

## ON A PROBLEM ASSOCIATED WITH APPROXIMATION BY EXPONENTIAL FUNCTIONS

R.A. SHARIPOV

**Abstract.** While formalizing a certain problem of numeric signal processing there arises a mathematical problem on approximating a square integrable function defined on some finite interval of the real line by linear combinations of exponential functions. This problem is solved as an optimization problem by means of minimizing the root-mean-square deviation with respect to the coefficients of the linear combination and with respect to the exponents of the exponential functions. In some cases, minimizing with respect to the exponents, a computational singularity occurs due to small denominators. In the present paper this singularity is shown to be removable and a mechanism of its removal is described.

**Keywords:** spectrum of a signal, approximation by exponential functions, root-mean-square deviation, small denominators.

**Mathematics Subject Classification:** 46E30, 41A30, 65D15, 68W25

### 1. INTRODUCTION

Signals being a mixture of damped and non-damped sine oscillations with several frequencies appear in various measurements. The frequencies of particular components in the mixture form the signal spectrum, while the procedure of measuring consists in enhancing one or several components from the mixture and to determining their amplitudes. In the most cases the enhancing of the signals is performed by circuit solutions, specifically by resonant filters. But sometimes it is impossible to use such filters, for instance, in the case of very low frequencies, where the periods of sinusoidal components exceed the time of signal measuring. In this case the signal is digitized and the problem of splitting separating it into components by means of digital algorithms arises. In such formulation the problem was proposed to the author by A.S. Vishnevskii, the chief executive officer of the group of companies “PhysTech”.

A measured and digitized signal can be regarded as a function  $f(x)$  defined by its values at a finite set of points on the real axis. However, the discreteness of the argument is not an essential simplification in the mathematical formulation of the problem. This is why in what follows we regard the downstream signal  $f(x)$  as a function defined in some interval  $x \in [a, b] \subset \mathbb{R}$ . We denote by  $\phi(x)$  an approximating function for  $f(x)$  and we choose it as the linear combination

$$\phi(x) = a_1 \phi_1(x) + \dots + a_n \phi_n(x), \quad (1.1)$$

where functions  $\phi_1(x), \dots, \phi_n(x)$  are the exponential functions with pairwise distinct complex exponents  $\lambda_1, \dots, \lambda_n$ :

$$\phi_i(x) = e^{\lambda_i x}, \quad i = 1, \dots, n. \quad (1.2)$$

Exponential functions with complex exponents (1.2) in (1.1) simulate damped sinusoidal signals. They are complex-valued function and this is why coefficients  $a_1, \dots, a_n$  in (1.1) are to be regarded as complex numbers.

In practice signal  $f(x)$  is real-valued. But its realness does not lead to an essential simplification. Because of this reason and for more generality, in what follows function  $f(x)$  is assumed to be complex-valued.

The approximation of function  $f(x)$  by function  $\phi(x)$  in (1.1) means that their difference  $f - \phi$  is to be small in some sense. One of the options of understanding the smallness for the root-mean-square deviation  $f - \phi$  is in the sense of  $\|f - \phi\|$ . In our case  $\|f - \phi\|$  is calculated by the formula

$$\|f - \phi\|^2 = \frac{1}{b-a} \int_a^b |f(x) - \phi(x)|^2 dx. \quad (1.3)$$

The minimization of root-mean-square deviation  $\|f - \phi\|$  is the base of the least square method [1] often employed in signal processing for checking various dependencies.

In practice signal  $f(x)$  is a continuous function. But on the theoretical level we can decline this condition. As the only condition ensuring the well-definiteness of integral (1.3), we choose the square integrability of function  $f(x)$ . Without loss of generality we can assume that  $a = -\pi$  and  $b = +\pi$ . Then formula (1.3) becomes

$$\|f - \phi\|^2 = \frac{1}{2\pi} \int_{-\pi}^{+\pi} |f(x) - \phi(x)|^2 dx, \quad (1.4)$$

that makes sums (1.1) more similar to partial sums for series of exponential functions [2], whose particular cases are Fourier series.

As a result of the above formalization of signal processing, the following mathematical problem arises.

**Problem 1.1.** *Given a function  $f(x)$  in the Hilbert space of square integrable functions  $L^2([-\pi, +\pi])$ , for each fixed  $n$  find sum  $\phi(x)$  defined by (1.1) providing the best approximation of function  $f(x)$  in the sense of  $L^2$ -norm  $\|f - \phi\|$  introduced in (1.4).*

Root-mean-square deviation  $\|f - \phi\|$  in (1.4) depends on  $2n$  complex variables  $a_1, \dots, a_n$  and  $\lambda_1, \dots, \lambda_n$ . We introduce the notation

$$F(a_1, \dots, a_n, \lambda_1, \dots, \lambda_n) = \|f - \phi\|^2. \quad (1.5)$$

Then problem 1.1 is formulated as the problem of finding the global minimum of function (1.5). It is split into two problems. The first of them is to find the minimum for function (1.5) w.r.t. variables  $a_1, \dots, a_n$  under fixed variables  $\lambda_1, \dots, \lambda_n$ :

$$\Phi(\lambda_1, \dots, \lambda_n) = \min_{a_1, \dots, a_n} F(a_1, \dots, a_n, \lambda_1, \dots, \lambda_n). \quad (1.6)$$

