Varin V.P.

Methods without saturation for boundary value problems

Recommended form of bibliographic references: Varin V.P. Methods without saturation for boundary value problems. Keldysh Institute preprints, 2008, No. 1, 13 p. URL: http://library.keldysh.ru/preprint.asp?id=2008-1&lg=e

РОССИЙСКАЯ АКАДЕМИЯ НАУК
ОРДЕНА ЛЕНИНА ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
ИМЕНИ М.В. КЕЛДЫША

V.P. Varin

METHODS WITHOUT SATURATION FOR BOUNDARY VALUE PROBLEMS

Moscow, 2008.

УДК 521.1+531.314

V.P. Varin Methods without saturation for boundary value problems. Preprint of the Keldysh Institute of Applied Mathematics of RAS, Moscow, 2008.

We consider a class of numerical methods without saturation for boundary value problems which is oriented for computations with an arbitrary controlled precision. These methods differ from the traditional ones in the feedback which they have with the problem, when the results of computations allow to draw conclusions on analytical properties of solutions to the problem. These methods are little-known, and they are not implemented in existing computer algebra systems, although these systems made their practical realization possible.

В.П. Варин. Методы без насыщения для краевых задач. им. М.В. Келдыша РАН, Москва, 2008.

Рассматривается класс численных методов без насыщения для краевых задач, ориентированный на вычисления с произвольной контролируемой точностью. Отличие этих методов от традиционных состоит в наличии обратной связи с задачей, когда по результатам вычислений можно сделать выводы об аналитических свойствах решения задачи. Эти методы малоизвестны и не применяются в существующих системах компьютерной алгебры, хотя она сделала их практическую реализацию возможной.

Работа поддержана Российским фондом фундаментальных исследований, гранты 08-01-00082 и 08-01-00251, и программой Президиума РАН "Фундаментальные проблемы нелинейной динамики".

E-mail: varin@keldysh.ru
http: www.keldysh.ru
(see electronic library/ catalogue of publications/ preprints).

§ 1. Introduction

In this paper we consider generic linear boundary value problems (BVP)

$$L(x, y(x)) = R(x), \quad x \in [-1, 1] \tag{1}$$

with arbitrary boundary conditions, where $L(x, y(x))$ is a differential operator and $R(x)$ is the right hand side. Note that initial value problems for (1) fall in this category. We assume that all coefficients in L and the function $R(x)$ are analytical functions in some ellipse with foci at ± 1.

Numerical methods of solution of BVP (1) may be loosely split into two groups. The first group uses some difference approximation or numerical integration in some form over the interval. It is equivalent to writing the BVP (1) in a finite dimensional form $Ly = R$, where L is a band matrix of high order and R is a vector. The boundary conditions are accounted for in upper and lower rows of the matrix L.

The second group uses expansions of solutions to a BVP in series of some special functions usually related to the problem. Here we obtain algebraic linear problems with dense matrices of lower order.

The methods of the first group are fairly developed. They are implemented in computer algebra packages and widely used for solution of typical BVPs. The methods of the second group require preliminary analysis of the BVP, and, consequently, they can not be implemented without a prior knowledge about the problem.

All the methods that use difference approximation or numerical integration have one significant drawback. They can give very little information on analytical properties of a solution to a BVP. A simple example is a Runge-Kutta method of, say, fourth order. It has the local error estimate $O(h^5)$, where h is a stepsize. If BVP (1) has a solution $y(x)$ in the form of a polynomial of up to fourth order, then this Runge-Kutta method gives the exact solution. So the method knows analytical properties of a solution to some extent. But the method would not distinguish between polynomial solutions of higher order regardless of the stepsize. This property of numerical algorithms is called *saturation* [1].

The second group consists mostly of variations on Galerkin method. The term saturation here is not applicable, since Galerkin methods imply average convergence. But there are some very important exceptions.

Some BVPs require expansions of their solutions in specific special functions. It may happen that these expansions provide methods without saturation for these BVPs, but it must be proved in each case.

Consider expansion in Fourier series of a periodic solution to a BVP with periodic boundary conditions. As we know, the order in which coefficients of a

Fourier series decrease reflects the order of smoothness of the function which we expand. Besides, under the assumptions we made on the BVP, Fourier series will converge uniformly to the solution if it is finite. So we have a feedback on analytical properties of a solution to a BVP from results of numerical experiment. This feedback is made possible here by the fact that the error estimate of a solution depends, among other things, on the order of smoothness of the solution. Numerical methods with this property are called *methods without saturation* [1].

