Mathematics Internal Assessment:

# An Investigation into The Trapezoidal Rule for Numerical Integration: Determine Its Level of Accuracy

Pages:20

## Introduction:

Definite integrals have wide applications in many areas. While some definite integrals can be evaluated using fundamental theorem of calculus, others require numerical methods or series expansion to evaluate. Such integrals include $\int_a^b e^{x^2} dx$, $\int_a^b \pi \cdot (1 + sin^2 x)^2 dx$, $\int_a^b \frac{sin\, x}{x} dx$, etc.

Trapezoidal rule is a numerical method used in definite integral evaluation. Compared with upper and lower rectangle methods, it converges quicker. Although series expansion converges quicker than the trapezoid rule, but to write a function as an infinite series involves differentiation that becomes harder to proceed as the order increases.

It always amuses me that as partition intervals increase, trapezoid rule approximations are more accurate. I am interested in how the accuracy of the results are determined. Since existing formula only gives an error bound, in this investigation, I aim to find the common trend of trapezoidal rule approximation error with number of partition intervals. The comparisons are investigated using functions with known closed-form antiderivatives. The results of trapezoidal rule of different partition intervals on a given domain is compared with the results obtained using fundamental theorem of calculus. The trend of error is identified, proved and then used to calculate definite integrals without closed-form antiderivatives with attention to the significant figures.

## Trapezoid method and error bound formula:

Trapezoidal rule method is stated as follows: Given $\int_a^b f(x)dx$, $[a,b]$ is first partitioned into $n$ subintervals. The width of each interval is $h = \frac{b-a}{n}$.

Each subinterval is then approximated by the trapezoid with height $h$, base $f(a_{i-1})$ and $f(a_i)$, the area is:

$$S_i = \frac{1}{2}h(f(a_{i-1}) + f(a_i))$$

The error for each interval is:

$$E_{Ti} = \int_{a_{i-1}}^{a_i} f(x)dx - S_i$$

Figure 1 shows $[a,b]$ partitioned into 4 subintervals.

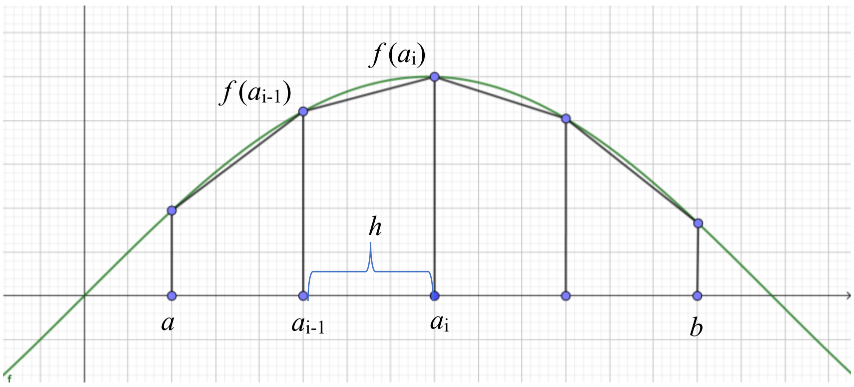



Figure 1. Interval $[a,b]$ with 4 subintervals

Then, for $[a,b]$ partitioned into n subintervals,

$$a_n = b$$

$$a_0 = a$$

$$\sum_{i=1}^{i=n} S_i = \frac{1}{2}h(f(a_0) + f(a_1)) + \frac{1}{2}h(f(a_1) + f(a_2)) + \frac{1}{2}h(f(a_2) + f(a_3)) + \cdots$$

$$+ \frac{1}{2}h(f(a_{n-1}) + f(a_n))$$

$$= h \times (\frac{1}{2}f(a_0) + f(a_1) + f(a_2) + \cdots + f(a_{n-1}) + \frac{1}{2}f(a_n))$$

The error for trapezoidal rule is:

$$E_T = \int_a^b f(x)dx - h \times (\frac{1}{2}f(a_0) + f(a_1) + f(a_2) + \cdots + f(a_{n-1}) + \frac{1}{2}f(a_n))$$

The error bound for trapezoidal rule in $[a, b]$ is[1]

$$|E_T| \le \frac{M \cdot (b-a) \cdot h^2}{12} = \frac{M \cdot (b-a)^3}{12 \cdot n^2}$$

where $M$ is the maximum of $f''(x)$ on $[a, b]$.

However, the value of $M$ depends on $f''(x)$ and the interval $[a, b]$ and it varies significantly. For example, let $f(x) = e^{x^2}$, so $f''(x) = 2e^{x^2} + 4x^2e^{x^2}$. Different evaluation intervals give error bounds with significantly different orders of magnitude. Thus, I would like to find the common trend of trapezoidal rule approximation errors with number of partition intervals.

---

1 *Section 7.7 Deriving the Trapezoidal Rule Error*. math.ucsd.edu/~ebender/20B/77_Trap.pdf. Accessed 18th, Feb., 2021.

## Comparison:

As the error bound formula suggests, for a given definite integral, increasing the partition intervals $n$ will reduce error bound by $n^2$.

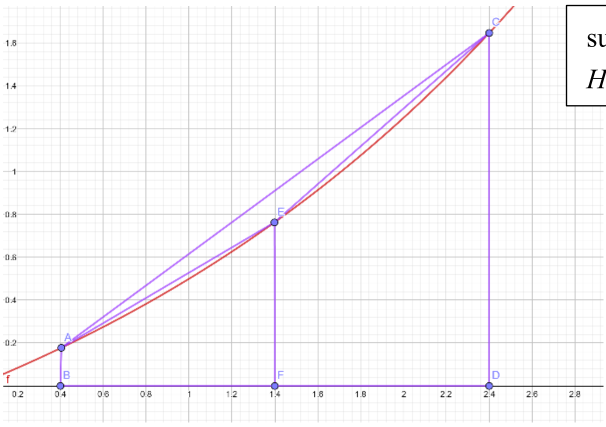



Figure 2. A portion of an interval with *n* partitions

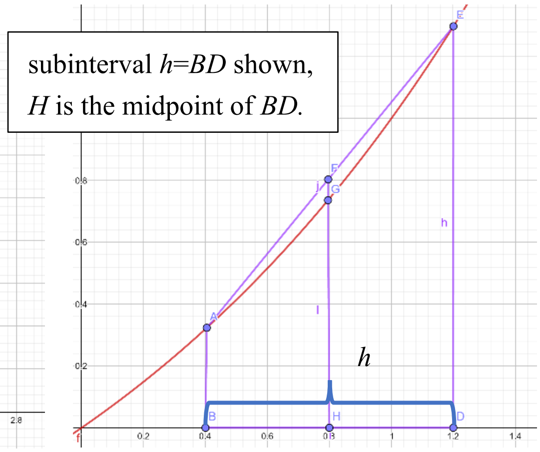



Figure 3. A portion of an interval with 2*n* partitions

As figure 2 shows, when $n$ is increased, such as the subinterval $BD$ is further divided into $BF$ and $FD$, the errors (from the area of arch $AC$ to area of arch $AE$ and arch $EC$) decrease significantly. The increase in accuracy is related to the decrease in the error. If one feature of the function can be used to reflect the change of $n$ on the errors, then this relationship can be modelled to find the common trend.

