Determining the value of the integral of a polynomial being a combination of Chebyshev polynomials of the first and second kind using the Trapezoidal rule and Simpson's method

Project 1

Authors
Hubert Kowalski
Adam Kaniasty

November 25, 2023

Contents

1 Description of the mathematical method

1.1 Chebyshev Polynomials

Chebyshev polynomials of the first kind are denoted by $T_n(x)$, where n is the degree of the polynomial. They are defined inductively as:

$$T_0(x) = 1, \quad T_1(x) = x$$

$$T_n(x) = 2x * T_{n-1}(x) - T_{n-2}(x)$$

Directly from the definition, it can be shown that these polynomials have even symmetry for even n and odd symmetry for odd n, i.e., $T_n(-x) = (-1)^n T_n(x)$. This information may be important in the context of integrating Chebyshev polynomials. The polynomial $T_n(x)$ also has n different roots in the interval $[-1, 1]$ determined by the formula:

$$x_i = cos(\frac{2i+1}{2n}\pi) \quad i = 0, 1, 2...n-1$$

These polynomials can be represented by trigonometric functions in the interval $[-1, 1]$ through the formula:

$$T_n(x) = cos(n * arccos(x))$$

Chebyshev polynomials of the second kind are denoted by $U_n(x)$, where n is the degree of the polynomial. They are defined inductively as

$$U_0(x) = 1, \quad U_1(x) = 2x$$

$$U_n(x) = 2x * U_{n-1}(x) - U_{n-2}(x)$$

These polynomials also have even symmetry, i.e., $U_n(-x) = (-1)^n U_n(x)$. The polynomial $U_n(x)$ also has n different roots in the interval $[-1, 1]$ determined by the formula:

$$x_k = \cos\left(\frac{k}{n+1}\pi\right), \quad k = 1, \dots, n$$

It is easy
to show that by expanding the polynomial $T_{n+1}(x)$ into product form we obtain:

$$T_{n+1}(x) = 2^n (x - x_0)(x - x_1) \dots (x - x_n)$$

Using such a form, one can easily bound the expression:

$$(x - x_0) \dots (x - x_n) \leqslant 2^{-n}$$

1.2 Trapezoidal Rule

The trapezoidal rule is a technique for numerical integration, i.e., approximating the value of a definite integral:

$$I(f) = \int_a^b f \mathrm{d}x$$

The composite trapezoidal rule is based on dividing the integration interval into smaller subintervals $[x_k, x_{k+1}],\ k = 0, \ldots, N-1$, where N - number of subintervals and interpolating the function f with a polynomial of degree 1. If $f(x_k) > 0$ and $f(x_{x+1}) > 0$, then the quadrature value on this interval can be interpreted as the area of the trapezoid formed by the respective segments. Summing the quadrature values on the subintervals, we obtain an approximation of the integral $S(f)$.

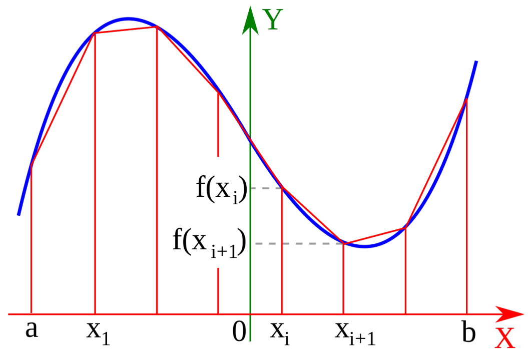


Figure 1: Trapezoidal method
Source: Wikipedia

The formula for the trapezoidal method is as follows:

$$\int_a^b f(x)dx \approx \frac{h}{2}\left[f(a) + 2\sum_{k=1}^{N-1} f(a + kh) + f(b)\right]$$

where $h = \frac{b-a}{N}$ is the length of the subinterval, and N is the number of subintervals.

The error of the trapezoidal method is the difference between the integral's value and the numerical approximation. It can be estimated using the formula:

$$|E| \leqslant \frac{(b-a)^3}{12N^2} \max_{a \leqslant x \leqslant b} |f''(x)|$$

where $f''(x)$ is the second derivative of the integrand function. This formula shows that the error decreases with the number of subintervals. However, it does not guarantee a low error value due to $\max_{a \leqslant x \leqslant b} |f''(x)|$, which can be arbitrarily large.