It is well known that expansions in Fourier series give the best finite dimensional approximations to periodic functions. It corresponds to the fact that the optimal distribution of the nodes of interpolation for periodic functions is uniform [1-2]. Further we will not consider periodic boundary conditions. So why we think that numerical methods without saturation exist for BVPs (1)?

The differential operator $L(x, y(x))$ together with the boundary conditions defines some closed operator $\mathcal{L}$ in some functional Banach space $\mathcal{F}$. The inverse operator $\mathcal{L}^{-1}$ is a compact operator in $\mathcal{F}$. It can be written in integral form with the help of the Green function. A compact operator can be written as a sum of a finite rank operator and an operator of arbitrary small norm. In other words, a compact operator has a good finite dimensional approximation. So we have to find this finite dimensional approximation and do it, if we can, in the best possible way.

It is hardly possible to suggest a numerical method without saturation for a BVP (1) in such a general statement. So we will not attempt this task. Instead, in this paper we describe some building blocks, which allow to construct a numerical method for such a BVP semi-automatically using symbolic computations. It is possible then to solve a BVP numerically and verify that the method is indeed without saturation. In this case we have a guaranteed controlled accuracy of the solution.

§ 2. Optimal approximation with Chebyshev polynomials

The problem of choosing the best numerical method for a BVP (1) is closely related to the problem of finding the best finite dimensional approximations to the solution $y(x)$. This problem is already solved for BVPs with periodic boundary conditions and periodic solutions. Let Π_n be the projection operator from the space of trigonometric polynomial $P_n(x)$ of the order n onto the space of analytical periodic functions. The norm of this operator is called the Lebesgue constant. It does not depend on the interval on which we approximate the function and depends only on the nodes of approximation. It is known that the uniform distribution of the nodes minimizes the Lebesgue constant for periodic functions [1- 2].

But if we apriori know nothing about the nature of the solution $y(x)$ to the BVP (1), then the best modus operandi is to approximate the solution with Chebyshev polynomials

$$y(x) \approx \sum_{k=0}^{m} a_k T_k(x), \quad T_n(x) = \cos(n \arccos x) \tag{2}$$

where m is the order of the approximation. Let us substantiate this statement.

First, for an arbitrary analytical function on the interval $[-1, 1]$, the optimal distribution of the nodes for its polynomial approximation is unknown. But it is known that Chebyshev nodes are close to the optimal distribution, and asymptotically they are optimal [1-2]. This fact indicates that the Lebesgue constant will be close to its minimal value, and that means that finite dimensional approximations to the operator $L(x, y(x))$ will have minimal dimensions for a given error of approximation.

Chebyshev polynomials are exceptional among all other systems of orthogonal polynomial. Their roots are known explicitly. These roots can be chosen for the nodes of approximation. For other polynomials, we would have to compute the roots numerically.

If a function is analytical in some ellipse with foci at ± 1, then it has an expansion in series of Chebyshev polynomials. Coefficients of these series will decrease exponentially in direct analogy with the Fourier series. These coefficients provide an accuracy control of the numerical solution, which is the feedback of numerical method on the problem that we discussed in previous section. Note that this follows from the uniform convergence of Chebyshev series and the minimal deviation of Chebyshev polynomials from zero on the interval $[-1, 1]$. Galerkin methods do not have this advantage due to their average convergence.

Finally, we need not work with coefficients of Chebyshev series. We can work with values of a function at the Chebyshev nodes and obtain the coefficients of Chebyshev series by fast Fourier-Chebyshev transform. Since ± 1 do not belong to Chebyshev nodes, we can solve BVPs that are close to singular or have some peculiar boundary conditions.

§ 3. Working with boundary conditions

There are mostly two ways to account for the boundary conditions when we use Chebyshev approximation for solution of the BVP (1). The first one uses a direct approach when a solution to the BVP is expanded in Chebyshev series (2). Coefficients in the differential operator are also expanded as well as all known functions taking part in BVP (1). When these series are differentiated and multiplied, some recurrent relations appear for the unknown coefficients in the expansion of solution $y(x)$. This infinite system of equations is truncated to a

finite order. The boundary conditions are used to derive additional equations for unknown coefficients that make the system closed and hopefully solvable. This approach was developed in [3].