The second problem is to find the minimum for function (1.6) w.r.t.  $\lambda_1, \dots, \lambda_n$ :

$$F_{\min} = \min_{\lambda_1 \neq \dots \neq \lambda_n} \Phi(\lambda_1, \dots, \lambda_n). \quad (1.7)$$

The problem of finding the minimum in (1.6) appears to be linear and solvable by standard technique of linear algebra. Its solution is provided below, cf. Section 2. Problem of finding the minimum in (1.7) is nonlinear. This is why problem 1.1 is not always solvable in its original form. In the present work we study one of such cases, where the minimum (1.7) does not exist, while the infimum

$$F_{\min} = \inf_{\lambda_1 \neq \dots \neq \lambda_n} \Phi(\lambda_1, \dots, \lambda_n)$$

is formed in the vicinity of a singular point, where at least one of non-coinciding condition  $\lambda_i \neq \lambda_j$  for  $i \neq j$  fails. Work [3], where problem 1.1 was studied for function  $f(x) = \text{sign}(x)$  as  $n = 1$  and  $n = 2$ , shows that it is indeed sometimes the case in numerical practice.

The main result of the present paper is the proof that singular points violating the condition  $\lambda_i \neq \lambda_j$  for  $i \neq j$  are removable. Despite of a small denominator in the formula for  $\Phi(\lambda_1, \dots, \lambda_n)$ , function  $\Phi(\lambda_1, \dots, \lambda_n)$  has a finite limit at these points. This limit can be calculated explicitly and then we can formulate a generalized formulation for problem (1.1) admitting the repeated numbers in the sequence  $\lambda_1, \dots, \lambda_n$  with arbitrary multiplicities.

## 2. SOLUTION TO LINEAR PROBLEM

The linear part of problem 1.1 is to find the minimum in (1.6). By (1.5) and (1.1) we get easily the formula

$$F = \|f\|^2 - \sum_{j=1}^n a^j \langle f | \phi_j \rangle - \sum_{i=1}^n \overline{a^i} \langle \phi_i | f \rangle + \sum_{i=1}^n \sum_{j=1}^n g_{ij} \overline{a^i} a^j. \quad (2.1)$$

By means of the angle brackets in (2.1) we denote the scalar product in  $L^2$ :

$$\langle a | b \rangle = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \overline{a(x)} b(x) dx. \quad (2.2)$$

The overline in (2.1) and (2.2) stands for the complex conjugation. By  $g_{ij}$  in (2.1) we denote the entries of Gram matrix  $G$ :

$$g_{ij} = \langle \phi_i | \phi_j \rangle. \quad (2.3)$$

By (2.2) we obtain the following property of the  $L^2$ -scalar product:

$$\overline{\langle a | b \rangle} = \langle b | a \rangle. \quad (2.4)$$

Then by (2.4) we get the identities

$$g_{ij} = \overline{g_{ji}}, \quad \langle f | \phi_i \rangle = \overline{\langle \phi_i | f \rangle}. \quad (2.5)$$

The minimum of function (1.5) w.r.t. the variables  $a_1, \dots, a_n$  is determined by the vanishing of its derivatives w.r.t. these variables:

$$\frac{\partial F}{\partial a^i} = 0, \quad \frac{\partial F}{\partial \overline{a^i}} = 0 \quad \text{where } i = 1, \dots, n. \quad (2.6)$$

Calculating derivatives (2.6) for function (2.1), we obtain the equations

$$\sum_{i=1}^n g_{ij} \overline{a^i} = \langle f | \phi_j \rangle, \quad \sum_{j=1}^n g_{ij} a^j = \langle \phi_i | f \rangle, \quad (2.7)$$

where  $i = 1, \dots, n$ . Equations (2.7) are linear w.r.t. variables  $a_1, \dots, a_n$ . Because of this fact, the problem of finding the minimum (1.6) is called the linear problem.

Due to (2.5), two systems of equations (2.7) are distinguished just by the complex conjugation. This means that it is sufficient to solve only one of these systems of equations.

We shall solve the second system of equations in (2.7). It is solved by means of the inverse Gram matrix  $G^{-1}$ . We denote by  $g^{ij}$  the entries of the matrix transpose to the inverse of Gram matrix:  $(G^{-1})^\top$ . Then quantities  $g_{ij}$  and  $g^{ij}$  related as follows:

$$\sum_{k=1}^n g_{ik} g^{jk} = \delta_i^j, \quad \sum_{k=1}^n g^{kj} g_{ki} = \delta_i^j. \quad (2.8)$$

Here  $\delta_i^j$  are the entries of the unit matrix. In linear algebra and in tensor calculus it is called the Kronecker delta.

Applying the second relation in (2.8) to the second system of equations (2.7), we find its solution. It is given by the formula

$$a^j = \sum_{i=1}^n g^{ij} \langle \phi_i | f \rangle \quad \text{where } j = 1, \dots, n. \quad (2.9)$$

Substituting (2.9) into (2.1), we obtain

$$\Phi(\lambda_1, \dots, \lambda_n) = \|f\|^2 - \sum_{i=1}^n \sum_{j=1}^n g^{ij} \langle \phi_i | f \rangle \langle f | \phi_j \rangle. \quad (2.10)$$

Formulae (2.9) and (2.10) give the solution to the linear problem of finding minimum (1.6).

