the constant M, which is the order of the polynomial of coordinates of the point Y expressing the function $f(X,Y)$. It is clear that higher value of M means that the function $f(X,Y)$ can better fit (as a function of Y) the measured values at the measurement points in the neighbourhood of the calculation point X, but the number of coefficients of this function (and thus also the number of equations of the system (35)) is higher. On the other hand, if $M = 0$, we obtain according to the formulae (36) – (38) the interpolating function $f(X)$ in the form (12), which has been rejected as unsatisfactory. The reason of this behaviour is now clear: the discussion in the present section shows that the interpolating function $f(X)$ is in some sufficiently small neighbourhood of any measurement point X_m very close to the function $f(X_m, X)$. If $M = 0$, the function $f(X,Y)$ is (as a function of Y) constant; therefore the interpolating function $f(X)$ is in some small neighbourhood of the point X_m almost constant and its first derivatives are almost zero. The case $M = 1$ is slightly better, as then the function $f(X,Y)$ is a linear polynomial of coordinates of the point Y, and thus the interpolating function $f(X)$ is in some small neighbourhood of the point X_m almost linear and its first derivatives are almost constant. The lowest satisfactory value of M is $M = 2$: the function $f(X,Y)$ is then a quadratic polynomial of coordinates of the point Y, and thus it can have (as a function of Y) a local extremum whose position can be arbitrary (it need not to lie in the vicinity of any measurement point). Therefore the local extrema of the interpolating function $f(X)$ are not forced to lie only in the vicinity of some measurement point (what is the case for $M = 0$). Thus, if the simplicity (both theoretical and numerical) has the priority, the most suitable value of the constant M is $M = 2$; in this case the number K of equations of the system (35) is according to (23) equal to $(n+1)(n+2)/2$. From the condition (74) we obtain the bound $2L > n+4$ and then the condition (73) is also satisfied.

If $M = 2$, we can easily obtain the explicit formula for the regularization matrix $R_{kl}(X)$. For brevity we first define the sets $I_1(n)$, $I_2(n)$, $I_3(n)$, as the sets of all natural numbers k such that $2 \le k \le n+1$, $n+2 \le k \le 2n+1$, $2n+2 \le k \le (n+1)(n+2)/2$, respectively. Using the definition of the monomials $Q_k(X,Y)$ and the formulae (28) – (30), we get from (37) that the elements of the matrix $R_{kl}(X)$ are nonzero only within three diagonal blocks: if $k \in I_1(n)$, $l \in I_1(n)$,

$$R_{kl}(X) = w(d_1)\,\frac{d_1^2}{n}\,\delta_{kl}, \tag{79}$$

if $k \in I_2(n)$, $l \in I_2(n)$,

$$R_{kl}(X) = w(d_1)\,\frac{d_1^4}{n(n+2)}\,(1 + 2\delta_{kl}), \tag{80}$$

and, if $k \in I_3(n)$, $l \in I_3(n)$,

$$R_{kl}(X) = w(d_1)\,\frac{d_1^4}{n(n+2)}\,\delta_{kl}. \tag{81}$$

Now we can recapitulate the properties of our interpolating function $f(X)$. We have shown that this function conforms to the principle C0 and that it satisfies the condition C1 in the sense of the approximate interpolation: if all measurement points are mutually different, the condition (2) can be satisfied by a suitable choice of the smoothing distance d_0. In the case that there are some identical measurement points, the condition (2) can be satisfied if we replace in the formula (1) the measured values by the modified measured values defined by the formula (75).

According to the formulae (17), (22) and (35) – (38) we can conclude that the coefficients $a_k(X)$ ($k \in I(K)$) (which are the unique solution of the system (35)) transform by any rotation of the coordinate system at the point X in such a way that the function $f(X,Y)$ is a scalar quantity (thus it is unchanged by these rotations). If the weighting function $w(d)$ belongs (as a function of d^2) for $d \ge 0$ to the class C_∞ (this is the case if this function is defined by the formula (78)), and the

smoothness condition on the space S (see the paragraphs before the formula (17)) is satisfied, then the function $f(X,Y)$ belongs as a function of X in the whole space S to the class C_∞. Moreover, it belongs to the same class as a function of Y (for any $Y \in S_e$) and as a function of each point X_i ($i \in I(N)$, $X_i \in S$). Thus, according to (16), the interpolating function $f(X)$ belongs as a function of X and all X_i ($i \in I(N)$) in the whole space S to the class C_∞. Further, the behaviour of the function $f(X)$ in the neighbourhood of any measurement point has no anomalous properties (if we have $M \geq 2$). Thus we can conclude that this function satisfies the conditions C2, C3 and also C5 (as the position of measurement points can be arbitrary).