Since the decrease in the error is largely due to the decrease in the area of the arch, to quantify this change, in every subinterval, length of $FG$ can be used to reflect the change of $n$ on the errors. $FG$ is the difference between the function value at $H$ and y-coordinate of $AE$'s midpoint.

$$FG = HF - HG$$

$$\text{Let } FG = d$$

$$d_i = f(a_{i-1}) + 0.5 \cdot (f(a_i) - f(a_{i-1})) - f(a_{i-1} + 0.5 \cdot h)$$

$$= 0.5 \cdot (f(a_i) + f(a_{i-1})) - f(a_{i-1} + 0.5 \cdot h)$$

For a total of $n$ subintervals, the average value of $d_i$ can be used to reflect the change of $n$ on the errors:

$$\overline{d} = \frac{1}{n}\sum_{i=1}^{i=n} d_i$$

## The common trend:

Several basic functions with known closed-form antiderivative are used to find the trend using MATLAB programs. They are listed below:

1. $f(x) = e^{k \cdot x}$ $\quad\quad \int f(x)dx = \frac{1}{k} \cdot e^{k \cdot x}$
2. $f(x) = cos(k \cdot x)$ $\quad\quad \int f(x)dx = \frac{1}{k} \cdot sin(k \cdot x)$
3. $f(x) = x^k$ $\quad\quad \int f(x)dx = \frac{1}{k+1} \cdot x^{k+1}$

where $k = 1,2,3,4,5,6$ and partition intervals $n = 10,10^2,10^3,10^4,10^5,10^6$

Steps for finding trend of trapezoidal rule error and $\overline{d}$ as *n* changes in MATLAB:

1. Calculate definite integrals of basic functions using fundamental theorem of calculus.
2. For a given *n*, trapezoidal rule is used to approximate the definite integral.
3. Calculate the error $E_T$ between 2 results.
4. Calculate $\overline{d}$.
5. Calculate $\frac{E_T}{\overline{d}}$.

Sample calculation:

1. $\int_a^b e^{k \cdot x} dx$

   Let $a = 0, b = 3$ and $b - a = 3$:

Table 1. Several computed values for *k*=1

| $k = 1$ | | | |
|---|---|---|---|
| $n$ | $E_T$ | $\overline{d}$ | $\frac{E_T}{\overline{d}}$ |
| 10 | 0.142927274 | 0.07143687 | 2.000749398 |
| $10^2$ | 0.001431394 | 0.000715694 | 2.0000075 |
| $10^3$ | 1.43142E-05 | 7.15708E-06 | 2.000000072 |
| $10^4$ | 1.43142E-07 | 7.15708E-08 | 2.000000766 |
| $10^5$ | 1.43156E-09 | 7.15708E-10 | 2.000199572 |
| $10^6$ | 1.37241E-11 | 7.15626E-12 | 1.917780695 |

The ratio of $\frac{E_T}{\overline{d}}$ is around 2 for $k = 1$. other $k$ values are computed on table 2.

Table 2. Computed values of $\frac{E_T}{\overline{d}}$ for different values of $n$ and $k$

| $\frac{E_T}{\overline{d}}$ | | $k$ | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 |
| $n$ | $10$ | 2.000749398 | 2.002990386 | 2.00670151 | 2.011847545 | 2.018380265 | 2.026239491 |
| | $10^2$ | 2.0000075 | 2.000029999 | 2.000067495 | 2.000119985 | 2.000187462 | 2.000269922 |
| | $10^3$ | 2.000000072 | 2.0000003 | 2.000000675 | 2.0000012 | 2.000001875 | 2.0000027 |
| | $10^4$ | 2.000000766 | 1.999999928 | 2.000000005 | 1.999999996 | 1.999999999 | 2.000000024 |
| | $10^5$ | 2.000199572 | 1.999970271 | 2.000007327 | 1.999996418 | 1.999998653 | 2.000002645 |
| | $10^6$ | 1.917780695 | 2.003471981 | 2.000440777 | 1.999639601 | 2.001984801 | 1.999328293 |

From table 2, all ratios are around 2, which equals $\frac{2}{3}(b-a)$. Other functions with different $(b-a)$ ranges are calculated similarly in MATLAB as the sample listed above. Likewise, the ratios $\frac{E_T}{\overline{d}}$ for different $n$ and $k$ are all around the same figure for a chosen function and a chosen interval. The results are listed in table 3. It is generated from table 2, 4, 5 and table 11, 13 in the appendices.

Table 3. Computed values and ratios for different functions

| Functions | $\frac{E_T}{\overline{d}}$ | $a$ | $b$ | $b-a$ | $\frac{\frac{E_T}{\overline{d}}}{b-a}$ |
|---|---|---|---|---|---|
| $f(x)=x^k$ | 1.333 | 3.00 | 5.00 | 2.00 | 0.6667 |
| $f(x)=e^{k\cdot x}$ | 2.000 | 0.0 | 3.00 | 3.00 | 0.6667 |
| | 10.00 | 5.0 | 20.0 | 15.0 | 0.6667 |
| | 0.267 | 0.1 | 0.500 | 0.400 | 0.6675 |
| $f(x)=\cos(k\cdot x)$ | 0.174 | 0.0 | $\frac{\pi}{12}\approx 0.262$ | 0.262 | 0.6646 |

From table 3, $\frac{\frac{E_T}{\overline{d}}}{b-a}$ are around 0.667, which is $\frac{2}{3}$. This suggests a possible relationship between $\overline{d}$ and $E_T$:

$$E_T \approx \frac{2}{3}(b-a)\cdot\overline{d}$$

Caution:

Table 4. Results for $\int_a^b e^{k \cdot x} dx$ when $a = 0.1$ and $b = 0.5$

| $\frac{E_T}{d}$ | | $k$ | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 |
| $n$ | 10 | 0.266668444 | 0.266673777 | 0.266682665 | 0.266695105 | 0.266711095 | 0.266730634 |
| | $10^2$ | 0.266666684 | 0.266666738 | 0.266666827 | 0.266666951 | 0.266667111 | 0.266667307 |
| | $10^3$ | 0.266666668 | 0.266666664 | 0.26666667 | 0.266666668 | 0.266666671 | 0.266666673 |
| | $10^4$ | 0.266684162 | 0.266664853 | 0.266667152 | 0.266666918 | 0.266666667 | 0.266666434 |
| | $10^5$ | 0.267532107 | 0.266148844 | 0.266808376 | 0.266774476 | 0.266713967 | 0.266706029 |
| | $10^6$ | -0.191981221 | 0.197286349 | 0.302519247 | 0.262635796 | 0.236267646 | 0.270401919 |