### 3. CONVERGENCE OF SUBSPACES IN HILBERT SPACES

Geometrically, the value of function  $\Phi(\lambda_1, \dots, \lambda_n)$  calculated by formula (2.10) is interpreted as the  $L^2$ -distance from function  $f(x) \in L^2([-\pi, +\pi])$  to  $n$ -dimensional subspace

$$L = \langle \phi_1(x), \dots, \phi_n(x) \rangle \quad (3.1)$$

generated by the exponential functions (1.2) in Hilbert space  $L^2([-\pi, +\pi])$ . Subspace (3.1) depends on  $\lambda_1, \dots, \lambda_n$ . To describe such a dependence, we shall make use of the following definition.

**Definition 3.1.** *We say that a sequence  $L_q$  of  $n$ -dimensional subspaces in Hilbert space  $\mathcal{H}$  converges to an  $n$ -dimensional subspace  $M$  in this space, if in each subspace  $L_q$  there exists a basis  $\mathbf{e}_{1q}, \dots, \mathbf{e}_{nq}$  and in  $M$  there exists a basis  $\mathbf{e}_1, \dots, \mathbf{e}_n$  such that*

$$\mathbf{e}_{iq} \rightarrow \mathbf{e}_i \quad \text{as } q \rightarrow \infty$$

*in the sense of the norm of Hilbert space  $\mathcal{H}$ .*

The definition of convergence of subspaces in a finite-dimensional complex space can be found in [4]. Definition 3.1 agrees with the definition in [4] and is its natural generalization for Hilbert spaces (see Corollary 1.5.5 of Proposition 1.5.4 in [4]).

Each of exponential functions  $\phi_i = e^{\lambda_i x}$  depends continuously on parameter  $\lambda_i$  in the sense of norm in space  $L^2([-\pi, +\pi])$ . It implies the continuous dependence of scalar products  $\langle \phi_i | f \rangle$  and  $\langle f | \phi_j \rangle$ , as well as of the entries of Gram matrix  $g_{ij}$  in (2.3) on parameters  $\lambda_1, \dots, \lambda_n$ . But apart from the mentioned ingredients, formula (2.10) involves also entries  $g^{ij}$  of matrix  $(G^{-1})^\top$ . At singular points, where the condition  $\lambda_i \neq \lambda_j$  for  $i \neq j$  is violated, exponential functions (1.2) become linearly dependent and the associated Gram matrix  $G$  degenerates. Hence,

$$\det G \rightarrow 0 \quad (3.2)$$

while approaching singular points. Calculating of the entries for the inverse matrix  $G^{-1}$  via minor determinants and cofactors (algebraic adjuncts) of the original matrix  $G$  involves the division by  $\det G$  [5]. Thus, relation (3.2) implies the following result.

**Theorem 3.1.** *The entries  $g^{ij}$  of matrix  $(G^{-1})^\top$  in formula (2.10) calculated by exponential functions (1.2) depend continuously on parameters  $\lambda_1, \dots, \lambda_n$  everywhere except singular points at which the condition  $\lambda_i \neq \lambda_j$  for  $i \neq j$  is violated.*

The calculation of the entries of matrix  $(G^{-1})^\top$  for particular  $n$  shows that the explicit formulae for  $g^{ij}$  involve small denominators vanishing at the singular points, where the condition  $\lambda_i \neq \lambda_j$  for  $i \neq j$  is violated.

4. CLUSTERING OF EXPONENTS IN THE VICINITY OF SINGULAR POINTS

We consider a singular point at which the condition  $\lambda_i \neq \lambda_j$  for  $i \neq j$  is violated. It means that there are repetitions in sequence  $\lambda_1, \dots, \lambda_n$ . Excluding the repetitions, we obtain a smaller sequence  $\Lambda_1, \dots, \Lambda_m$  with no repetitions. In a small neighbourhood of the singular point, numbers  $\lambda_1, \dots, \lambda_n$  are pairwise distinct but they are grouped into clusters attracted by numbers  $\Lambda_1, \dots, \Lambda_m$ . In this situation it is reasonable to renumber  $\lambda_1, \dots, \lambda_n$  writing

$$\lambda_{ij} \quad \text{where} \quad i = 1, \dots, m \quad \text{and} \quad j = 1, \dots, k_i. \tag{4.1}$$

Here  $i$  is the number of a cluster, and  $j$  is an index of the number within the cluster. Numbers  $k_1, \dots, k_m$  in (4.1) are called the multiplicities of clusters. The identity

$$k_1 + \dots + k_m = n \tag{4.2}$$

holds true. Identity (4.2) expresses the preserving of the total amount of exponents under clustering (4.1). Due to (4.1) and (4.2), identity (3.1) can be rewritten as

$$L = \langle \{e^{\lambda_{ij}x}\}_{j=1, \dots, k_i}^{i=1, \dots, m} \rangle \tag{4.3}$$

or as  $L = \langle \{e^{\lambda_{ij}x}\} \rangle$  for the sake of brevity. The passage to a singular point from the surrounding regular ones is expressed as the passage to the limit:

$$\lambda_{ij} \rightarrow \Lambda_i. \tag{4.4}$$

Considering limit (4.4), it is convenient to introduce the small parameter

$$\varepsilon = \max\{|\lambda_{ij} - \Lambda_i|\}_{j=1, \dots, k_i}^{i=1, \dots, m}. \tag{4.5}$$

In view of (4.5), formula (4.4) becomes

$$\lambda_{ij} \rightarrow \Lambda_i \quad \text{as} \quad \varepsilon \rightarrow 0. \tag{4.6}$$

As it was mentioned above, under the passage to the limit (4.6) the exponential functions  $e^{\lambda_{ij}x}$  forming the basis of subspace  $L$  in (4.3) become linearly dependent and do not form a basis anymore. They do not fit Definition 3.1. Our further efforts are focused on proving the existence of another basis in  $L$  satisfying Definition 3.1.