1.3 Simpson's Method

Simpson's method is a technique for numerical integration that uses parabolic interpolation of the function to approximate the value of a definite integral:

$$I(f) = \int_a^b f(x)\,\mathrm{d}x$$

Simple Simpson's rule is based on dividing the interval of integration into two subintervals of equal length and approximating the function f by a second-degree polynomial. It is assumed that the function f is well approximated by a parabola passing through the points $(a, f(a))$, $\left(\frac{a+b}{2}, f\left(\frac{a+b}{2}\right)\right)$, $(b, f(b))$. The quadrature value for a single interval is then given

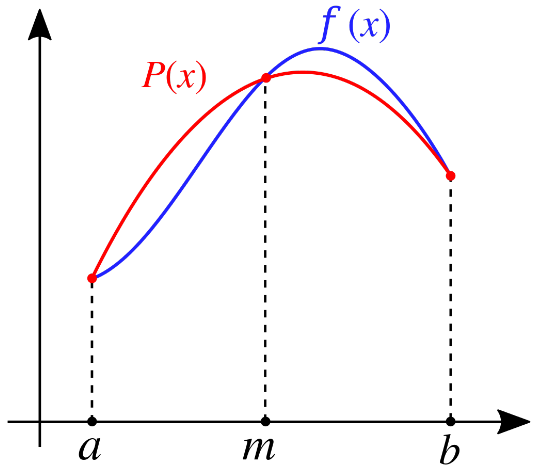


Figure 2: Simpson's Method
Source: Wikipedia

by:

$$S(f) = \frac{h}{3}\left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right]$$

where $h = \frac{b-a}{2}$ is the length of the subinterval. The error of the simple Simpson's rule is

$$E(f) = -\frac{h^5}{90} f^{(4)}(\xi)$$

where ξ is a point in the interval (a, b). Therefore, the error is zero when $f^{(4)}(x) = 0$. This implies that the approximation error is zero for polynomials of degree three or lower.

Composite Simpson's rule uses the simple Simpson's rule, dividing the integration interval into N equal subintervals, where N is the number of intervals for Simpson's quadrature. The formula for composite Simpson's rule is:

$$S(f) = \sum_{k=1}^{N} \frac{H}{6}\left(f(x_{k-1}) + 4f\left(x_{k-1} + \frac{H}{2}\right) + f(x_k)\right).$$

Here we see the use of simple Simpson's methods. However, we use a transformed formula in calculations:

$$S(f) = \frac{H}{6}\left(f(a) + f(b) + 2\sum_{k=1}^{N-1} f(a + kH) + 4\sum_{k=0}^{N-1} f\left(a + kH + \frac{H}{2}\right)\right).$$

where $H = \frac{b-a}{N}$, and $x_k = a + kH$ for $k = 0, \ldots, N$.

The error of Simpson's method can be estimated using the formula:

$$|E| \leqslant \frac{(b-a)^5}{180N^4} \max_{a \leqslant x \leqslant b} |f^{(4)}(x)|$$

where $f^{(4)}(x)$ is the fourth derivative of the integrand. This formula indicates that the error decreases with a higher power of N than in the trapezoidal method.

1.4 Error Measurements

The absolute error is defined as:

$$\Delta = \bar{x} - x,$$

where $\bar{x}$ is the approximate value, and x is the exact value.

Assuming that $x \neq 0$, the relative error is expressed by the formula:

$$\delta = \frac{\bar{x} - x}{x} = \frac{\Delta}{x},$$

which gives the ratio of the absolute error to the exact value.

1.4.1 Optimization Method for the Number of Intervals

We have prepared a function `crankyKowalskiError` for finding the optimal number of subintervals N for a numerical integral, which minimizes the total error of numerical integration. This error considers not only the difference between the approximated and the exact value of the integral but also includes a penalty for increasing the number of intervals, aiming to account for computational cost. The penalty for increasing the number of intervals grows logarithmically with N.

2 Program Description

The solution implementation was carried out using the *MATLAB* environment. The core consists of 3 functions:

```
trapezoidal(a, b, N, f, a_k),
simpson(a, b, N, f, coefficients),
chebyshev_combination(a_k, x)
```

The function *trapezoidal* is used to calculate the approximate integral of the polynomial $w_n(x) = \sum_{k=0}^{n} a_k T_k(x) U_k(x)$ on the interval [a, b] using the composite trapezoidal method consisting of N subintervals. Similarly, the function *simpson* is used to calculate the same integral but using Simpson's method. This function only works when the function f provided as an argument takes two parameters: a_k - the vector of coefficients for the combination, and x - the point at which the value is calculated.