Table 5. Results for $\int_a^b e^{k \cdot x} dx$ when $a = 5$ and $b = 20$

| $\frac{E_T}{d}$ | | $k$ | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 |
| $n$ | 10 | 10.09190133 | 10.34697589 | 10.71337656 | 11.1302732 | 11.54725271 | 11.93257444 |
| | $10^2$ | 10.00093731 | 10.00374699 | 10.00842227 | 10.01495193 | 10.02332034 | 10.03350755 |
| | $10^3$ | 10.00000937 | 10.0000375 | 10.00008437 | 10.00015 | 10.00023436 | 10.00033748 |
| | $10^4$ | 10.00000012 | 10.00000037 | 10.00000085 | 10.0000015 | 10.00000235 | 10.00000338 |
| | $10^5$ | 10.00000963 | 9.999989808 | 9.999998143 | 9.999997422 | 9.999997341 | 9.999998327 |
| | $10^6$ | 9.98867081 | 10.00054901 | 9.999891065 | 10.0002318 | 9.999840316 | 10.00004637 |

When $n <= 10^5$, the ratios of $\frac{E_T}{d}$ are consistent. However, as partition increases, $h$ and $d$ decreases. When $n = 10^6$, for some evaluation intervals, such as in table 5, the ratios are not consistent with those obtained when $n <= 10^5$. this is due to limitations in the computing power. In computer programs, the calculations are carried within a certain accuracy level. For my program, it is 15 significant figures. So, the rounding errors in computer programs would make some results incorrect when $n = 10^6$. For larger partition interval $[a, b]$ and function values, the ratio is also consistent when $n = 10^6$. As a result, in general, when using trapezoidal rule to approximate the definite integrals, the partition intervals can take to be $10^5$.

## Proof of $E_T \approx \frac{2}{3}(b-a)\cdot\overline{d}$:

The error formula: [2]

From introduction of trapezoid method,

$$E_{Ti+1} = \int_{x_i}^{x_{i+1}} f(x)dx - S_{i+1}$$

$$\therefore \int_{x_i}^{x_{i+1}} f(x)dx = S_{i+1} + E_{Ti+1}$$

The integral for one subinterval is $\int_{x_i}^{x_{i+1}} f(x)dx$, using integration by parts it can be written as:

$$\int_{x_i}^{x_i+1} f(x)dx$$

$$= \int_0^h f\left(t+x_i\right)dt$$

$$= \left[(t+A)f\left(t+x_i\right)\right]_0^h - \int_0^h (t+A)f'\left(t+x_i\right)dt$$

$$= \left[(t+A)f\left(t+x_i\right)\right]_0^h - \left[\left(\frac{(t+A)^2}{2}+B\right)f'\left(t+x_i\right)\right]_0^h + \int_0^h \left(\frac{(t+A)^2}{2}+B\right)f''\left(t+x_i\right)dt$$

$A$ and $B$ are arbitrary constants. To express $\int_{x_i}^{x_i+1} f(x)dx = S_{i+1} + E_{Ti+1}$, $A$ and $B$ can take specific values so that:

$\left[(t+A)f\left(t+x_i\right)\right]_0^h = S_{i+1}$ and $\left[\left(\frac{(t+A)^2}{2}+B\right)f'\left(t+x_i\right)\right]_0^h = 0$

Since $S_{i+1} = \frac{h}{2}\cdot(f\left(x_{i+1}\right)+f\left(x_i\right))$ and

$\left[(t+A)f\left(t+x_i\right)\right]_0^h = (h+A)f\left(h+x_i\right) - A\cdot f\left(x_i\right) = (h+A)f\left(x_{i+1}\right) - A\cdot f\left(x_i\right)$

Let $A = -\frac{h}{2}$, $\left[(t+A)f\left(t+x_i\right)\right]_0^h$ is the area for the trapezoid.

$$\because \left[\left(\frac{(t+A)^2}{2}+B\right)f'\left(t+x_i\right)\right]_0^h = \left(\frac{(h+A)^2}{2}+B\right)f'\left(h+x_i\right) - \left(\frac{A^2}{2}+B\right)f'\left(x_i\right)$$

---

[2] *Section 7.7 Deriving the Trapezoidal Rule Error*. math.ucsd.edu/~ebender/20B/77_Trap.pdf. Accessed 18th, Feb., 2021.

$$=\left(\frac{(h-\frac{h}{2})^2}{2}+B\right)f'\left(x_{i+1}\right)-\left(\frac{(-\frac{h}{2})^2}{2}+B\right)f'\left(x_i\right)$$

$$=\left(\frac{h^2}{8}+B\right)f'\left(x_{i+1}\right)-\left(\frac{h^2}{8}+B\right)f'\left(x_i\right)$$

$\therefore$Let $B=-\frac{h^2}{8}$, so that $\left[\left(\frac{(t+A)^2}{2}+B\right)f'\left(t+x_i\right)\right]_0^h=0$

$$\therefore \int_{x_i}^{x_i+1} f(x)dx=\int_0^h f\left(t+x_i\right)dt=\left[(t-\frac{h}{2})f\left(t+x_i\right)\right]_0^h+\int_0^h\left(\frac{(t+A)^2}{2}-\frac{h^2}{8}\right)f''\left(t+x_i\right)dt$$

The error $E_T$ when interval $[a,b]$ is partitioned into *n* intervals is:

$$E_T=E_{T_1}+E_{T_2}+\cdots+E_{T_n}$$

$$=\int_0^h\left(\frac{(t-\frac{h}{2})^2}{2}-\frac{h^2}{8}\right)f''\left(t+x_0\right)dt\quad+\cdots+\int_0^h\left(\frac{(t-\frac{h}{2})^2}{2}-\frac{h^2}{8}\right)f''\left(t+x_{n-1}\right)dt$$

$$=\int_0^h\left(\frac{(t-\frac{h}{2})^2}{2}-\frac{h^2}{8}\right)\left(f''\left(t+x_0\right)+\cdots+f''\left(t+x_{n-1}\right)\right)dt$$

Likewise, when the interval is cut into *2n* intervals, $h'=\frac{h}{2}$, the error $E_T'$ is:

$$E_T'=E_{T_1}'+E_{T_2}'+\cdots+E_{T_{2n}}'$$

$$=\int_0^{\frac{h}{2}}\left(\frac{(t-\frac{0.5\cdot h}{2})^2}{2}-\frac{(0.5\cdot h)^2}{8}\right)\left(f''\left(t+x'_0\right)+\cdots+f''\left(t+x'_{2n-1}\right)\right)dt$$

Ratio of $\frac{E_T}{E_T'}$:

$A_{Trapezoid}$ represents the value of definite integral approximated using $n$ trapezoid partitions. $A'_{Trapezoid}$ represents the value of definite integral approximated using $2n$ trapezoid partitions.