5. TAYLOR EXPANSIONS FOR EXPONENTIAL FUNCTIONS

This section is preliminary. Suppose for a while that we have just one cluster (i.e.,  $m = 1$ ) with  $\Lambda_1 = 0$ . Then we can employ the original notations  $\lambda_1, \dots, \lambda_n$  for the exponents and we can write formula (4.4) as  $\lambda_i \rightarrow 0$ . Exponential functions in (1.2) have the following Taylor expansions at zero:

$$\begin{aligned} e^{\lambda_1 x} &= 1 + \lambda_1 x + \dots + \frac{\lambda_1^{n-1} x^{n-1}}{(n-1)!} + \dots = \sum_{q=0}^{\infty} \frac{\lambda_1^q x^q}{q!}, \\ &\dots \\ e^{\lambda_n x} &= 1 + \lambda_n x + \dots + \frac{\lambda_n^{n-1} x^{n-1}}{(n-1)!} + \dots = \sum_{q=0}^{\infty} \frac{\lambda_n^q x^q}{q!}. \end{aligned} \tag{5.1}$$

Partial sums of power series (5.1) define the polynomials

$$p_i(x) = 1 + \lambda_i x + \dots + \frac{\lambda_i^{n-1} x^{n-1}}{(n-1)!} \quad \text{where} \quad i = 1, \dots, n. \tag{5.2}$$

Employing polynomials (5.2), we introduce the following equations for variables  $\alpha_1, \dots, \alpha_n$  involved in these polynomials:

$$\alpha_1 p_1(x) + \dots + \alpha_n p_n(x) = \frac{x^{n-1}}{(n-1)!}. \tag{5.3}$$

Polynomial equation (5.3) is equivalent to the matrix equation

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \dots & \lambda_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \dots & \lambda_n^{n-1} \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}. \quad (5.4)$$

Matrix (5.4) is obtained by transposing the Vandermonde matrix [6]:

$$W = \begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & \dots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^2 & \dots & \lambda_2^{n-1} \\ 1 & \lambda_3 & \lambda_3^2 & \dots & \lambda_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_n & \lambda_n^2 & \dots & \lambda_n^{n-1} \end{pmatrix}. \quad (5.5)$$

Vandermonde matrix (5.5) is non-degenerate for distinct  $\lambda_i$ , i.e., if  $\lambda_i \neq \lambda_j$  as  $i \neq j$ . In this case the inverse matrix  $U = W^{-1}$  is well-defined [7]. In order to write matrix  $U$  explicitly, we shall make use the following polynomials in  $\lambda$ :

$$P_q(\lambda) = \frac{\prod_{s \neq q}^n (\lambda - \lambda_s)}{\prod_{s \neq q} (\lambda_q - \lambda_s)} \quad \text{where } q = 1, \dots, n. \quad (5.6)$$

It is easy to see that polynomials (5.6) satisfy the relation

$$P_q(\lambda_i) = \begin{cases} 1 & \text{as } q = i, \\ 0 & \text{as } q \neq i. \end{cases} \quad (5.7)$$

If we represent polynomials (5.6) as sums of monomials

$$P_q(\lambda) = \sum_{r=1}^n U_{rq} \lambda^{r-1} = U_{1q} + U_{2q} \lambda + \dots + U_{nq} \lambda^{n-1}, \quad (5.8)$$

then relation (5.7) is written as

$$\sum_{r=1}^n \lambda_i^{r-1} U_{rq} = \begin{cases} 1 & \text{for } q = i, \\ 0 & \text{for } q \neq i. \end{cases} \quad (5.9)$$

In view of the structure of matrix (5.5) one can see that relation (5.9) is equivalent to the matrix identity  $W \cdot U = 1$ , where  $U$  is the matrix formed by the coefficients of polynomials (5.8):

$$U = \begin{pmatrix} U_{11} & U_{12} & U_{13} & \dots & U_{1n} \\ U_{21} & U_{22} & U_{23} & \dots & U_{2n} \\ U_{31} & U_{32} & U_{33} & \dots & U_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ U_{n1} & U_{n2} & U_{n3} & \dots & U_{nn} \end{pmatrix}. \quad (5.10)$$

The explicit formula for matrix (5.10) follows from (5.6):

$$U_{rq} = \frac{1}{(r-1)!} \left. \frac{d^{r-1} P_q(\lambda)}{d\lambda^{r-1}} \right|_{\lambda=0}. \quad (5.11)$$

The identity  $W \cdot U = 1$  implied by (5.9) means that matrix (5.10) with entries (5.11) is inverse for the Vandermonde matrix.