The condition C4 is satisfied by the presence of the regularization term; the function $f(X)$ tends at points far from the measurement domain to the mean value of all measured values f_i ($i \in I(N)$). The condition C6 is evidently satisfied: any new measurement point only increases the value of N in the formulae (36) and (38). Also the conditions C7 and C8 are evidently satisfied, as the interpolating function was constructed according to their requirements. If the number n of dimensions of the space S grows, the only substantial change is the increase of the number K of the monomials $Q_k(X,Y)$, and thus also of the number of equations of the system (35). There are also minor changes in the definition (17) of the distance $d_e(X,Y)$ and in the regularization term (see the formulae (28) – (30)). The basic formulae (22) and (35) – (38) remain unchanged, and thus the condition C9 is satisfied.

The interpolating function $f(X)$ was constructed to be maximally simple; the most complicated step in the calculation of this function is the solving of the system (35) of K equations (note that we need to calculate only the coefficient $a_1(X)$). The choice (78) of the form of the weighting function $w(d)$ and the choice $M = 2$ (quadratic approximation by the function $f(X,Y)$) make the interpolating function $f(X)$ still simpler. Therefore we can say that the condition C10 is satisfied. The condition C11 is satisfied by the presence of the regularization term: the numerical calculation of the solution of the system (35) can be performed for any positions of the calculation point X and the measurement points X_i; if we choose the regularization distance d_1 properly, it cannot happen that the numerical calculation would be broken by the rounding errors. Finally, the interpolating function $f(X)$ belongs with respect to its dependence on the number N of the measurement points to the class F_1, and thus, the condition C12, which is the most important from the viewpoint of the numerical calculation, is satisfied.

Moreover, as a consequence of satisfying the principle C0, the interpolating function $f(X)$ has an important additional property: as the weighting function $w(d)$ decreases with the distance d, the contribution of the measurement points that are sufficiently distant from the calculation point X can be so small (with respect to the contribution of the nearby measurement points) that it can be numerically negligible. Of course, we have to take into account that the particular functions $A_{kl}(X)$ and $F_k(X)$ (defined by the formulae (36) and (38)) behave in this respect differently (see the discussion in the paragraph preceding the formula (74)). However, it is evident that in the case $M = 2$, it suffices to consider the behaviour of quantities

$$C_1(X) = A_{11}(X), \quad C_2(X) = \sum\nolimits_{k \in I_1(n)} A_{kk}(X), \quad C_3(X) = \sum\nolimits_{k \in I_2(n)} \sum\nolimits_{l \in I_2(n)} A_{kl}(X), \quad (82)$$

$$C_4(X) = F_1(X), \quad C_5(X) = \sum\nolimits_{k \in I_2(n)} F_k(X), \quad (83)$$

(for the definitions of $I_1(n)$ and $I_2(n)$ see the paragraph preceding the formula (79)), which are invariant with respect to the rotations of the coordinate system at the point X. Moreover, according to the definition of functions $Q_k(X,Y)$ and the formula (36), each of the quantities $C_1(X)$ – $C_3(X)$ can be expressed as sum of contributions of all measurement points and all these contributions are nonnegative. From the previous discussion it is clear that the quantity $C_3(X)$ is the most sensitive to the contribution of the distant measurement points.

Now, let ε ($\varepsilon > 0$) be the relative accuracy of the numerical calculation (for example, the smallest real number which can be internally expressed by the given number of significant digits).

Let us consider the quantity $C_3(X)$ at any given calculation point X: for any distance $d > 0$ we can divide the measurement points X_i ($i \in I(N)$) into two groups, the first one containing points with the distance $d_e(X, X_i) \leq d$ and the second one containing points with the distance $d_e(X, X_i) > d$, and we can calculate the contributions of both groups to the value of $C_3(X)$. It is evident that there always exists the smallest distance d such that the ratio of the contribution of the second group to the contribution of the first group is smaller than (or equal to) ε. We denote this boundary distance as $d_b(X)$; it depends on the position of the point X, as the largest contribution to the quantity $C_3(X)$ comes from the measurement point which is the nearest to the point X. However, the most important parameter governing the value of $d_b(X)$ is the value of the exponent L appearing in the definition of the weighting function (78): the distance $d_b(X)$ decreases with the increasing value of L.