The function *chebyshev_combination* calculates the value of the polynomial $w_n(x) = \sum_{k=0}^{n} a_k T_k(x) U_k(x)$ at point x for the coefficients a_k using a recursive relationship.

The use of the function *chebyshev_combination* in conjunction with *trapezoidal* or *simpson* to calculate the approximate value of an integral from a polynomial of the form $w_n(x)$ involves setting the desired values in the script:

a - start of the integration interval

b - end of the integration interval

N - number of subintervals in the composite trapezoidal quadrature

a_k - coefficients of the Chebyshev polynomial combinations (where the number of coefficients ensures the appropriate degree of the resulting polynomial $w_n(x)$)

and then calling the function *trapezoidal*:

```
w = trapezoidal(a, b, N, @chebyshev_combination, a_k)
```

or *simpson*:

```
w = simpson(a, b, N, @chebyshev_combination, a_k)
```

However, to calculate the exact integral using the built-in function *integral*, it is necessary to provide it with a function that takes only one argument - x. Therefore, the solution also includes functions named:

chebyshev_example_i, i = 1,2 ... 12

The solution also contains the functions *trapezoidal_general(a, b, N, f)* and *simpson_general(a, b, N, f)* that allow calculating the approximate integral using the trapezoidal method (or Simpson's method) from any function $y = f(x)$. If you want to calculate the integral for one of the functions *chebyshev_example_i*, you should use *trapezoidal_general* or *simpson_general*:

```
wt = trapezoidal_general(a, b, N, @chebyshev_example_1)
ws = simpson_general(a, b, N, @chebyshev_example_1)
```

2.1 GUI

The GUI contains three main tabs:

- **ExamplesTab** - a section where errors are generated for selected examples chosen by us.
- **ErrorsTab** - a tab where random combinations of Chebyshev polynomials of a degree chosen by the user are generated. It shows the errors of both methods.
- **HeatmapTab** - a section where the user examines the dependence of the error on the number of intervals and the degree of the polynomial. For the value of **Max Subintervals** specified by the user, errors of integration in the interval [-1, 1] are generated for combinations of first-degree polynomials, then first and second, etc., up to the tenth.

For visualizing errors, but this time depending on the integration interval for selected example combinations of Chebyshev polynomials, one can use the application *errorsGUI.mlapp*.

3 Examples

3.1 Description of Examples

1. chebyshev_example_1 - combination coefficients take the form $\{4, 2\}$. Simpson's method should calculate the exact value of the integral of this polynomial.

2. chebyshev_example_2 - combination coefficients take the form $\{15\}$. The Trapezoidal method should calculate the exact value of the integral of this polynomial.

3. chebyshev_example_3 - combination coefficients take the form $\{1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1\}$. The values of the integral of this polynomial over an interval greater than (-1,1) are very large and increase rapidly.

4. chebyshev_example_4 - the polynomial coefficients are identical to those in chebyshev_example_3. The values of the integral of this polynomial within the interval (-1,1) are relatively low.

5. chebyshev_example_5 - combination coefficients take the form $\{0.2916, 0.1978, 1.5877, -0.8045\}$. The value of the integral of this polynomial on the interval [-1,1] for a large number of subintervals N is more accurate than for a smaller N.

6. Example 6 - the polynomial is identical to the one in example 5 to show that for an increased number of subintervals N, the error (both relative and absolute) is smaller.

7. chebyshev_example_7 - combination coefficients take the form $\{1, 1, 1, 1, 1, -1\}$. This example demonstrates that outside the interval [-1,1], the value of the polynomial is determined solely by the last term of the sum of the form a_k*Un(x)*Tn(x) (the coefficient at the highest power).

8. chebyshev_example_8 - combination coefficients take the form $\{\ 10^{-8}, 10^{-8}, 10^{-8}, 10^{-8}, 10^{-8}, 10^{-8}\ \}$. All coefficients are equal to 10^{-8}. This does not prevent large values of the polynomial outside the interval [-1,1], but within this interval, the value of the integral is close to 0.