Since  $(W^\top)^{-1} = U^\top$ , we can make use of the transposed matrix  $U^\top$  to solve matrix equation (5.4):

$$\alpha_q = U_{nq} = \frac{1}{(n-1)!} \left. \frac{d^{n-1} P_q(\lambda)}{d\lambda^{n-1}} \right|_{\lambda=0} = \frac{1}{\prod_{s \neq q}^n (\lambda_q - \lambda_s)}. \quad (5.12)$$

Along with solving matrix equation (5.4), quantities (5.12) solve also polynomial equation (5.3).

Employing quantities (5.12) determined as coefficients, we introduce the following linear combination of exponential functions (1.2):

$$f_n(x) = \alpha_1 e^{\lambda_1 x} + \dots + \alpha_n e^{\lambda_n x}. \quad (5.13)$$

If we take into consideration (5.1), (5.2), and (5.3), function  $f_n(x)$  in (5.13) can be represented as the power series

$$f_n(x) = \frac{x^{n-1}}{(n-1)!} + \sum_{q=1}^{\infty} \frac{B_{nq} x^{n-1+q}}{(n-1+q)!}. \quad (5.14)$$

Coefficients  $B_{nq}$  in (5.14) are determined by the formulae

$$B_{nq} = \sum_{i=1}^n \alpha_i \lambda_i^{n-1+q} = \sum_{i=1}^n \frac{\lambda_i^{n-1+q}}{\prod_{s \neq i}^n (\lambda_i - \lambda_s)}. \quad (5.15)$$

Formally speaking, power series (5.14) involves quantities (5.15) with  $q \geq 1$ . But formula (5.15) can be extended for the case  $q = 0$ . In this case, substituting  $q = 0$  into (5.15) and using matrix identity (5.4) for quantities  $\alpha_1, \dots, \alpha_n$ , we obtain the identity

$$B_{n0} = 1 \quad \text{for each } n \geq 1. \quad (5.16)$$

**Lemma 5.1.** *For arbitrarily many  $\lambda_1, \dots, \lambda_n, \lambda_{n+1}, \dots$ , pairwise distinct numbers  $\lambda_i \neq \lambda_j$  and  $q \geq 1$ , quantities  $B_{nq}$  in (5.15) satisfy the recurrent identity*

$$B_{n+1q} = B_{nq} + \lambda_{n+1} B_{n+1q-1}. \quad (5.17)$$

Lemma 5.1 can be easily proved by straightforward calculations based on formula (5.15). As  $n = 1$ , formula (5.12) becomes  $\alpha_1 = 1$ . Then (5.15) yields

$$B_{1q} = \lambda_1^q \quad \text{for each } q \geq 0. \quad (5.18)$$

Formulae (5.16) and (5.18) and recurrent identities (5.17) are sufficient to determine all  $B_{nq}$  by induction w.r.t. two parameters  $n$  and  $q$ .

**Lemma 5.2.** *Quantities  $B_{nq}$  determined by pairwise distinct numbers  $\lambda_i \neq \lambda_j$  via formula (5.15) are calculated explicitly:*

$$B_{nq} = \sum_{1 \leq i_1 \leq \dots \leq i_q \leq n} \lambda_{i_1} \cdot \dots \cdot \lambda_{i_q} \quad \text{where } n \geq 1 \quad \text{and } q \geq 1. \quad (5.19)$$

In order to prove Lemma 5.2 it is sufficient to make sure that identities (5.17) and (5.18) are fulfilled once we substitute (5.19). After that formulae (5.19) and (5.16) determine all  $B_{nq}$  explicitly.

We denote by  $N_{nq}$  the number of terms in formula (5.19) for  $B_{nq}$ . This number is estimated by the following inequality:

$$N_{nq} \leq n^q. \quad (5.20)$$

Note that in this section we consider the special case, where the number of clusters is  $m = 1$ , the multiplicity is  $k_1 = n$ , and  $\Lambda_1 = 0$ . In view of this, we let

$$\varepsilon = \max(|\lambda_1|, \dots, |\lambda_n|). \quad (5.21)$$

Notation (5.21) is a version of notation (4.5) applied to the special case. We apply (5.20) and (5.21) to estimate the terms in power series (5.14). For  $q \geq 1$  we obtain

$$\left| \frac{B_{nq} x^{n-1+q}}{(n-1+q)!} \right| \leq \frac{n^q \varepsilon^q |x|^{n-1+q}}{(n-1)! n (n+1) \cdot \dots \cdot (n+q-1)}. \quad (5.22)$$

For our purposes it is sufficient to have a weaker estimate. It follows from (5.22) that

$$\left| \frac{B_{nq} x^{n-1+q}}{(n-1+q)!} \right| \leq \frac{\varepsilon^q |x|^{n-1+q}}{(n-1)!} \quad \text{where } q \geq 1. \quad (5.23)$$

It implies easily the estimate in the norm of  $L^2([- \pi, + \pi])$ :

$$\left\| \frac{B_{nq} x^{n-1+q}}{(n-1+q)!} \right\| \leq \sqrt{\frac{2\pi}{2n+2q-1}} \frac{\pi^{n-1}}{(n-1)!} (\pi \varepsilon)^q. \quad (5.24)$$