Although this approach provides a method without saturation if applicable, in practice, it is very time and effort consuming, and it cannot be pre-programmed or automated. Indeed, if we change some coefficients in the differential operator, then the recurrent relations on coefficients will also change.

Another approach is to transform the BVP (1) into an integral equation with all boundary conditions included using the Green function. If it is done, then we can approximate the solution in the following form

$$y(x) = \sum_{i=1}^{n} y(x_i) p_{ni}(x), \quad x_i = \cos \frac{(2i-1)\pi}{2n}, \quad i = 1, \ldots, n \tag{3}$$

where x_i are Chebyshev roots, and

$$p_{ni}(x) = \frac{T_n(x)}{(x - x_i) T_n'(x_i)}, \quad i = 1, \ldots, n \tag{4}$$

are Chebyshev fundamental polynomials of Lagrange interpolation. Then we have a linear system of equations for the unknowns $y(x_1), \ldots, y(x_n)$. The values of the function $y(x)$ at other points in $[-1, 1]$ can be obtained by the formula (3) with very high accuracy. And the accuracy itself is controlled by coefficients of Chebyshev series (2).

This approach requires a deep insight into the problem, and, obviously, cannot be pre-programmed.

So we see that although Chebyshev polynomials are highly effective and provide exceptionally good numerical methods of solution of BVPs, their implementation is a serious problem in itself. Until recently, this problem was always treated on individual basis by the researcher, since it is inextricable from the BVP and its properties. Or is it?

In some of his last works, K.I. Babenko suggested a means that allows to approximate a differential operator with boundary conditions in such a way that it is equivalent to approximation of the corresponding integral operator [1]. It best be shown on an example.

Consider the BVP $f'(x) = g(x)$, $-1 \le x \le 1$, $f(-1) = 0$. It is, obviously, a Cauchy problem. If we use Lagrange interpolation (3) for the unknown function $f(x)$, then we obtain a finite dimensional linear problem $Af = g$, where f and g are now vectors, and A is the matrix of approximation. However, the matrix A is degenerate and the problem is unsolved. Let us rewrite Cauchy problem in integral form

$$f(x) = \int_{-1}^{1} G(x, y) g(y) dy, \tag{5}$$

where

$$G(x,y)=\begin{cases}1, & \text{if } y\le x\\ 0, & \text{if } y>x\end{cases}$$

is the Green function. Now we can approximate the functions in (5) using Lagrange interpolation (3), and we obtain a finite dimensional linear problem $f=Bg$, i.e. the problem is solved. The matrix B is non-singular, since the function g is found uniquely from the integral equation (5).

Let us return to Cauchy problem and use Lagrange interpolation (3) but with different fundamental polynomials of interpolation

$$f(x)=\sum_{i=1}^{n} f(x_i)q_{ni}(x),\quad q_{ni}(x)=\frac{1+x}{1+x_i}\,p_{ni}(x),\quad i=1,\dots,n. \tag{6}$$

Fundamental polynomials $q_{ni}(x)$, and, consequently, the function $f(x)$ in (6), satisfy the boundary condition. For the function g, we use interpolation (3) as before.

We obtain a finite dimensional linear problem $Cf=g$, but now the matrix C is non-singular. In fact, $C=B^{-1}$. Indeed, let $g\in\mathcal{P}_n$, where $\mathcal{P}_n$ is the n-dimensional linear space of polynomials whose degree is no greater than $n-1$. Then $f\in\mathcal{P}_{n+1}$ is given by the formula (5), which is written as $f=Bg$ for the values of polynomials at the nodes. Conversely, the polynomial $f\in\mathcal{P}_{n+1}$ is restored uniquely from the values $f(x_1),\dots,f(x_n)$ by the formula (6); it satisfies the boundary condition $f(-1)=0$ and the equation (5). Consequently, $f'=g$, which is written as $Cf=g$ for the values at the nodes.