9. chebyshev_example_9 - combination coefficients take the form $\{1, 10^{-3}, 10^{-6}, 10^{-9}, 10^{-12}, 10^{-15}\ \}$. Subsequent elements of the combination are weighted by increasingly smaller weights. As a result, even on an interval larger than [-1,1], the values of the integral are relatively small.

10. chebyshev_example_10 - combination coefficients take the form $\{1, 2, 3, 4, 5, 6\}$. Based on this example, we demonstrate that due to the parity of the polynomial $w_5(x) = \sum_{k=0}^{5} a_k T_k(x) U_k(x)$, the areas under the graph of this polynomial on the interval $[-a, 0]$ and $[0, a]$ are equal.

11. example_11 - combination coefficients take the form $\{1, 1, 1, 1, 1\}$. This polynomial is a sum of the form $\sum_{k=0}^{n} a_k * T_k(x)$. It is, therefore, a combination but only of Chebyshev polynomials of the first kind.

12. example_12 - finally, one can test the performance of the trapezoidal and Simpson methods on a function of the form $y = 5 * \sin(x) - 3 * \cos(x)$.

3.2 Result Table

Test #	a	b	N	Integral	Trapezoidal	Trap Abs Error	Trap Rel Error	Simpson	Simp Abs Error	Simp Rel Error
1	-2	10	50	2.2915e+12	2.3037e+12	1.2156e+10	0.0053	2.2915e+12	2.1095e+06	9.2054e-07
2	-2	10	50	180	180	0	0	180	0	0
3	1	3	50	1.4678e+14	1.4780e+14	1.0168e+12	0.0069	1.4678e+14	2.8371e+08	1.9329e-06
4	-0.5000	0.5000	50	5.7523	5.7529	5.3565e-04	9.3118e-05	5.7523	5.8079e-07	1.0097e-07
5	-1	1	5	1.8327	1.6373	0.1955	0.1066	1.7985	0.0342	0.0187
6	-1	1	50	1.8327	1.8294	0.0033	0.0018	1.8327	3.5719e-06	1.9489e-06
7	-1.5000	-1	50	-700.9092	-701.4640	0.5548	7.9152e-04	-700.9092	1.1801e-05	1.6836e-08
8	-1	1	50	7.8782e-08	8.0000e-08	1.2179e-09	0.0155	8.0000e-08	1.2179e-09	0.0155
9	-5	5	50	10.1769	10.1771	1.6403e-04	1.6118e-05	10.1769	1.3869e-09	1.3628e-10
10	-0.7000	0	50	7.4369	7.4372	3.2101e-04	4.3165e-05	7.4369	1.5738e-07	2.1162e-08
11	-1	-0.5000	50	0.1125	0.1126	1.0833e-04	9.6293e-04	0.1125	1.3878e-17	1.2336e-16
12	-10	10	100	3.2641	3.2532	0.0109	0.0033	3.2641	1.8147e-06	5.5594e-07

Figure 3: Results of the trapezoidal and Simpson methods on example functions

The results of the Simpson's method and the trapezoidal method on the given examples are consistent with the predicted properties. For instance, one can observe that for test number 3, i.e., the function chebyshev_example_3, the absolute error calculated by both methods is of the order of 10^{10}. However, due to the expected huge value of the actual integral in the given range, the relative error (divided by the actual value) remains small. It is of the order of 10^{-2} for the trapezoidal method and 10^{-8} for Simpson's method.
By comparing example number 5 with example number 6, one can observe that increasing the number of subintervals from 5 to 50 reduces the relative error of the integral calculation by a factor of 100 for the trapezoidal method and by 10,000 for Simpson's method.
Based on example number 9, it can be concluded that, as expected, the application of low coefficients (of the order of 10^{-15}) at the highest powers of the polynomial reduces the value of the integral on an interval close to [-1, 1] and improves the properties of numerical integration.
It is also worth noting that when applying a different combination of polynomials, namely the combination of the form $\sum_{k=0}^{n} a_k * T_k(x)$ (see example 10), the error of numerical integration of such functions is smaller than in the case of previously studied combinations of Chebyshev polynomials of both the first and second kind.

4 Error Measurements

Error analysis for the Simpson's method and the trapezoidal method shows that Simpson's method generally achieves smaller errors compared to the trapezoidal method. This is consistent with expectations since Simpson's method is more accurate assuming appropriate smoothness of the integrated function.