Table 6: related subscripts $i$ when interval $[a,b]$ is partitioned from $n$ pieces to $2n$ pieces

| $2n$ partitions ($x'_i$) | 0 | 1 | 2 | 3 | 4 | 5 | 6 | … | $2n$-3 | $2n$-2 | $2n$-1 | $2n$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ partitions ($x_i$) | 0 | | 1 | | 2 | | 3 | … | | $n$-1 | | $n$ |

$x_i$ represents $x$-coordinate when definite integral is approximated using $n$ trapezoid partitions. $x'_i$ represents $x$-coordinate when definite integral is approximated using $2n$ trapezoid partitions.

To make following statements more concise,

let $H_1 = \left( \frac{(t-\frac{h}{2})^2}{2} - \frac{h^2}{8} \right)$ and $H_2 = \left( \frac{(t-\frac{0.5\cdot h}{2})^2}{2} - \frac{(0.5\cdot h)^2}{8} \right)$:

$E_T$ is spit into 2 integrals from 0 to $\frac{h}{2}$ and $\frac{h}{2}$ to $h$. This would help link $E_T$ to $E_T'$:

$$E_T = \int_0^h H_1\left( f''(t+x_0) + f''(t+x_1) + \cdots + f''(t+x_{n-2}) + f''(t+x_{n-1}) \right) dt$$

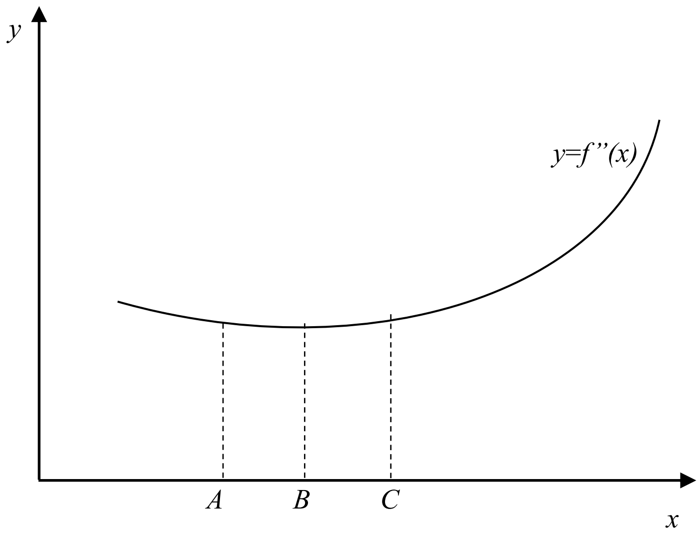
$$= \int_0^{\frac{h}{2}} H_1\left( f''(t+x_0) + f''(t+x_1) + \cdots + f''(t+x_{n-1}) \right) dt +$$

$$\int_{\frac{h}{2}}^{h} H_1\left( f''(t+x_0) + f''(t+x_1) + \cdots + f''(t+x_{n-1}) \right) dt$$

$$= \int_0^{\frac{h}{2}} H_1 dt \cdot \int_0^{\frac{h}{2}} \left( f''(t+x_0) + f''(t+x_1) + \cdots + f''(t+x_{n-1}) \right) dt +$$

$$\int_{\frac{h}{2}}^{h} H_1 dt \cdot \int_{\frac{h}{2}}^{h} \left( f''(t+x_0) + f''(t+x_1) + \cdots + f''(t+x_{n-1}) \right) dt$$

According to table 4, some subscripts represent the same value, e.g., $x$-coordinate of $x_2$ for $n$ partitions is the same with $x_4$ for $2n$ partitions.




$AC=h$

$AB=BC=\frac{h}{2}$

| A | B | C |
|---|---|---|
| $x_i$ | | $x_{i+1}$ |
| $x_{2i}$ | $x_{2i+1}$ | $x_{2i+2}$ |

Graph 1. illustration of relationship between coordinates

Graph 1 shows part of partition between interval $[a,b]$, it can be seen that

$$\int_{\frac{h}{2}}^{h} f''(t+x_i)\,dt = \int_0^{\frac{h}{2}} f''(t+x'_{2i+1})\,dt$$

Convert the above subscripts for $n$ partitions into subscripts for $2n$ partitions, we get:

$$E_T = \int_0^{\frac{h}{2}} H_1 dt \cdot \int_0^{\frac{h}{2}} \left( f''(t+x'_0) + f''(t+x'_2) + \cdots + f''(t+x'_{2n-2}) \right) dt +$$
$$\int_{\frac{h}{2}}^{h} H_1 dt \cdot \int_0^{\frac{h}{2}} \left( f''(t+x'_1) + f''(t+x'_3) + \cdots + f''(t+x'_{2n-1}) \right) dt$$

Let

$$F_1 = \int_0^{\frac{h}{2}} \left( f''(t+x'_0) + f''(t+x'_2) + \cdots + f''(t+x'_{2n-2}) + f''(t+x'_1) + f''(t+x'_3) + \cdots + f''(t+x'_{2n-1}) \right) dt$$
$$F_2 = \int_0^{\frac{h}{2}} \left( f''(t+x'_1) + f''(t+x'_3) + \cdots + f''(t+x'_{2n-1}) \right) dt$$
$$F_3 = \int_0^{\frac{h}{2}} \left( f''(t+x'_0) + f''(t+x'_2) + \cdots + f''(t+x'_{2n-2}) \right) dt$$

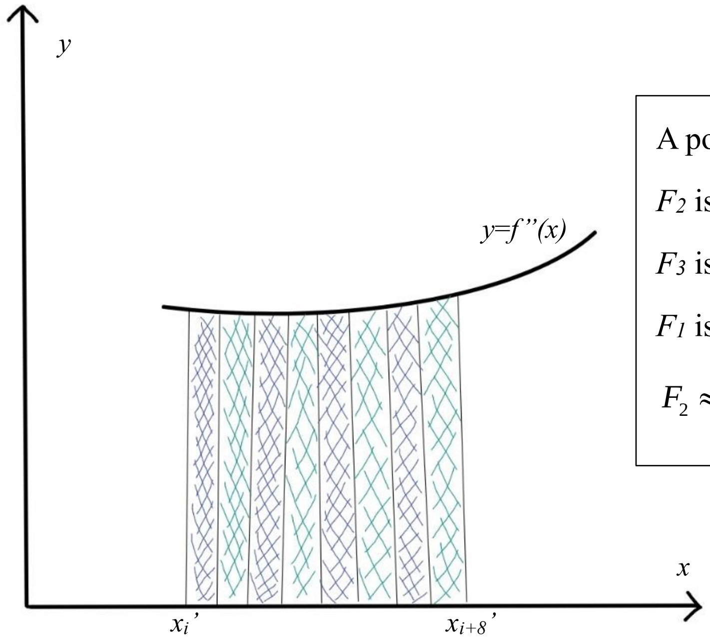
$\therefore F_1 = F_3 + F_2$

Since $F_2$ involves half of the integration terms in $F_1$ alternatively,

$$\therefore F_2 \approx \frac{F_1}{2}$$

This is shown in graph 2. Although this is only an approximation, however, in real evaluations, as partition intervals increase, the difference between $f''(t+x'_n)$ and

$f''\left(t+x'_{n+1}\right)$ will decrease, so this approximation can be used in later statements.