So, instead of rewriting the problem as an integral equation and then using Chebyshev approximation (3), we can approximate the differential operator and then invert a matrix. In fact, even in this simple example, the matrix B is not easily computed, whereas the matrix C is easily found:

$$C_{i,j}=\frac{(-1)^{i-j}(1+x_i)\sqrt{\dfrac{1-x_j^2}{1-x_i^2}}}{(x_i-x_j)(1+x_j)},\ i\ne j,\quad C_{i,i}=\frac{2-x_i}{2(1-x_i^2)}. \tag{7}$$

To conclude this example, we note that the matrix C is quite useful. It provides Gauss quadrature formulas for definite integration over the interval $[-1,1]$. We recall, that Gauss quadrature formulas of a certain order are quadratures that are exact on polynomials of this order. So Gauss quadratures as a whole give another example of numerical method without saturation.

§ 4. Building blocks for numerical method

The principle of approximating differential instead of integral operators can be applied to differential operators d^m/dx^m with m simple conditions imposed

on the values of the function and its derivatives at the points ± 1. The boundary conditions are *correct*, if for any $g \in \mathcal{P}_n$, the BVP

$$\frac{d^m f}{dx^m} = g \tag{8}$$

is solvable, and $f \in \mathcal{P}_{n+m}$. The correctness of boundary conditions means that the linear system of equations for coefficients of the polynomial f is non-degenerate.

The fundamental polynomials of Chebyshev interpolation for the left hand side of (8) can be written as

$$q_{ni}(x) = \frac{r_{mi}(x)}{r_{mi}(x_i)}\, p_{ni}(x), \quad i = 1, \ldots, n, \tag{9}$$

where polynomials $p_{ni}(x)$ are defined in (4), and $r_{mi}(x)$ are polynomials of degree m such that fundamental polynomials $q_{ni}(x)$ satisfy the boundary conditions. For the function g, we use interpolation (3) as before. If boundary conditions are correct, then each polynomial $r_{mi}(x)$ is defined uniquely modulus constant factor, and we obtain the matrix Q of a finite dimensional approximation to the BVP (8) such that Q^{-1} is the matrix of approximation to the corresponding BVP in integral form. It is proved in the same way as for the BVP (5). The matrix Q is computed by the formula

$$Q_{k,i} = \frac{d^m q_{ni}}{dx^m}(x_k), \qquad k, i = 1, \ldots, n, \tag{10}$$

although some indeterminate forms would need to be evaluated for $k = i$ in (10).

In fact, interpolation (3) with fundamental polynomials (9) is Hermite interpolation. We recall that Hermite polynomials of interpolation are Lagrange polynomials but with some additional conditions imposed on the function and/or its derivatives at some, may be additional, nodes.

We just described one of the building blocks that can be used for numerical solution of BVPs (1). These blocks are realized as small computer algebra programs that take boundary conditions and the dimension n as an input and produce matrices of finite dimensional Chebyshev approximation for differential operators (8). For each set of boundary conditions, we need this program written only once. For some boundary conditions it is fairly simple, as for the matrix C in previous section; other boundary conditions require very bulky symbolic computations.

These building blocks can be stored and used to model other boundary conditions without additional computations. For example, fundamental polynomials for the BVP

$$\frac{d^2 f}{dx^2} = g \tag{11}$$

with zero boundary conditions are written as

$$q_{ni}(x) = \frac{1-x^2}{1-x_i^2}\, p_{ni}(x), \quad i = 1, \ldots, n.$$

Let M be the corresponding matrix of Chebyshev approximation for the BVP (11). It is easy to see that the matrix M^2 gives Chebyshev approximation for the BVP

$$\frac{d^4 f}{dx^4} = g \tag{12}$$

with the boundary conditions $f(-1) = f(1) = f''(-1) = f''(1) = 0$.

Now we write the BVP (1) in the form

$$\sum_{k=0}^{m} a_k(x)\, \frac{d^k y}{dx^k} = R(x), \quad x \in [-1, 1] \tag{13}$$

and we use appropriate matrix of finite dimensional approximation (10) for each differential operator $d^k y/dx^k$ in (13) taking as many boundary conditions for each block as it needs. Obviously, only the block of the highest order m in (13) is determined uniquely. For other blocks we have a certain flexibility as to which k of the m boundary conditions to take. The coefficients $a_k(x)$ in (13) are written as diagonal matrices.

So we have obtained a linear algebraic problem that models a given BVP. Potentially, it may be closest to the best possible finite dimensional approximation meaning the lowest dimension and minimal global error estimate. However the real power of this method is unfolded when we use it dynamically, increasing (or decreasing) the dimension, if we need, after the analysis of Fourier-Chebyshev coefficients (2) of the solution.