Since  $n = \text{const}$  in (5.1), (5.14), and (5.24), estimate (5.24) can be simplified. In order to do it, we introduce the following constant independent of  $q \geq 1$ :

$$C_n = \sqrt{\frac{2\pi}{2n+1}} \frac{\pi^{n-1}}{(n-1)!}. \quad (5.25)$$

Employing constant (5.25), estimate (5.24) can be rewritten as

$$\left\| \frac{B_{nq} x^{n-1+q}}{(n-1+q)!} \right\| \leq C_n (\pi \varepsilon)^q \quad \text{where } q \geq 1. \quad (5.26)$$

For  $\pi \varepsilon < 1/2$ , estimate (5.26) produces an estimate for function  $f_n(x)$  defined by formula (5.13) and represented by series (5.14):

$$\left\| f_n(x) - \frac{x^{n-1}}{(n-1)!} \right\| \leq \frac{C_n \pi \varepsilon}{1 - \pi \varepsilon} \leq 2 C_n \pi \varepsilon. \quad (5.27)$$

Due to formula (5.13), function  $f_n(x)$  is a linear combination of exponential functions  $e^{\lambda_1 x}, \dots, e^{\lambda_n x}$ , i.e., it belongs to subspace  $L$  in (3.1). Hence, we can formulate the following theorem.

**Theorem 5.1.** *For any  $n$  pairwise distinct complex quantities  $\lambda_1, \dots, \lambda_n$  tending to zero we can construct a complex-valued function  $f(x)$  belonging to subspace  $L$  generated by the exponential functions  $e^{\lambda_1 x}, \dots, e^{\lambda_n x}$  in Hilbert space  $H = L^2([- \pi, + \pi])$  of square integrable functions such that*

$$\|f(x) - x^{n-1}\| \rightarrow 0, \quad (5.28)$$

as  $\lambda_1, \dots, \lambda_n$  tend to zero.

Theorem 5.1 follows immediately from inequality (5.27). It is important to stress that the norm convergence in (5.28) depends neither on any relations for  $\lambda_1, \dots, \lambda_n$  no on individual convergence rates for particular  $\lambda_i \rightarrow 0$ .

Given  $n$  pairwise distinct complex parameters  $\lambda_1, \dots, \lambda_n$  tending to zero, we can select part of them  $\lambda_1, \dots, \lambda_s$ , where  $1 \leq s \leq n$ . Applying Theorem 5.1 to each such sample, we obtain the following result.



**Theorem 5.2.** *For any  $n$  pairwise distinct complex  $\lambda_1, \dots, \lambda_n$  tending to zero, we can construct  $n$  complex-valued functions  $f_1(x), \dots, f_n(x)$  belonging to subspace  $L$  generated by the exponential functions  $e^{\lambda_1 x}, \dots, e^{\lambda_n x}$  in Hilbert space of square integrable functions  $H = L^2([-\pi, +\pi])$  such that*

$$\|f_s(x) - x^{s-1}\| \rightarrow 0, \quad s = 1, \dots, n, \tag{5.29}$$

as  $\lambda_1, \dots, \lambda_n$  tend to zero.

Again, it is important to stress that the norm convergence (5.29) depends neither on relations between  $\lambda_1, \dots, \lambda_n$  nor on individual rate of convergence for particular  $\lambda_i \rightarrow 0$ .

Comparing Theorem 5.2 with Definition 3.1 of the convergence of subspace, we arrive immediately at the following result.

**Theorem 5.3.** *For any  $n$  pairwise distinct complex quantities  $\lambda_1, \dots, \lambda_n$  tending to zero, the linear span of exponential functions*

$$L = \langle e^{\lambda_1 x}, \dots, e^{\lambda_n x} \rangle$$

in Hilbert space  $H = L^2([-\pi, +\pi])$  of square integrable functions converges to the linear span of polynomials

$$M = \langle 1, x, \dots, x^{n-1} \rangle$$

as  $\lambda_1, \dots, \lambda_n$  tend to zero.

## 6. CASE OF SEVERAL CLUSTERS

We proceed to studying general singular points in the vicinity of which the exponents  $\lambda_1, \dots, \lambda_n$  are grouped into  $m$  clusters with multiplicities  $k_1, \dots, k_m$  (see (4.1)). They converge to  $m$  pairwise distinct complex numbers  $\Lambda_1, \dots, \Lambda_m$  in accordance with (4.4) and (4.6). This is why we introduce small deflections

$$\theta_{ij} = \lambda_{ij} - \Lambda_i$$

tending to zero and instead of (5.1) we write

$$\begin{aligned} e^{\lambda_{i1}x} &= e^{\Lambda_i x} \left( 1 + \theta_{i1} x + \dots + \frac{\theta_{i1}^{n-1} x^{n-1}}{(n-1)!} \right) + \dots = \sum_{q=1}^{\infty} \frac{\theta_{i1}^q x^q e^{\Lambda_i x}}{q!}, \\ &\dots\dots\dots \\ e^{\lambda_{ik_i}x} &= e^{\Lambda_i x} \left( 1 + \theta_{ik_i} x + \dots + \frac{\theta_{ik_i}^{n-1} x^{n-1}}{(n-1)!} \right) + \dots = \sum_{q=1}^{\infty} \frac{\theta_{ik_i}^q x^q e^{\Lambda_i x}}{q!}. \end{aligned}$$

As a result, instead of Theorem 5.1 we obtain the following theorem.