We see that by the numerical calculation of the interpolated value at any point X only a limited number of the measurement points will be actually used: these are the points X_i satisfying the condition $d_e(X, X_i) \leq d_b(X)$. It is possible to develop algorithms which allow to determine efficiently the value of the boundary distance in the course of the calculation of quantities $A_{kl}(X)$ and $F_k(X)$ and thus to limit the calculation of contributions of measurement points lying beyond this distance in a maximal manner. However, it seems to be more efficient (at least in the case that the set of the measurement points is not very irregular) to estimate the contribution of all measurement points lying beyond certain fixed distance d_b (independent on the position of the calculation point X) for each quantity $A_{kl}(X)$ and $F_k(X)$ (in fact, it is sufficient to estimate this contribution for the quantities $C_1(X)$ – $C_5(X)$). This can be done in a manner similar to the considerations in the paragraph preceding the formula (74), if we assume that the measurement points, whose distance from the point X is greater than d_b, are distributed with a constant mean density and the measured values at these points are equal to the mean value of all measured values. These assumptions can be considered as acceptable because the contribution of the measurement points lying beyond a suitably chosen distance d_b will be much more smaller than the contribution of the nearby measurement points. Moreover, it can be shown that the contribution of the distant measurement points obtained in this way has a similar (though not identical) effect as the regularization term and that the boundary distance d_b should satisfy the condition $d_b \geq d_1$. The number of the measurement points lying within this boundary distance can be estimated as $\Omega_n (d_b/d_c)^n$, where Ω_n is the volume of the n-dimensional unit ball.

5 Conclusion

The interpolation method developed in the previous sections is universal, as it can be used for the spaces of any dimension and geometry and for the measured data for any configuration of the measurement points. However, it is universal also in a broader sense: it can be used besides for the approximate interpolation of the measured data (what is its primary purpose) also for the smoothing of the measured data (if we adopt a value of the smoothing distance d_0 as some multiple of the rms-minimal distance d_c), and finally, it can be used for the approximation of the measured data (if there are several measured values at each measurement point). Moreover, in a limited sense this method allows to perform the extrapolation of the measured data (if the extrapolated value should behave in general as some polynomial of coordinates of the order M). All this can be accomplished with the single formula (only with the possible adjustment of the two free parameters d_0 and d_1).

The regularization term assures that the calculation always leads to a unique result; it can be even shown that its presence has the following consequence: if the space S is not euclidean and it is embedded in the euclidean space S_e, we need not to perform the interpolation as described here (thus using at any point X only the coordinates in the tangential space $S(X)$), but we can work directly in the space S_e and use all n_e coordinates (of course, then the regularization term

also uses all n_e coordinates). This has the advantage that we then work in an euclidean space and the calculation can be simpler (at the cost of the increased number of parameters of the function $f(X, Y)$).

The presented method has the property that the numerical calculation of the interpolated value with the prescribed accuracy can be performed using only a limited number of measurement points lying within certain distance from the calculation point. Therefore this method can be used to perform the interpolation even for a huge data sets; the calculation time represents no problem, as the calculation of the interpolated values is evidently parallelizable.

We can finally mention that the described method can be easily generalized in several ways (if we drop the condition of the maximal simplicity). For example, we can define the function $f(X, Y)$ (see the formula (22)) as a linear combination of any basic functions (provided that the independence on the coordinate system and sufficient smoothness of the solution is assured). There is a straightforward generalization to a rational function of coordinates: the nominator will be the same polynomial of coordinates as on the rhs of (22), while the denominator can be chosen as some fixed rotationally invariant polynomial of these coordinates. This choice has the advantage that for even values of M we can get the function $f(X, Y)$ that is bounded everywhere what may further improve the properties of the interpolating function (it may reduce the number of data points needed for the numerical calculation of this function).

At last, we can define the weighting function $w(X, Y)$ (see the formula (26)) more generally; we can also introduce in the sum in the formula (27) any positive multiplicity factor (depending on i). We can adopt any other form of the regularization term, provided that it assures the existence and uniqueness of the solution of the system of equations (35). All this shows that the presented method can be adapted to a very wide class of particular situations and requirements on the solution.

References

[1] Billings S. D., Beatson R. K., Newsam G. N., 2002: Interpolation of geophysical data using continuous global surfaces. Geophysics **67**, 6, 1810–1822.

[2] Pohánka V., 1982: A simple formula for interpolation in n-dimensional space. Studia Geoph. et Geod., **26**, 1, 17–23.