Examples: 20 Poly degree: 10 Start

Test #	True Integral	Simpson	Simpson Diff	Simpson Error (%)	Trapezoidal	Trapezoidal Diff	Trapezoidal Error (%)
1.0000	34.9221	34.9233	0.0012	0.0034	35.4304	0.5083	1.4555
2.0000	43.6781	43.6813	0.0032	0.0073	44.5442	0.8661	1.9830
3.0000	25.1414	25.1433	0.0018	0.0073	25.5026	0.3612	1.4366
4.0000	31.4758	31.4774	0.0016	0.0050	31.9745	0.4987	1.5844
5.0000	28.1822	28.1832	0.0010	0.0036	28.4906	0.3084	1.0943
6.0000	40.9996	41.0033	0.0037	0.0091	41.9138	0.9142	2.2297
7.0000	26.5631	26.5652	0.0021	0.0081	27.1713	0.6082	2.2896
8.0000	28.5824	28.5843	0.0019	0.0065	29.1942	0.6117	2.1403
9.0000	31.9535	31.9549	0.0015	0.0046	32.4567	0.5032	1.5749
10.0000	22.7508	22.7525	0.0017	0.0073	23.1751	0.4243	1.8648
11.0000	28.2144	28.2167	0.0023	0.0080	28.7574	0.5430	1.9244
12.0000	34.1720	34.1755	0.0034	0.0101	34.9873	0.8152	2.3857
13.0000	18.2585	18.2586	0.0002	0.0009	18.4255	0.1671	0.9150
14.0000	33.8998	33.9005	0.0008	0.0022	34.2549	0.3551	1.0474
15.0000	21.5059	21.5078	0.0019	0.0088	22.0882	0.5823	2.7077
16.0000	22.9032	22.9034	0.0002	0.0010	23.0665	0.1634	0.7133
17.0000	28.9199	28.9205	0.0006	0.0020	29.1333	0.2134	0.7378

Figure 4: Comparison of errors for Simpson's method and the trapezoidal method with varying numbers of subintervals.

As seen in Figure 4, for combinations of 10th-degree polynomials, Simpson's method achieves error values less than or equal to 0.01%, while for the trapezoidal method the errors are less than or equal to 2%. This is significant from the point of view of choosing a numerical method for specific applications where high precision is required.

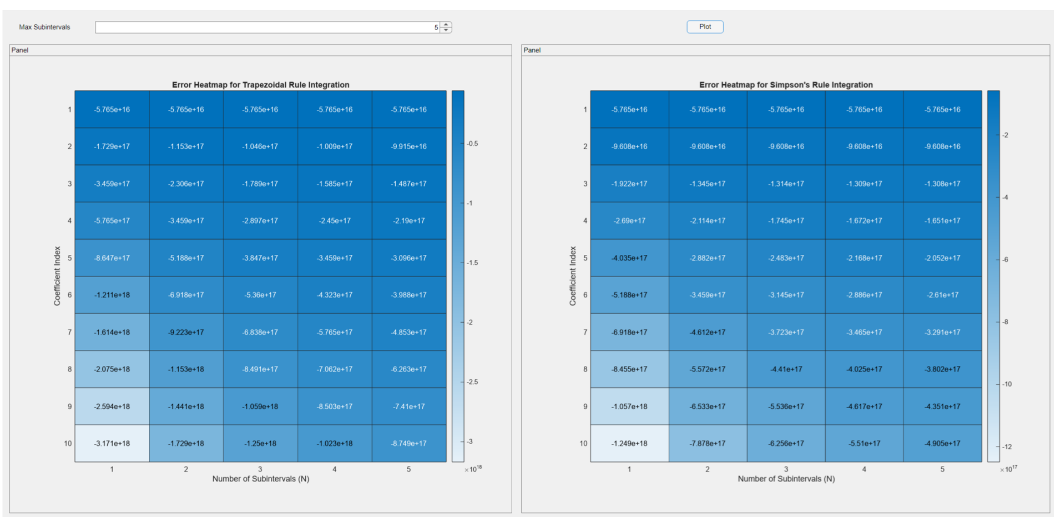


Figure 5: Heatmap of errors for the trapezoidal method (left) and Simpson's method (right) depending on the number of subintervals and the coefficient index.

