Integrals

- Problem:

  given a continuous function $f(x)$ over an interval $[a,b]$, find its definite integral:

$$I = \int_a^b f(x)dx$$

- The value of I is NOT possible to be computed every time, it depends on the function $f(x)$.

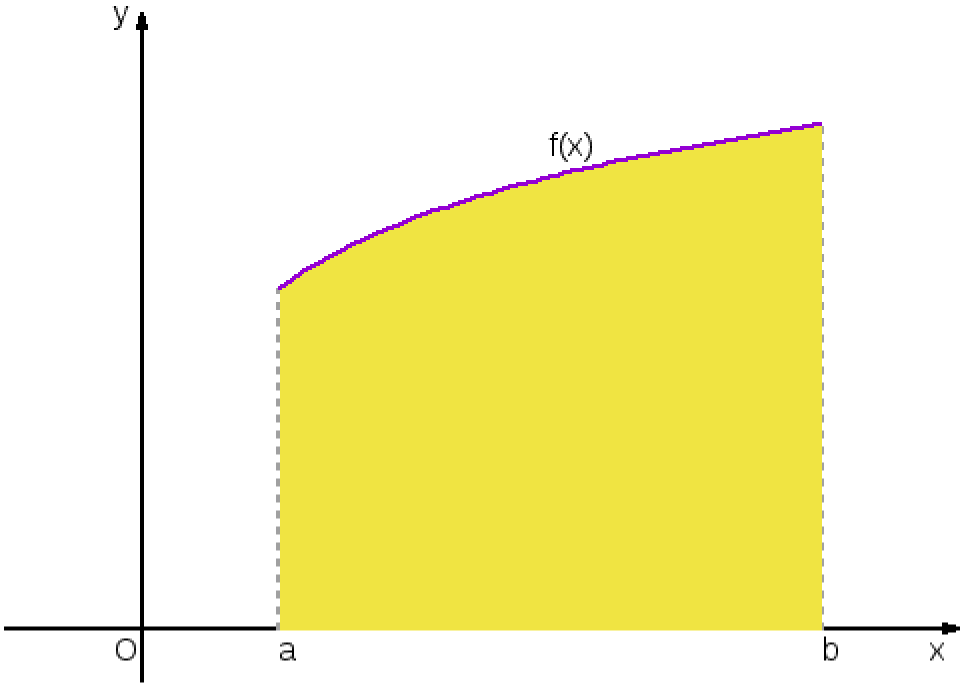
  Example: $$I = \int_0^z \frac{\sin x}{x} dx$$

Integrals

- There are two possible distinct cases:
  - The function $f(x)$ is known, so its values can be calculated for each x ;
  - The function $f(x)$ is unknown, but its values are known at some grid-points x_j (for instance, one has a set of tabulated values coming from some experiment)
- If we know how to solve the second case, we can always solve the first one, since we can always choose some points in $[a,b]$ which are representative of the function's behavior.

Integrals

- Geometrically, the computation of I is equivalent to assess the area below the curve representing $f(x)$:



Integrals

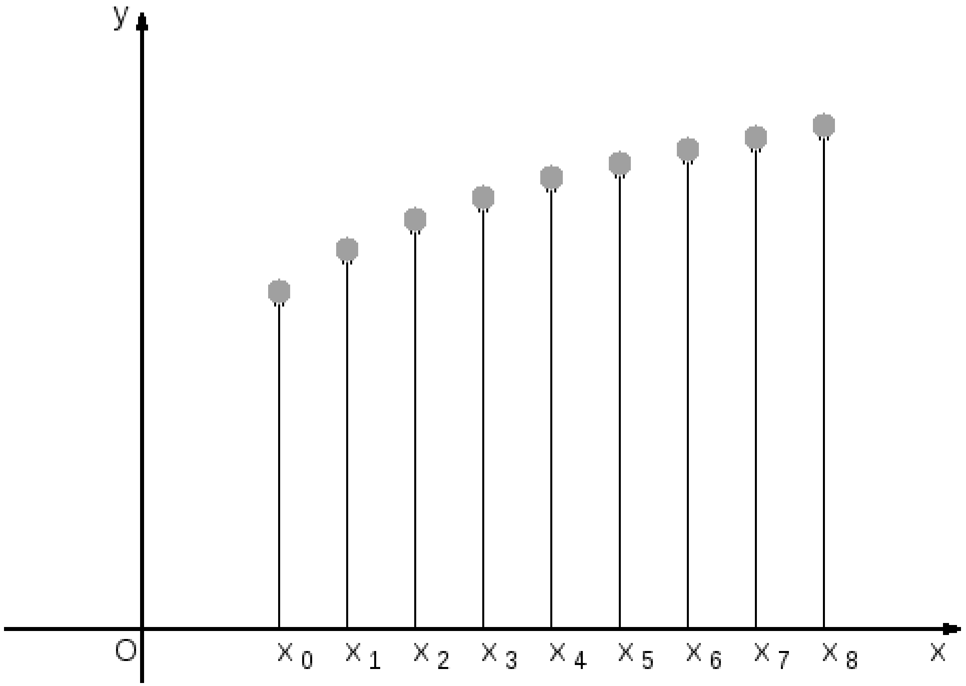
- When we know the values of the function $f(x)$ on $n+1$ discrete points x_j (not necessarily equidistant one from the other!) we could think of approximating the function with a polynomial of degree equal to n, passing from the same values $f(x_j)$:

$$f(x) \simeq \Pi_n(x) \quad \Rightarrow \quad I \simeq \int_a^b \Pi_n(x)dx$$

where: $$f(x_j) = \Pi_n(x_j) \quad \forall j = 0, \ldots, n$$

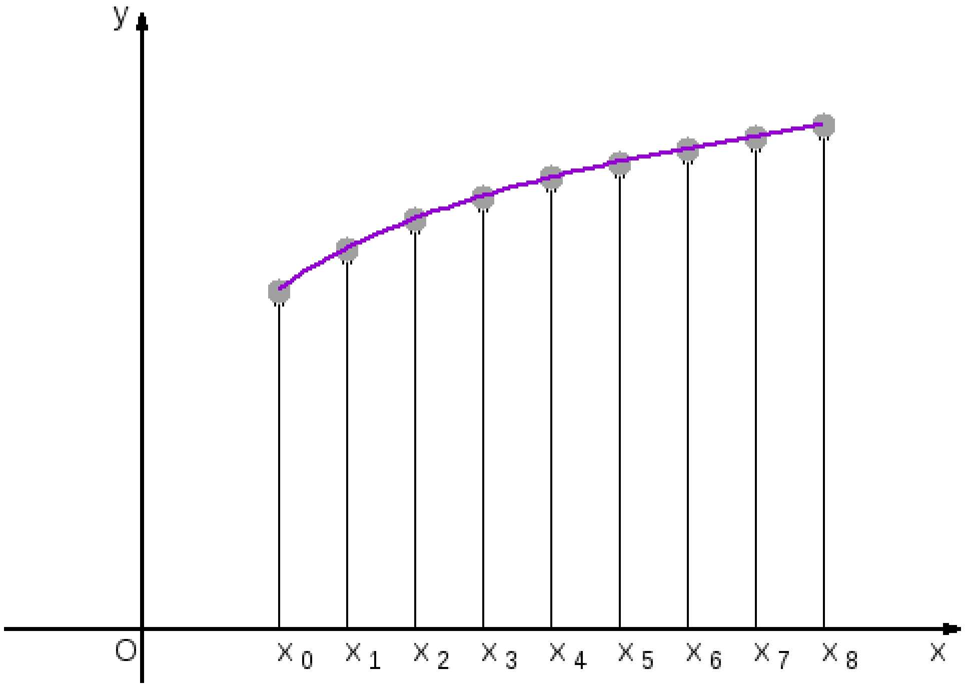
Integrals

- The philosophy is that we always know how to integrate a polynomial function of any degree, therefore if we found an approximation of the values with a polynomial we would solve the problem:



Integrals

- The philosophy is that we always know how to integrate a polynomial function of any degree, therefore if we found an approximation of the values with a polynomial we would solve the problem:



Integrals

- For instance, in the previous case (n = 8), we have:

$$\Pi_8(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_7 x^7 + a_8 x^8$$

where the nine coefficients a_i are such that:

$$\Pi_8(x_i) = f(x_i)$$

and then:

$$I = \int_a^b f(x)dx \simeq \int_a^b \Pi_8(x)dx =$$

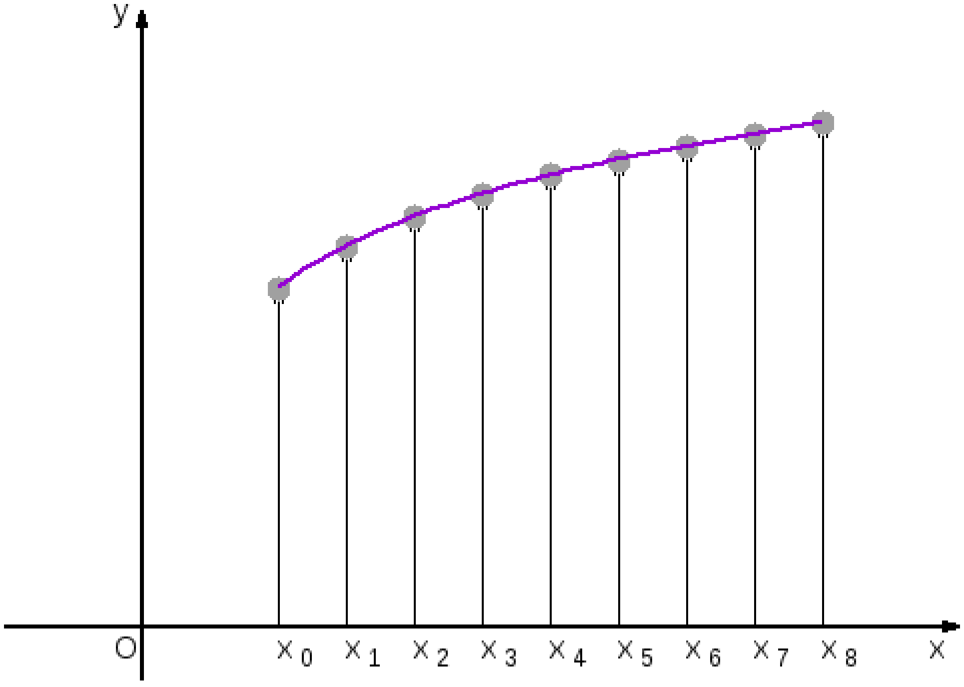
$$= \left[a_0 x + a_1 \frac{x^2}{2} + a_2 \frac{x^3}{3} + \ldots a_7 \frac{x^8}{8} + a_8 \frac{x^9}{9} \right]_a^b$$

Integrals

- Unfortunately, this technique only works for **small values** of n. As soon as n increases, the polynomial approximation starts to show strong oscillations in between subsequent points x_i and the evaluation of the integral is wrong!
- However, we can argue that, instead of searching a unique polynomial approximation for the whole set of n+1 points xi, we can search different low-degree polynomials for each subinterval $[x_i,x_{i+1}]$, or couples or triplets, etc., of subintervals.