**Theorem 6.1.** *For a set of pairwise distinct complex ingredients  $\lambda_{ij}, i = 1, \dots, m, j = 1, \dots, k_i$ , tending to  $m$  pairwise distinct complex numbers  $\Lambda_1, \dots, \Lambda_m$  there exists a function  $f_s(x)$  belonging to subspace  $L = \langle \{e^{\lambda_{ij}x}\}_{j=1, \dots, k_i}^{i=1, \dots, m} \rangle$  in Hilbert space of square integrable functions  $\mathcal{H} = L^2([-\pi, +\pi])$  such that*

$$\|f_s(x) - x^{k_s-1} e^{\Lambda_s x}\| \rightarrow 0 \quad \text{as } \lambda_{ij} \rightarrow \Lambda_i.$$

The proof of Theorem 6.1 follows the same lines as the proof of Theorem 5.1. The only difference is that instead of polynomial  $x^{n-1}$  here we deal with an exponential polynomial<sup>1</sup>. The exponential polynomial  $x^{k_s-1} e^{\Lambda_s x}$  is not unique in Theorem 6.1. Reproducing the arguments in deriving Theorem 5.2 from Theorem 5.1, we can get as many exponential polynomials as many exponential functions we initially have.

<sup>1</sup>In Russian mathematical literature the “квазиполином” is widely used for exponential polynomial, which is translated literally as “quasipolynomial”. It denotes expressions  $c x^q e^{\lambda x}$  with positive integer exponent  $q$  and

**Theorem 6.2.** For a set of pairwise distinct complex parameters  $\lambda_{ij}$ ,  $i = 1, \dots, m$ , and  $j = 1, \dots, k_i$ , tending to  $m$  pairwise distinct complex numbers  $\Lambda_1, \dots, \Lambda_m$ , there exist the set of functions  $f_{ij}(x)$ ,  $i = 1, \dots, m$ ,  $j = 1, \dots, k_i$ , belonging to subspace  $L = \langle \{e^{\lambda_{ij}x}\}_{j=1, \dots, k_i}^{i=1, \dots, m} \rangle$  in Hilbert space  $\mathcal{H} = L^2([-\pi, +\pi])$  such that

$$\|f_{ij}(x) - x^{j-1} e^{\Lambda_i x}\| \rightarrow 0 \quad \text{as} \quad \lambda_{ij} \rightarrow \Lambda_i.$$

An analogue of Theorem 5.3 reads as follows.

**Theorem 6.3.** For a set of pairwise distinct complex parameters  $\lambda_{ij}$ ,  $i = 1, \dots, m$ ,  $j = 1, \dots, k_i$ , tending to  $m$  pairwise distinct complex numbers  $\Lambda_1, \dots, \Lambda_m$ , the linear span of exponential functions

$$L = \langle \{e^{\lambda_{ij}x}\}_{j=1, \dots, k_i}^{i=1, \dots, m} \rangle$$

in Hilbert space  $H = L^2([-\pi, +\pi])$  of square integrable functions, as  $\lambda_{ij} \rightarrow \Lambda_i$ , converges to the linear span of exponential polynomials

$$M = \langle \{x^{j-1} e^{\Lambda_i x}\}_{j=1, \dots, k_i}^{i=1, \dots, m} \rangle. \quad (6.1)$$

## 7. REMOVING SINGULARITIES

We return back to the solution of linear approximation problem (1.6) represented by formula (2.10). As it was mentioned above in Section 3, the geometric interpretation of function  $\Phi(\lambda_1, \dots, \lambda_n)$  is the  $L^2$ -distance from function  $f(x) \in L^2([-\pi, +\pi])$  to  $n$ -dimensional subspace  $L$  in (3.1). It means that  $\Phi(\lambda_1, \dots, \lambda_n)$  depends on  $L$  but is independent of a particular basis in subspace  $L$  employed in formula (2.10). On the other hand, due to formula (2.10), quantity  $\Phi(\lambda_1, \dots, \lambda_n)$  depends continuously on the basis provided this basis changes together with subspace  $L$  not stopping to be a basis. It implies the following result.

**Theorem 7.1.** Function  $\Phi(\lambda_1, \dots, \lambda_n)$  being a solution to linear approximation problem (1.6) is continuous in regular points, where the condition  $\lambda_i \neq \lambda_j$  as  $i \neq j$  holds true, and has a finite limit if this condition is violated.

To prove Theorem 7.1, it is sufficient to replace basis of exponential functions (1.2) in  $L$  by the basis of functions  $f_{ij}(x)$  whose existence is ensured by Theorem 6.2. To calculate the limiting value of function  $\Phi(\lambda_1, \dots, \lambda_n)$  at a singular point determined by numbers  $\Lambda_1, \dots, \Lambda_m$  and multiplicities  $k_1, \dots, k_m$ , we replace  $L$  by the limiting subspace  $M$  in (6.1) and employ basis of exponential polynomials from formula (2.10).

Theorem 7.1 means that function  $\Phi(\lambda_1, \dots, \lambda_n)$  can be extended to a function continuous in the whole space  $\mathbb{C}^n$ . This fact is important in composing numerical algorithms aimed on solving nonlinear approximation problem and Problem 1.1. Problem 1.1 should be treated in generalized sense including together with the minimization over exponential functions the minimization over exponential polynomials.