The heatmap in Figure 5 displays the distribution of errors depending on the number of subintervals (horizontal axis) and the number of coefficients (vertical axis). The colors represent the magnitude of the error, where darker shades of blue indicate larger errors. The heatmap allows us to conclude that increasing the number of subintervals (N) leads to a reduction in errors, especially in Simpson's method. It is worth noting what happens for combinations of first-degree polynomials. Although both methods should be accurate, we can observe minimal errors resulting from the representation of numbers in computer memory.

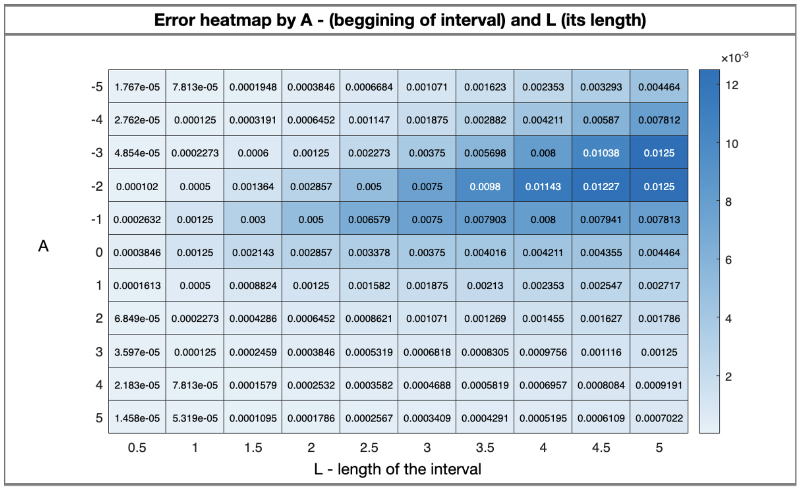


Figure 6: Heatmap of errors for the trapezoidal method depending on the start of the interval (A) and its length (L)

Based on the heatmap in Figure 6, it can be stated that the longer the integration interval, the greater the relative error in calculating the approximate value of the integral. This dependency generally does not depend on the starting point of the integration interval. However, it is important to note that the relative error value strongly depends on the location where the function is integrated. For example, the integral over the interval [5, 10] was calculated with an error twenty times smaller than the integral over the interval [-2,3]. This may be a consequence of the properties of higher-degree Chebyshev polynomials, which behave oscillatory within the interval [-1,1].