A portion of $f''(x)$ is shown left

$F_2$ is the area shaded in green

$F_3$ is the area shaded in dark blue

$F_1$ is the total shaded area

$$F_2 \approx \frac{F_1}{2}$$

Graph 2. illustration of $F_2 \approx \frac{F_1}{2}$

$$\therefore E_T = \int_0^{\frac{h}{2}} H_1 dt \cdot F_3 + \int_{\frac{h}{2}}^{h} H_1 dt \cdot F_2$$

Relate $E_T$ to $E_T'$:

$$E_T = \int_0^{\frac{h}{2}} H_1 dt \cdot F_3 + (\int_0^{\frac{h}{2}} H_1 dt \cdot F_2 - \int_0^{\frac{h}{2}} H_1 dt \cdot F_2) + \int_{\frac{h}{2}}^{h} H_1 dt \cdot F_2$$

$$= \int_0^{\frac{h}{2}} H_1 dt \cdot (F_3 + F_2) - \int_0^{\frac{h}{2}} H_1 dt \cdot F_2 + \int_{\frac{h}{2}}^{h} H_1 dt \cdot F_2$$

And

$$E_T' = \int_0^{\frac{h}{2}} H_2 \left( f''\left(t+x'_0\right) + f''\left(t+x'_1\right) + \cdots + f''\left(t+x'_{2n-2}\right) + f''\left(t+x'_{2n-1}\right)\right) dt = \int_0^{\frac{h}{2}} H_2 \cdot F_1$$

The integral values involved are:

$$\int_0^{\frac{h}{2}} H_1 dt = \frac{1}{8} \bullet \int_0^{\frac{h}{2}} \left( 4(t-\frac{h}{2})^2 - h^2 \right) dt = \int_0^{\frac{h}{2}} \frac{1}{8} \bullet (4 \cdot t^2 - 4 \cdot t \cdot h) dt$$

$$= \frac{1}{8} \bullet (\frac{4}{3} \cdot t^3 - 2 \cdot t^2 \cdot h) \Big|_0^{\frac{h}{2}} = \frac{1}{8} \bullet (\frac{h^3}{6} - \frac{h^3}{2}) = -\frac{h^3}{24}$$

$$\int_{\frac{h}{2}}^{h} H_1 dt = \frac{1}{8} \bullet \int_{\frac{h}{2}}^{h} \left( 4(t-\frac{h}{2})^2 - h^2 \right) dt$$

$$= \frac{1}{8} \bullet (\frac{4}{3} \cdot t^3 - 2 \cdot t^2 \cdot h) \Big|_{\frac{h}{2}}^{h} = \frac{1}{8} \bullet (\frac{4h^3}{3} - 2h^3) = -\frac{h^3}{24}$$

$$\int_0^{\frac{h}{2}} H_2 dt = \frac{1}{8} \bullet \int_0^{\frac{h}{2}} \left( 4(t-\frac{h}{4})^2 - (\frac{h}{2})^2 \right) dt = \int_0^{\frac{h}{2}} \frac{1}{8} \bullet (4 \cdot t^2 - 2 \cdot t \cdot h) dt$$

$$= \frac{1}{8} \bullet (\frac{4}{3} \cdot t^3 - t^2 \cdot h) \Big|_0^{\frac{h}{2}} = \frac{1}{8} \bullet (\frac{h^3}{6} - \frac{h^3}{4}) = -\frac{h^3}{96}$$

So,

$$\frac{E_T}{E_T'} = \frac{\int_0^{\frac{h}{2}} H_1 dt \cdot F_1 - \int_0^{\frac{h}{2}} H_1 dt \cdot F_2 + \int_{\frac{h}{2}}^{h} H_1 dt \cdot F_2}{\int_0^{\frac{h}{2}} H_2 dt \cdot F_1}$$

$$= \frac{\int_0^{\frac{h}{2}} H_1 dt \cdot F_1}{\int_0^{\frac{h}{2}} H_2 dt \cdot F_1} + \frac{-\int_0^{\frac{h}{2}} H_1 dt \cdot F_2 + \int_{\frac{h}{2}}^{h} H_1 dt \cdot F_2}{\int_0^{\frac{h}{2}} H_2 dt \cdot F_1}$$

$$\approx 4 + \frac{-\int_0^{\frac{h}{2}} H_1 dt \cdot \frac{F_1}{2} + \int_{\frac{h}{2}}^{h} H_1 dt \cdot \frac{F_1}{2}}{\int_0^{\frac{h}{2}} H_2 dt \cdot F_1}$$

$\approx$ is used here since $F_2 \approx \frac{F_1}{2}$

$$= 4 + \frac{-\int_0^{\frac{h}{2}} H_1 dt + \int_{\frac{h}{2}}^{h} H_1 dt}{2\int_0^{\frac{h}{2}} H_2 dt} = 4 + \frac{\frac{h^3}{24} - \frac{h^3}{24}}{2 \cdot (-\frac{h^3}{96})} = 4$$

As a result, when the partitions are increased from *n* to *2n* pieces, the error decreases by a factor of 4.

$$\therefore \frac{E_T}{E_T'} \approx 4$$

## **Proving the relationship:**

In following equations, $a_i$ represents $x$-coordinate with $n$ trapezoid partitions. $a'_i$ represents $x$-coordinate with $2n$ trapezoid partitions.

$$\overline{d} = \frac{1}{n}\sum_{i=1}^{i=n} d_i$$

$$\because d_i = 0.5 \cdot (f(a_i) + f(a_{i-1})) - f(a_{i-1} + 0.5 \cdot h)$$

$$\therefore \overline{d} = \frac{1}{n} \cdot \sum_{i=0}^{i=n-1} (0.5 \cdot (f(a_{i+1}) + f(a_i)) - f(a_i + 0.5 \cdot h))$$

$$= \frac{1}{n}[(0.5 \cdot f(a_0) + f(a_1) + f(a_2) + \ldots + f(a_{n-1}) + 0.5 \cdot f(a_n)) - \sum_{i=0}^{i=n-1} f(a_i + 0.5 \cdot h)]$$

$$\because A_{Trapezoid} = h \times (\frac{1}{2}f(a_0) + f(a_1) + f(a_2) + \cdots + f(a_{n-1}) + \frac{1}{2}f(a_n))$$

$$\therefore \overline{d} = \frac{1}{n} \cdot \frac{A_{Trapezoid}}{h} - \frac{1}{n}(f(a_0 + \frac{h}{2}) + f(a_1 + \frac{h}{2}) + \ldots + f(a_{n-1} + \frac{h}{2}))$$