Another advantage over the finite-difference approach is the possibility of solution of singular problems (or very close to such). The stability or instability of a solution here is not an issue. The stiffness of the problem is also meaningless. It may turn out that a solution to such a problem has a good polynomial approximation, but we may never know it with a finite-difference method.

§ 5. An example of application

We consider a model of human cochlea which is described by the following equation (A.G. Petrov, a private communication):

$$\frac{d^2}{dx^2}\left(\Sigma(x)\frac{d^2}{dx^2}H\,(x) + \alpha^2 M(x)\Omega^2 H(x)\right) + 2\left(iB^3(x)\Omega\,\beta - B(x)\Omega^2\right)H\,(x) = 0. \tag{14}$$

Here the complex function $H(x)$ is on the interval $[0, 1]$ and satisfies the following boundary conditions

$$H(0) = H(1) = 0, \quad H''(0) = -1, \; H''(1) = (i\beta\Omega - \Omega^2)(A_1 + A_2 \int_0^1 H(x)dx). \tag{15}$$

Physical meaning of the functions and constants taking part in the BVP (14, 15) is not relevant in this context. The main stumbling block for some attempts of solution of this BVP made with ordinary finite-difference methods is the function $\Sigma(x)$, which is monotonously decreasing from $\Sigma(0) = 1$ to a very small value $\Sigma(1) > 0$. This property of the function $\Sigma(x)$ makes the problem very close to singular. Another potential difficulty for numerical methods requiring some numerical integration is the constant Ω, which can vary from 1 to 2000 (corresponding to the hearing range). Big values of Ω can make numerical solutions unstable.

As was stated in previous section, all these difficulties are not relevant with the method we use.

We denote the unknown boundary value $H''(1) = K$ and introduce the polynomial

$$f(x) = \frac{1}{6}(K - 2)x + \frac{1}{2}x^2 - \frac{1}{6}(K + 1)x^3.$$

Obviously, $f(0) = f(1) = 0$, $f''(0) = 1$, $f''(1) = -K$, and $f(x)$ is determined uniquely by these boundary conditions. Then we make the substitution $h(x) = H(x) + f(x)$ in the BVP (14, 15). We obtain the BVP for the function h with the boundary conditions

$$h(0) = h(1) = 0, \quad h''(0) = h''(1) = 0,$$

which were discussed in the previous section.

We have all the building blocks we need for Chebyshev finite dimensional approximation of the obtained BVP (see formulas (7), (11, 12), with obvious linear transformations). The constant K is not included in the differential operator but only in the right hand side of the BVP. After we solved the linear system of equations (symbolically, since K is unknown), the constant K is determined from the linear equation

$$K = (i\beta\Omega - \Omega^2)(A_1 + A_2 \int_0^1 (h(x) - f(x))dx). \tag{16}$$

The integral here is computed with the Gauss quadrature formula (see the end of Sect. 3).

Some numerical experiments with frozen coefficients (when the BVP (14, 15) is solved explicitly) have shown that the dimension of approximation $n = 30$ with

the precision of 30 decimal places give the global accuracy of the solution $H(x)$ of the order 10^{-25}. So let us consider more realistic settings.

We have chosen the precision of 14 decimal places (float[8] in Maple notation) and the following parameters: $\alpha = 1/15$, $beta = 3/2$, $\Omega = 100$, $A_1 = 1$, $A_2 = 2$, $\Sigma(x) = \exp(-3x)$, $B(x) = x + 1/2$, $M(x) = (1 - \cos(\pi x))/10$. And we have taken the dimension of approximation $n = 50$. Fig. 1 shows real and imaginary parts of the function $H(x)$ both at the nodes (broken lines) and interpolated by the formula (3). Computations with higher precision indicate that the global accuracy of the solution is of the order 10^{-10}, although boundary values are less accurate. The constant $K = -3197.5673 - i177.26132$ is determined from the equation (16) with 8 valid decimal places.

Fig. 2 shows absolute values of Chebyshev coefficients (2) of the function $H(x)$ in logarithmic scale. It is clear that Chebyshev coefficients a_k decrease exponentially for $k < 40$, and then they fluctuate at the value 10^{-12}. Fig. 2 explains the obtained accuracy of the solution and predicts the upper bound for the increase of accuracy with current settings. Indeed, higher dimension of approximation will be useless unless we use higher precision. In fact, computations with $n = 40$ gave the same accuracy as with $n = 50$.