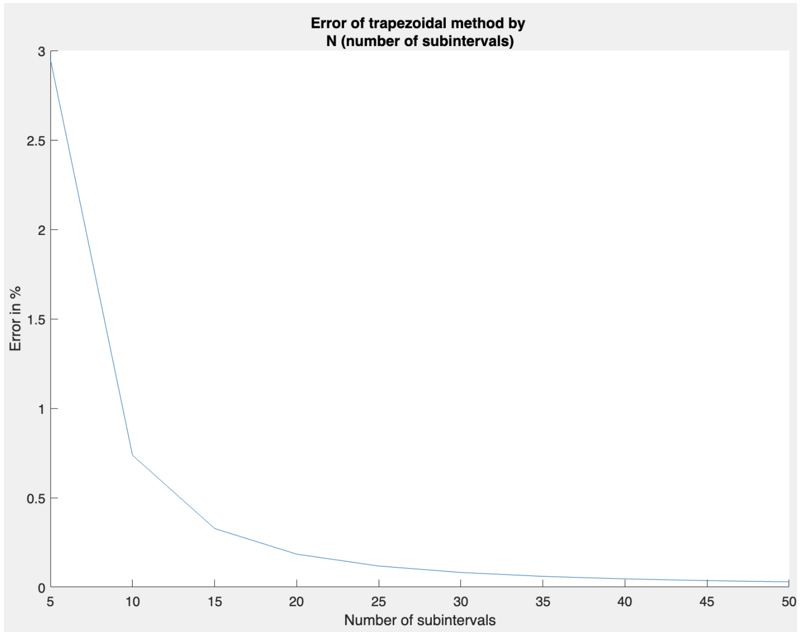


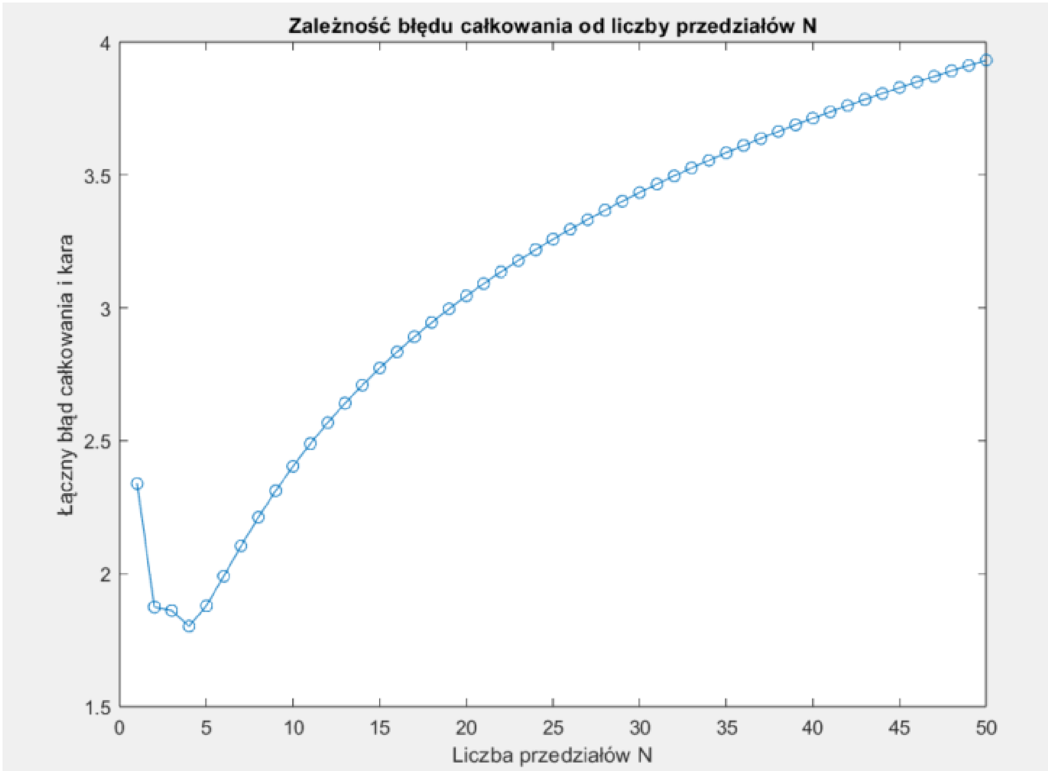
Figure 7: Graph of the dependence of the relative error expressed in (%) on the number of subintervals in integration by the trapezoidal method

Based on the graph in Figure 7, it can be observed that for a certain example combination of Chebyshev polynomials, the integral of the resulting polynomial is calculated more accurately when the integration interval is divided into more subintervals (N). This property is consistent with theoretical expectations, according to which the error estimation is inversely proportional to the value of N:

$$|E| \leqslant \frac{(b-a)^3}{12N^2} \max_{a \leqslant x \leqslant b} |f''(x)|$$

4.1 Optimization of the Number of Intervals

We applied our test optimization function to the example of a Chebyshev polynomial combination with coefficients [1, 2, 3, 4, 5], with a maximum number of intervals of 50, in the range (-10,10).



For the analyzed function, the "optimal" number of intervals is 4, which seems to be too low a value to use in applications. We suspect that there are better penalizing functions that take into account the addition of the number of intervals and the expected degree of accuracy. We present our function as a curiosity that may be developed in the future.

5 Conclusions

1. **Accuracy of Methods:** Simpson's method generally provides greater accuracy compared to the trapezoidal method, especially for functions with higher smoothness. This is evident in error analysis, where Simpson's method achieves significantly smaller relative errors compared to the trapezoidal method, which is consistent with the theoretical error estimates for both methods.

2. **Behavior of Chebyshev Polynomials:** Chebyshev polynomials have unique oscillatory properties on the interval $[-1, 1]$, which affects the results of numerical integration. For intervals outside $[-1, 1]$, the values of these polynomials grow very rapidly, which can lead to large errors if the coefficients are not appropriately adjusted.

3. **Effect of the Number of Subintervals:** Increasing the number of subintervals N leads to a reduction in the error of the integral approximation for both methods. This is consistent with theoretical error estimates, which are inversely proportional to the square (for the trapezoidal method) or the fourth power (for Simpson's method) of the number N.

4. **Applications:** Depending on the required accuracy and characteristics of the integrand function, the choice between the trapezoidal method and Simpson's method should be made considering both computational cost and expected precision.