## 8. DISCUSSIONS AND CONCLUSIONS

The present work was discussed on November 19, 2014 in Bashkir State University in the seminar which now has the status of the city seminar on theory of functions named after A.F. Leont'ev. During the discussion N.F. Valeev expressed the idea that theorems like 5.1, 5.2, 5.3 and 6.1, 6.2, 6.3 can be proved in the framework of the theory of linear ordinary differential equations. R.S. Yulmukhametov conjectured that results similar to these theorems can be contained in earlier works of 1960 not accessible via the Internet. After two days he sent to the author the e-mail with a sophisticated and shorter proof of Theorem 5.1 found by himself.

---

various sums of such expressions. However, in English mathematical literature the term "quasipolynomial" has a different meaning [8].

In view of these circumstances, the author does not pretend to the priority of Theorems 5.1, 5.2, 5.3 and 6.1, 6.2, 6.3, but he provides them with the proofs for the completeness and because these proofs are rather constructive and can be employed for quantitative estimates in composing numerical algorithms. The main result of the work is Theorem 7.1 having a direct influence on solving Problem 1.1 numerically.

Along with a practical application, Problem 1.1 requires a further theoretical study. The author thinks that it has a certain potential for generating new results and deserves a popularization.

#### 9. ADDENDUM UPON THE FIRST REFEREEING AND DISCUSSION IN A CONFERENCE

On December 2–5, 2014, the second all-Russian scientific youth conference “Topical problems of nano- and microelectronics” was held in Bashkir State University. It was organized by R.Z. Bakhtizin. The author presented his report in the section “Digital processing of information and automation of measurements in nano- and microelectronics”. During the report the original applied problem and the algorithm of its solving were discussed. The referee also emphasizes the applied problem and points to the paper [9], where the nonhomogeneous discrete Fourier transform (NDFT) and its application to creating digital frequency filters were considered. Some of above formulae can be found in [9]. For instance, it concerns formula (5.5) for Vandermonde matrix and formula (5.6) for interpolation polynomial. In some other papers in book [9] quadratic  $L^2$ -norms and root-mean-square deviations were considered. However, in book [9] we found no formulae being in direct correspondance with the above formulae in Sections 2-4. Employing of wavelets, which were mentioned by the referee, means to replace exponential functions by more complicated functions [10].

Many of the papers in book [9] use somehow Nyquist-Shannon sampling theorem [11]. The author learned on this theorem from R.Z. Bakhtizin. It describes the reconstruction of a signal with a bounded spectrum by a discrete sample in an infinite sequence of points. Signals in the applied problem from the group of companies “PhysTech” are assumed to have a bounded spectrum. But Nyquist-Shannon sampling theorem is not used in the present work. It can be required in future while increasing the number of spectral points  $\lambda_1, \dots, \lambda_n$ .

The choice of optimal values  $\lambda_1, \dots, \lambda_n$  in the applied problem from the group of companies “PhysTech” resembles choosing of frequencies for NDFT filters. However, there is an important difference. The choice of the frequencies for filters is made once during their design. In our case the frequencies  $\lambda_1, \dots, \lambda_n$  should be determined on-fly in processing each incoming signal.

#### 10. ACKNOWLEDGMENTS

The author is grateful to N.F. Valeev, R.S. Yulmukhametov and all participants of the seminar on the theory of function for a fruitful discussion of the work. The author is grateful to B.N. Khabibullin with whom the author discussed Problem (1.1) about one year ago and who said that this problem is of certain interest. The author is grateful to R.Z. Bakhtizin for the invitation to take part in the organized conference, to S.S. Gotz, and to the participants of the section on digital processing of signals for useful discussion of the work during the conference and afterward. The author also thanks the referee of Ufa Mathematical Journal for the information on book [9] and the useful comments.

#### BIBLIOGRAPHY

1. *Least squares*. Wikipedia, Wikimedia Found. Inc., San Francisco, USA.
2. A.F. Leontiev. *Exponential series*. Nauka, Moscow (1976). (in Russian).
3. R.A. Sharipov. *On root mean square approximation by exponential functions* // Preprint: arXiv:1411.2467.

4. F. Chatelin. *Eigenvalues of matrices*. Classics in Applied Mathematics. **71**. SIAM, Philadelphia (2013).
5. A.G. Kurosh. *A course in higher algebra*. Nauka, Moscow (1968). (in Russian).
6. *Vandermonde matrix*. Wikipedia, Wikimedia Found. Inc., San Francisco, USA.
7. L.R. Turner. *Inverse of the Vandermonde matrix with applications*. NASA Technical Note D-3547, NASA, Washington (1966).
8. *Quasi-polynomial*. Wikipedia, Wikimedia Found. Inc., San Francisco, USA.
9. S. Bagchi, S.K. Mitra. *The nonuniform discrete Fourier transform* // in book “Nonuniform sampling: theory and practice”, F. Marvasti, ed., Kluwer Academic/Plenum Publishers, New York. 325–360 (2001).
10. V.A Baikov, A.V. Zhiber, R.D. Murtazina. *Foundation of wavelets theory*. Bashkir State University, Ufa (2012). (in Russian).
11. *Nyquist-Shannon sampling theorem*. Wikipedia, Wikimedia Found. Inc., San Francisco, USA.

Ruslan Abdulovich Sharipov,  
Bashkir State University,  
Zaki Validi str. 32,  
450074, Ufa, Russia  
E-mail: r-sharipov@mail.ru