§ 6. Conclusion

Although we stated from the beginning that we will solve linear BVPs (1), this approach is suitable for nonlinear problems as well. We will need to approximate the linear differential operator that is obtained from the equation in variations. So a finite dimensional approximation of the BVP will depend on the current approximation to the solution to the BVP. Then the BVP can be solved with Newton iterations.

Other examples of application of this method can be found in the book [1] and in [4]. The method is proved to be very reliable. Its development was restrained at the time it was invented by large amount of symbolic computations that had to be done manually.

References

1. Babenko K.I. Foundations of Numerical Analysis. Iszevsk: RCD, 2002, 848 p. (Russian).

2. de Boor C. Polynomial Interpolation // In: Proceedings of the International Congress of Mathematicians. Helsinky, 1980, **2**, p. 917-922.

3. Paszkowski S. Zastosowania Numeryczne Wielomianow i Szeregov Czebyszewa. Warszava: Panst. Wydawn. Nauk., 1975. (Polish) = Pashkovski S. Numerical

Applications of Chebyshev Polynomials and Series. M.: Nauka, 1983, 384 p. (Russian).

4. Afendikov A.L., Varin V.P. An analysis of periodic flows in the vicinity of the plane Poiseille flow // Eur. J. Mech., B/Fluids, 10:6 (1991) 577-603.