The terms $f(a_0 + \frac{h}{2}) + f(a_1 + \frac{h}{2}) + \ldots + f(a_{n-1} + \frac{h}{2})$ can be linked to $A'_{Trapezoid}$:

$$A'_{Trapezoid} = h \times (\frac{1}{2}f(a'_0) + f(a'_1) + f(a'_2) + \cdots + f(a'_{2n-1}) + \frac{1}{2}f(a'_{2n}))$$

$$\because f(a_0 + \frac{h}{2}) + f(a_2 + \frac{h}{2}) + \ldots + f(a_{n-1} + \frac{h}{2}) = f(a'_1) + f(a'_3) + \ldots + f(a'_{2n-1})$$

$$\therefore \overline{d} = \frac{1}{n} \cdot \frac{A_{Trapezoid}}{h} - \frac{2}{2n}(f(a'_1) + f(a'_3) + \ldots + f(a'_{2n-1}) + 0.5 \cdot f(a_0) + f(a_1) +$$

$$f(a_2) + \ldots + f(a_{n-1}) + 0.5 \cdot f(a_n)) + \frac{1}{n}(0.5 \cdot f(a_0) + f(a_1) + \ldots + f(a_{n-1}) + 0.5 \cdot f(a_n))$$

$$= \frac{1}{n} \cdot \frac{A_{Trapezoid}}{h} - \frac{2}{2n}(f(a'_1) + f(a'_3) + \ldots + f(a'_{2n-1}) + 0.5 \cdot f(a'_0) + f(a'_2) +$$

$$f(a'_4) + \ldots + f(a'_{2n-2}) + 0.5 \cdot f(a'_{2n})) + \frac{1}{n}(0.5 \cdot f(a_0) + f(a_1) + \ldots + f(a_{n-1}) + 0.5 \cdot f(a_n))$$

$$= \frac{2}{n} \cdot \frac{A_{Trapezoid}}{h} - \frac{2}{2n} \cdot \frac{A'_{Trapezoid}}{0.5h} = 2 \cdot (\frac{A_{Trapezoid}}{nh} - \frac{A'_{Trapezoid}}{2n \cdot 0.5h})$$

$$\bar{d} \cdot (b-a) = 2 \cdot (\frac{A_{Trapezoid}}{nh} - \frac{A'_{Trapezoid}}{nh}) \cdot (b-a) = 2 \cdot (A_{Trapezoid} - A'_{Trapezoid})$$
$$= 2 \cdot (\int_a^b f(x)dx - A'_{Trapezoid} - (\int_a^b f(x)dx) - A_{Trapezoid}))$$
$$= 2 \cdot (E'_T - E_T)$$
$$\because \frac{E_T}{E_T'} \approx 4$$
$$\therefore \bar{d} \cdot (b-a) = 2 \cdot (\frac{1}{4}E_T - E_T) = 2 \cdot (-\frac{3}{4}E_T)$$
$$\bar{d} \cdot (b-a) = 2 \cdot (-\frac{3}{4}E_T)$$
$$\frac{2}{3} \cdot \bar{d} \cdot (b-a) = -E_T$$

## Application of $E_T \approx \frac{2}{3}(b-a) \cdot \bar{d}$:

From the previous results, general steps to use trapezoidal rule is listed below:

1. For a given number of partitions, trapezoidal rule is used to approximate the definite integral.
2. Calculate $\bar{d}$.
3. Calculate $E_T$ using $E_T = -\frac{2}{3}(b-a) \cdot \bar{d}$
4. Repeat steps 1-3 for $n = 10$ to $n = 10^5$ $or$ $10^6$.
5. Using results of $E_T$ to approximate definite integral to a certain order of magnitude. Compare the last 2 rows of results to give the final evaluation.

Examples:

1. $\int_a^b e^{x^2} dx$

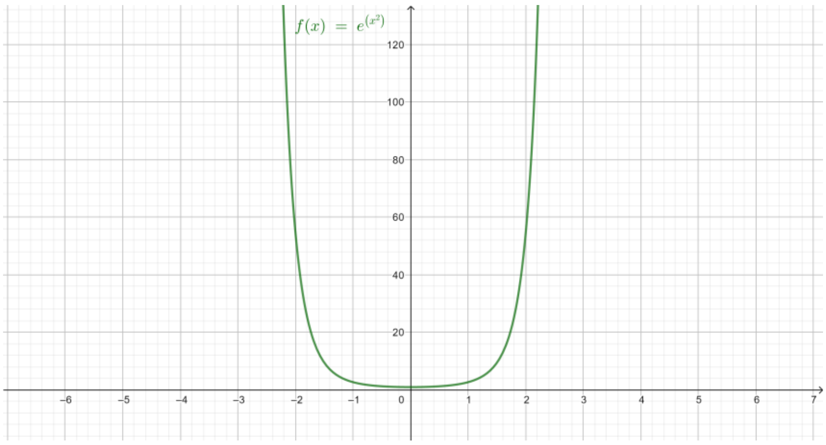



Figure 4. Graph of the function

Case 1: Let $a = 5, b = 9$:

Table 6. Computed values for different values of $n$

| $n$ | Trapezoidal approximation | Value corrected for error |
|---|---|---|
| 10 | 3.01747913548866E+34 | 1.12016167742544E+34 |
| $10^2$ | 8.77816694270312E+33 | 8.42062668040462E+33 |
| $10^3$ | 8.4234583898377E+33 | 8.41984415383871E+33 |
| $10^4$ | 8.41988022063325E+33 | 8.41984407433742E+33 |
| $10^5$ | 8.41984443579289E+33 | 8.41984407432946E+33 |
| $10^6$ | 8.41984407794411E+33 | 8.4198440743295E+33 |

Compare the last 2 rows, the result is correct to 13 figures, $8.419844074329 \times 10^{33}$.

The formula $E_T = -\frac{2}{3}(b-a) \cdot \overline{d}$ gives a more accurate approximation for the result than with the error bound formula.

