6.2 Trapezoidal Rule

As simplest polynomial interpolation, consider a linear approximation, i.e. the secant through the points $(a; f(a))$ and $(b; f(b))$. Then, to approximate the area under the curve, one calculates the area of the trapezoid with vertices $(a; 0)$, $(b; 0)$, $(b; f(b))$, $(a; f(a))$, which obviously is

$$T(f) = (b-a)\frac{f(a)+f(b)}{2}\,. \tag{6.3}$$

If one splits the full interval $[a, b]$ into n subintervals of equal length, one can apply the trapezoidal rule for each of the subintervals, which significantly improves the approximation. To this end, we define

$$\begin{aligned} h_n &= \frac{b-a}{n}, \\ x_0 &= a,\ x_1 = a + h_n,\ x_2 = a + 2h_n,\ \ldots,\ x_n = b. \end{aligned} \tag{6.4}$$

This allows to write

$$\begin{aligned} T_n(f) &= (x_1 - x_0)\frac{f(x_0)+f(x_1)}{2} + (x_2 - x_1)\frac{f(x_1)+f(x_2)}{2} + (x_3 - x_2)\frac{f(x_2)+f(x_3)}{2} \\ &\quad + \ldots + (x_n - x_{n-1})\frac{f(x_{n-1})+f(x_n)}{2} \\ &= h_n\frac{f(x_0)+f(x_1)}{2} + h_n\frac{f(x_1)+f(x_2)}{2} + \ldots + h_n\frac{f(x_{n-1})+f(x_n)}{2} \\ &= \frac{h_n}{2}\left[f(x_0) + 2f(x_1) + 2f(x_2) + \ldots + 2f(x_{n-1}) + f(x_n)\right]. \end{aligned} \tag{6.5}$$

Theorem 1: *error of the trapezoidal rule* (without proof)
Let f be twice continuously differentiable on $[a, b]$. For the approximation of the integral of f over $[a, b]$, $I(f) = \int_a^b f(x)\mathrm{d}x$, by the trapezoidal rule for n sub intervals, $T_n(f)$, it holds that

$$|I(f) - T_n(f)| = n\,\frac{h_n^3}{12}|f''(\epsilon_n)| = \frac{1}{12}\,\frac{(b-a)^3}{n^2}|f''(\epsilon_n)|; \quad \epsilon_n \in [a, b]. \tag{6.6}$$

It follows from Theorem 1 that doubling the number of sub intervals approximately quarters the absolute error of the trapezoidal rule.

Example:
Consider the following analytically determined integral of the function $f(x) = \frac{1}{1+x}$:

$$I(f) = \int\limits_0^1 \frac{\mathrm{d}x}{1+x} = \ln(1+x)\Big|_0^1 = \ln(2) - \ln(1) = \ln(2) \approx 0.69314718. \tag{6.7}$$

Applying the trapezoidal rule directly to this interval ($a = 0$, $b = 1$) yields

$$T_1(f) = (1-0)\frac{\frac{1}{1+0} + \frac{1}{1+1}}{2} = \frac{3}{4} = 0.75. \tag{6.8}$$

This result has an absolute error of about $|I(f) - T_1(f)| \approx 0.057$ (relative error ≈ 8 %). To improve the approximation, we divide the interval $[a, b]$ into two subintervals, by which the error should roughly be reduced by a factor of $1/4$:

$$\begin{aligned} T_2(f) &= \left(\frac{1}{2} - 0\right)\frac{\frac{1}{1+0} + \frac{1}{1+\frac{1}{2}}}{2} + \left(1 - \frac{1}{2}\right)\frac{\frac{1}{1+\frac{1}{2}} + \frac{1}{1+1}}{2} \\ &= \frac{1}{2}\left(\frac{1+\frac{2}{3}}{2} + \frac{\frac{2}{3}+\frac{1}{2}}{2}\right) = \frac{1}{2}\left(\frac{5}{6} + \frac{7}{12}\right) = 0.708\bar{3} \end{aligned} \tag{6.9}$$

This second approximation has an absolute error of $|I(f) - T_2(f)| \approx 0.015$ (rel. error ≈ 2 %), which indeed is roughly one quarter of the error of the first approximation (even a little less).