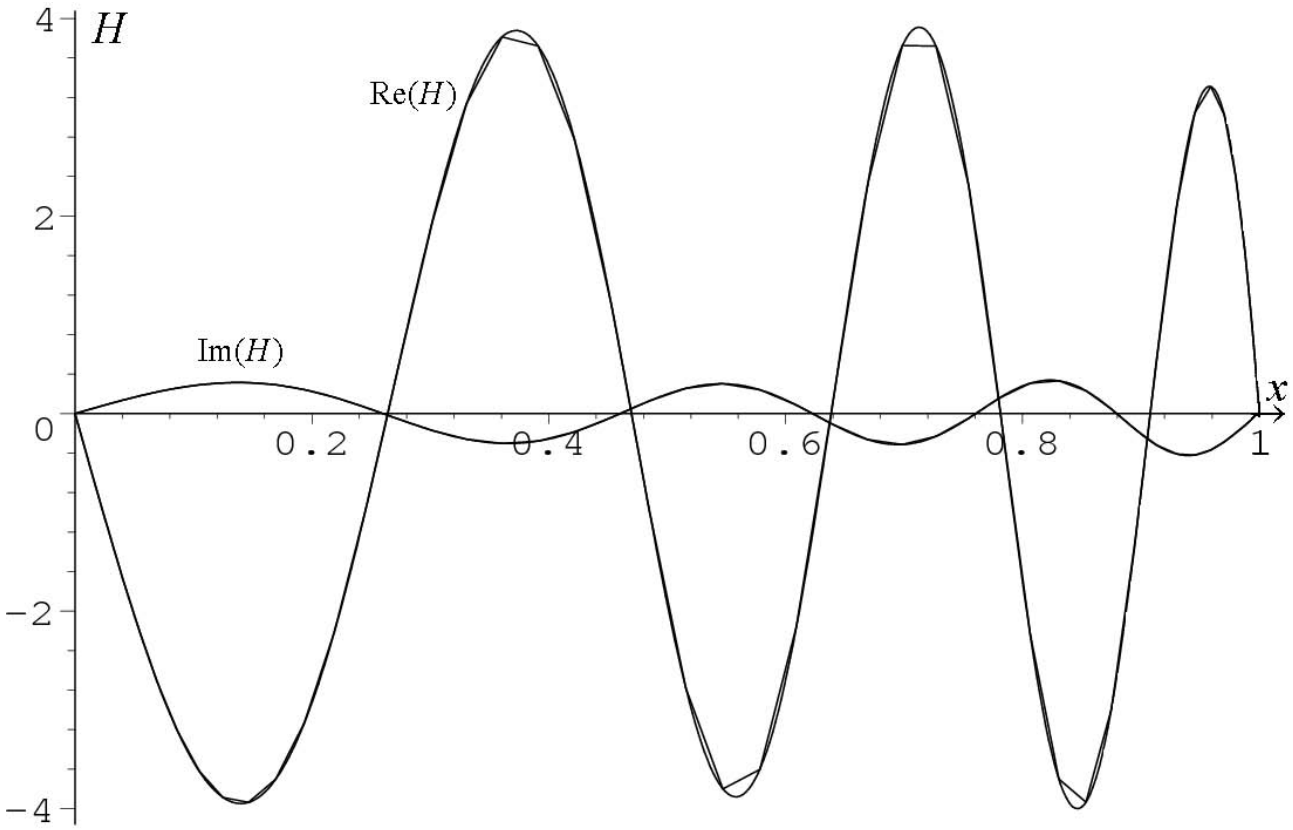

Fig.1. Real and imaginary parts of the function $H(x)$.

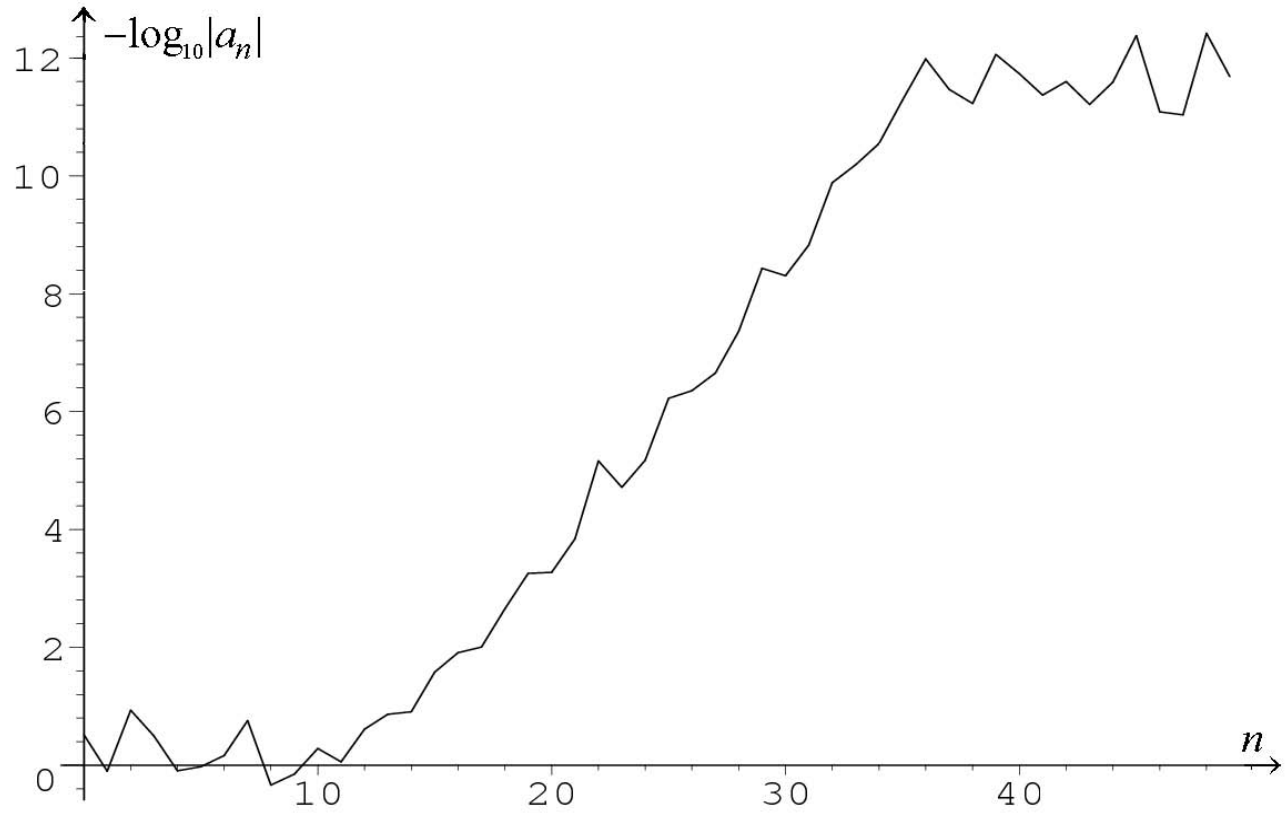

Fig.2. Absolute values of Chebyshev coefficients of the function $H(x)$.