Evaluate $\int_a^b e^{x^2} dx$ using $|E_T| \leq \frac{M \cdot (b-a) \cdot h^2}{12}$ for $a = 5, b = 9, n = 6$

$$\because f(x) = e^{x^2}$$

$$\therefore f''(x) = 2e^{x^2} + 4x^2 e^{x^2}$$

$$|E_T| \leq \frac{M \cdot (b-a) \cdot h^2}{12} = \frac{(2e^{9^2} + 4 \times (9^2) \times e^{9^2}) \cdot 5 \cdot (\frac{4}{10^6})^2}{12}$$

$$= 2.618601198 \times 10^{26}$$

Since $f(x)$ is above the *x-axis* and concave up, take into consideration for the error bound:

$$f(x) \approx 8.41984407794411 \times 10^{33} - 2.618601198 \times 10^{26}$$
$$= 8.4198438 \times 10^{33}$$

The result is correct to $8.41984 \times 10^{33}$, 6 figures, taken $10^6$ partitions. It is much less accurate than the result calculated using previous method.

Case 2: Let $a = 1, b = 2$:

Table 7 .Computed values for different values of *n*

| $n$ | Trapezoidal approximation | Value corrected for error |
|---|---|---|
| 10 | 15.1667841486261 | 14.9901395460705 |
| $10^2$ | 14.9917505872659 | 14.9899760360926 |
| $10^3$ | 14.9899937659298 | 14.9899760196017 |
| $10^4$ | 14.9899761970634 | 14.9899760196001 |
| $10^5$ | 14.9899760213746 | 14.9899760196 |

Compare the last 2 rows, the result is correct to 12 figures, 14.9899760196.

Likewise, this result is much more accurate than that calculated using error bound formula.

$$|E_T| \leq \frac{M \cdot (b-a) \cdot h^2}{12} = \frac{(2e^{2^2} + 4 \times (2^2) \times e^{2^2}) \cdot 1 \cdot (\frac{1}{10^6})^2}{12}$$
$$= 8.189722505 \times 10^{-9}$$

$$f(x) \approx 14.9899760213746 - 8.189722505 \times 10^{-9} = 14.98997601$$

The result is correct to $14.9899760$, 9 figures, taken $10^5$ partitions.

2. $\int_a^b \frac{\sin x}{x} dx$

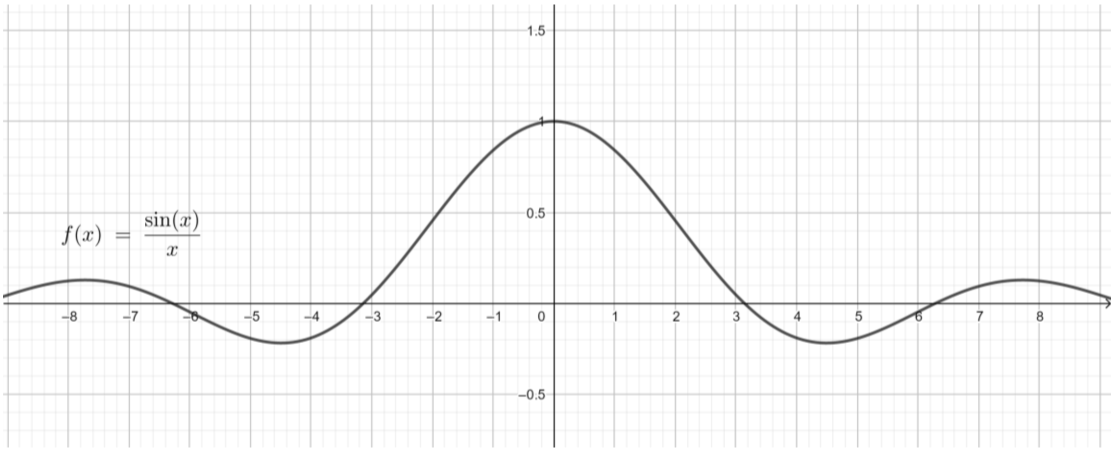



Figure 5. Graph of the function

Case 1: Let $a = 1, b = 2$:

Table 8. Computed values for different values of $n$

| $n$ | Trapezoidal approximation | Value corrected for error |
|---|---|---|
| 10 | 0.659218040541651 | 0.659329908514041 |
| $10^2$ | 0.65932878785888 | 0.65932990643572 |
| $10^3$ | 0.659329895249754 | 0.659329906435512 |
| $10^4$ | 0.659329906323654 | 0.659329906435511 |
| $10^5$ | 0.659329906434394 | 0.659329906435513 |

Compare the last 2 rows, the result is correct to 13 figures, 0.6593299064355.

Case 2: Let $a = 0.1, b = \pi$:

Table 9. Computed values for different values of $n$

| $n$ | Trapezoidal approximation | Value corrected for error |
|---|---|---|
| 10 | 1.74979408696729 | 1.75199290280498 |
| $10^2$ | 1.75197061818439 | 1.75199259090534 |
| $10^3$ | 1.75199237114853 | 1.75199259087419 |
| $10^4$ | 1.75199258867694 | 1.75199259087419 |
| $10^5$ | 1.75199259085225 | 1.75199259087422 |

Compare the last 2 rows, the result is correct to 13 figures, 1.751992590874.

## Conclusion:

In this investigation, I formulated and justified the relationship between the trapezoidal rule error and the difference $\overline{d}$ on the midpoint of the interval. That is:

$$E_T = -\frac{2}{3}(b-a)\cdot\overline{d}$$

Compared with the error bound formula, the above equation can calculate the error directly to a greater precision as showed using same number of partition intervals. The minus sign also indicates whether the trapezoidal result is smaller than or greater than the actual result. This also can be applied to calculators and numerical integration programs to calculate the integrals to a certain precision. One weakness of my equation is that the calculation involves the calculation of $\overline{d}$, which has more computation steps than the error bound formula. This weakness would not be significant when partition intervals are small. Computers can carry out billions of computations in one second. However, the weakness will be significant when partition intervals increase.

Although the error can be reduced by making $\overline{d}$ smaller using more partition intervals, cautions is still needed. This is because present computers carry out calculations to a certain precision. So, there is a limit to the precision of the error depending on the computing power. This is actually a limit to the precision of all the numerical methods. Also, the relationship between the number of intervals and precision of the equation can be further explored.

Another point to note is that the approximation of $F_2 \approx \frac{F_1}{2}$ to justify the equation above. This needs further justification beyond the level of this investigation.

# Works Cited

*Calculus: Graphical, Numerical, Algebraic*, by Ross L. Finney, 5th ed., Pearson Prentice Hall, 2016, pp. 314–319.

*Section 7.7 Deriving the Trapezoidal Rule Error*. math.ucsd.edu/~ebender/20B/77_Trap.pdf. Accessed 18th, Feb., 2021.

## Appendices:

### 1. MATLAB code:

### 1.1: Example of finding the trend:

```
format long e
filename='E:\exp0-3.xlsx';
a=double(0);
b=double(3);
h=double(0.0);
k=double(0.0);
sl=double(0.0);
err=double(0.0);

%para=6;

time=6;
aaa=zeros(10,20); %absolute value
bbb=zeros(10,20);  %difference
ddd=zeros(10,20);
eee=zeros(10,20);
fff=zeros(10,20);
ggg=zeros(10,20);

aa=double(aaa);
bb=double(bbb);
dd=double(ddd);
ee=double(eee);
ff=double(fff);
gg=double(ggg);


for para=1:time

tv=double(1./para.*(exp(b*para)-exp(a*para)));
disp(tv);


for i =1:7
    k=10.^i;
    h=(b-a)./k;
```

```
    ss=double(0.0);
    cc=double(0.0);
    for j=1:(k+1)
        evalue=double(exp(para.*(a+(j-1).*h)));
        ss=ss+evalue.*h;
        if j<(k+1)
            sl=(exp(para.*(a+j.*h))-evalue)./h;
            err=evalue+0.5*h*sl;
            cc=cc+err-exp(para.*(a+(j-1).*h+0.5.*h)); %sum of vertical
difference
        end
    end
    ss=ss-(exp(a*para)+exp(b*para)).*0.5.*h;
    %%ss=ss*h;
    aa(i,para)=ss;
    %%disp(ss);
    bb(i,para)=ss-tv;
    ee(i,para)=log10(bb(i,para));
    dd(i,para)=cc./k;
    ff(i,para)=log10(dd(i,para));
    gg(i,para)=bb(i,para)./dd(i,para);
end;
end;
xlswrite(filename,bb,1);
xlswrite(filename,dd,2);
xlswrite(filename,ee,3);
xlswrite(filename,ff,4);
xlswrite(filename,gg,5);
xlswrite(filename,aa,6);
```

1.2: Example of application of the trend:

```
format long g
filename='E:\testexp4.xlsx';
a=1;
b=2;
h=0;
k=0;
sl=0;


aa=zeros(10,1); %absolute value
bb=zeros(10,1);  %difference
dd=zeros(10,1);
ee=zeros(10,1);
for i =1:5
    k=10.^i;
    h=(b-a)./k;
    ss=0;
    cc=0;
    for j=1:(k+1)
        evalue=exp((a+(j-1).*h).^2);
        ss=ss+evalue.*h;
        if j<(k+1)
            sl=(exp((a+j.*h).^2)-evalue)./h;
            err=evalue+h*0.5*sl;
            cc=cc+err-exp((a+(j-1).*h+0.5.*h).^2);
        end
    end
    ss=ss-(exp(a.^2)+exp(b.^2)).*0.5.*h;
    %%ss=ss*h;
    aa(i,1)=ss;
    %%disp(ss);
    %bb(i,1)=tv-ss;
    dd(i,1)=cc./k;
    bb(i,1)=dd(i,1)*(b-a)*(2/3);
    ee(i,1)=aa(i,1)-bb(i,1);
end;
xlswrite(filename,aa,1);
xlswrite(filename,bb,2);
xlswrite(filename,dd,3);
xlswrite(filename,ee,4);
```

2. <u>Table of values:</u>

2.1 Function:

$$f(x) = cos(k \cdot x)$$

Let $a = 0,\ b = \frac{\pi}{12}$:

Table 10. Several computed values for *k*=2

| *k*=2 | | | |
|---|---|---|---|
| $n$ | $E_T$ | $\overline{d}$ | $\frac{E_T}{\overline{d}}$ |
| 10 | -5.71184E-05 | -0.000327268 | 0.174530931 |
| $10^2$ | -5.71158E-07 | -3.27249E-06 | 0.174532905 |
| $10^3$ | -5.71158E-09 | -3.27249E-08 | 0.174532927 |
| $10^4$ | -5.71157E-11 | -3.27249E-10 | 0.17453264 |
| $10^5$ | -5.69239E-13 | -3.27249E-12 | 0.173946928 |
| $10^6$ | -1.6126E-14 | -3.27272E-14 | 0.492739219 |

Table 11. Computed values of $\frac{E_T}{\overline{d}}$ for different values of *n* and *k*

| $\frac{E_T}{\overline{d}}$ | | $k$ | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 |
| $n$ | 10 | 0.174532427 | 0.174530931 | 0.174528439 | 0.17452495 | 0.174520463 | 0.174514978 |
| | $10^2$ | 0.17453292 | 0.174532905 | 0.17453288 | 0.174532845 | 0.174532801 | 0.174532746 |
| | $10^3$ | 0.174532929 | 0.174532927 | 0.174532924 | 0.174532924 | 0.174532924 | 0.174532923 |
| | $10^4$ | 0.174539797 | 0.17453264 | 0.174533116 | 0.174532677 | 0.174532688 | 0.174533347 |
| | $10^5$ | 0.178334171 | 0.173946928 | 0.174458069 | 0.174739592 | 0.174539014 | 0.174560473 |
| | $10^6$ | 0.432527907 | 0.492739219 | 0.122736128 | 0.210056168 | 0.131348176 | 0.189685802 |

2.2 Function:

$$f(x) = x^k$$

When $k = 1, f(x) = x$, so $E_T$ is 0.

Let $a = 3,\ b = 5$:

Table 12. Several computed values for *k*=2

| $k$=3 | | | |
|---|---|---|---|
| $n$ | $E_T$ | $\overline{d}$ | $\frac{E_T}{\overline{d}}$ |
| 10 | 1.30656 | 0.12 | 1.333333333 |
| $10^2$ | 0.013066656 | 0.0012 | 1.333333333 |
| $10^3$ | 0.000130667 | 1.2E-05 | 1.333333336 |
| $10^4$ | 1.30667E-06 | 1.2E-07 | 1.333336488 |
| $10^5$ | 1.30693E-08 | 1.20001E-09 | 1.332373855 |
| $10^6$ | 1.32786E-10 | 1.19932E-11 | 1.338953121 |

Table 13. Computed values of $\frac{E_T}{\overline{d}}$ for different values of *n* and *k*

| $\frac{E_T}{\overline{d}}$ | | $k$ | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 |
| $n$ | 10 | 0 | 1.333333333 | 1.33333333 | 1.33336055 | 1.33341179 | 1.333484527 |
| | $10^2$ | 0 | 1.333333333 | 1.33333333 | 1.33333361 | 1.33333412 | 1.333334845 |
| | $10^3$ | 0 | 1.333333317 | 1.33333334 | 1.33333334 | 1.33333335 | 1.333333347 |
| | $10^4$ | 0 | 1.333340507 | 1.33333649 | 1.33333498 | 1.33333452 | 1.333334183 |
| | $10^5$ | 0 | 1.332953463 | 1.33237386 | 1.33359566 | 1.33344547 | 1.333102775 |
| | $10^6$ | 0 | 2.603490777 | 1.33895312 | 1.35548721 | 1.2676491 | 1.277765952 |