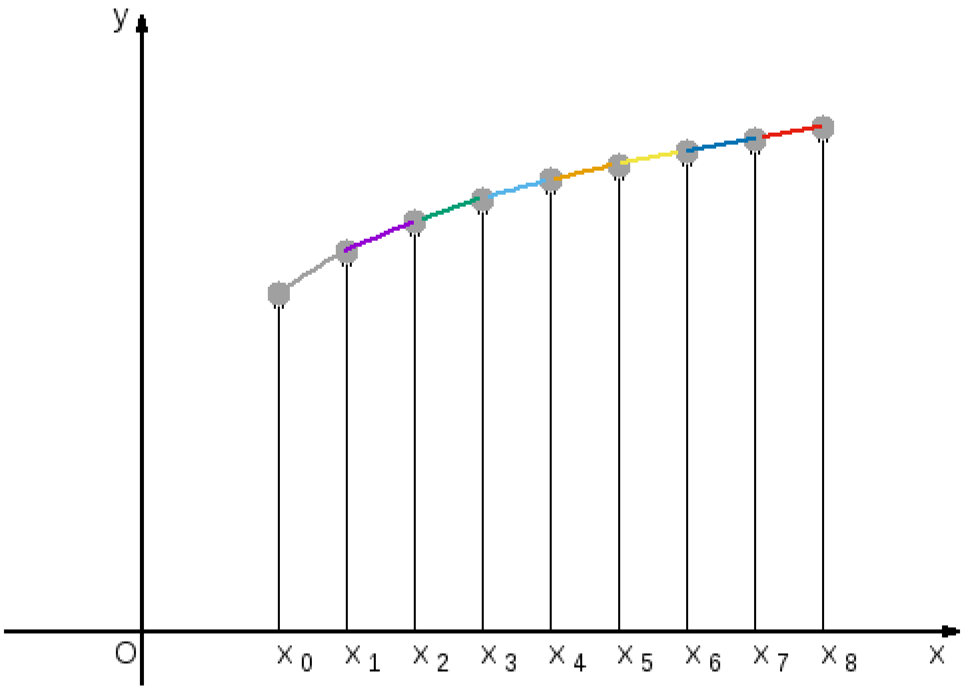
Integrals

- Instead of a 8-degree polynomial on 8 intervals:



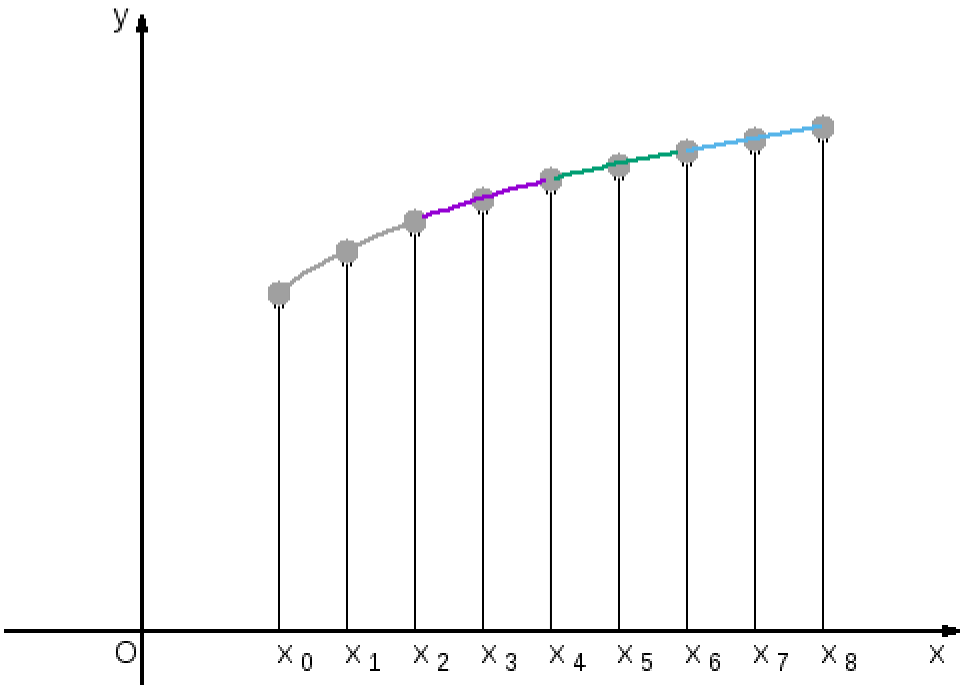
Integrals

- We can use a 1-degree polynomial on 1 interval:



Integrals

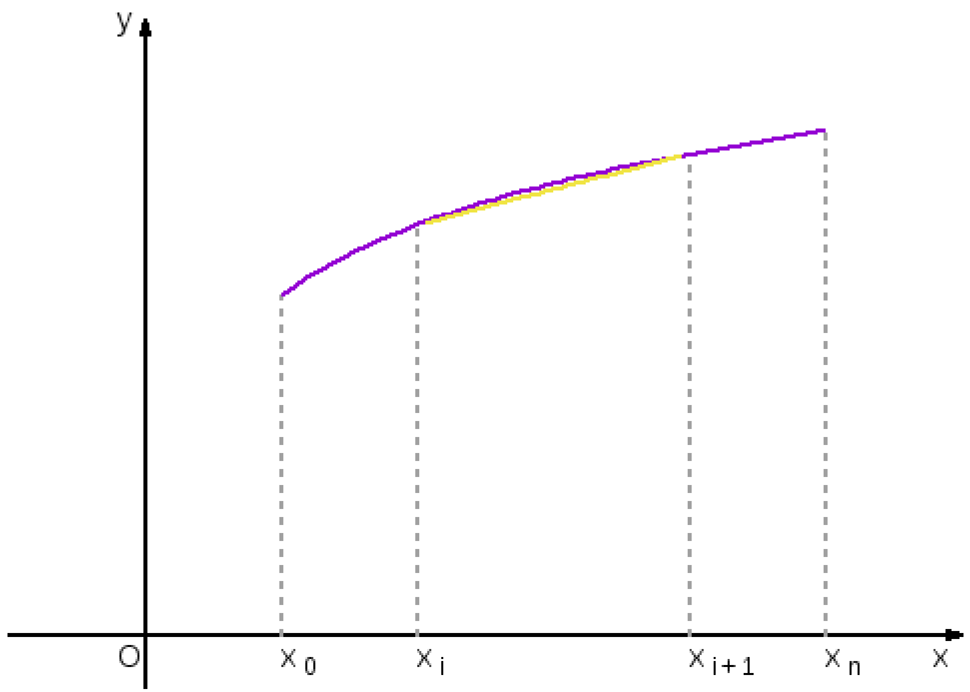
- Or a 2-degree polynomial on 2 adjacent intervals:



and so on...

Trapezoidal rule

- Let's study the first case: a first-degree polynomial in each subinterval $[x_i, x_{i+1}]$:



Trapezoidal rule

- The equation of the polynomial (straight line) passing through $f(x_i)$ and $f(x_{i+1})$ is:

$$\Pi_1(x) = f(x_i) + \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}(x - x_i)$$

- By letting: $h_i = x_{i+1} - x_i$ one has:

$$I_{[x_i, x_{i+1}]} \simeq \int_{x_i}^{x_{i+1}} \Pi_1(x)dx =$$

$$= \int_{x_i}^{x_{i+1}} f(x_i)dx + \frac{f(x_{i+1}) - f(x_i)}{h_i} \int_{x_i}^{x_{i+1}} (x - x_i)dx$$

Trapezoidal rule

$$I_{[x_i,x_{i+1}]} \simeq f(x_i)h_i + \frac{f(x_{i+1})-f(x_i)}{h_i}\left[\frac{x^2}{2} - x_i x\right]_{x_i}^{x_{i+1}} =$$
$$= f(x_i)h_i + \frac{f(x_{i+1})-f(x_i)}{h_i}\left[\frac{x_{i+1}^2}{2} - x_i x_{i+1} - \frac{x_i^2}{2} + x_i^2\right]$$
$$= \frac{f(x_{i+1})+f(x_i)}{2}h_i$$

which is actually the area of the trapezium having $f(x_{i+1})$ and $f(x_i)$ as major and minor bases, respectively and height equal to h_i!

Trapezoidal rule

- Then we get the approximation for the integral I by adding up all the contributions of the single trapeziums:

$$I = \sum_{i=0}^{n-1} \frac{f(x_{i+1}) + f(x_i)}{2} h_i$$

- If all the subintervals have the same width: h_i=h, we get:

$$I = \frac{h}{2} \sum_{i=0}^{n-1} [f(x_{i+1}) + f(x_i)] = \frac{h}{2} \{[f(x_0) + f(x_1)] +$$

$$+ [f(x_1) + f(x_2)] + \ldots + [f(x_{n-2}) + f(x_{n-1})] + [f(x_{n-1}) + f(x_n)]\}$$

Trapezoidal rule

- Except for the terms $f(x_0)$ and $f(x_n)$, all the others are added two times, that is:

$$I \simeq h \left[\frac{f(x_0) + f(x_n)}{2} + \sum_{i=1}^{n-1} f(x_i) \right]$$

which is the so-called "**trapezoidal rule**":

The approximation of the integral is obtained as the average of the values of the function at the first and last point plus the sum of the values on all internal points.

Error for the trapezoidal rule

- It is important to assess the error that we commit when we use the trapezoidal rule, at least as a magnitude order:

$$\mathcal{E} = I - I^{\mathrm{trap}}$$

- Of course, we cannot evaluate I in a exact way (otherwise we knew the solution of the problem) but we can use the Taylor's expansion to have an idea of how it depends on h:

$$f(x+h) = f(x) + h\,\frac{df}{dx}\bigg|_x + \frac{h^2}{2!}\,\frac{d^2 f}{dx^2}\bigg|_x + \frac{h^3}{3!}\,\frac{d^3 f}{dx^3}\bigg|_x + \ldots$$

Error for the trapezoidal rule

- We start by evaluating the error on the single subinterval $[x_i, x_{i+1}]$:

$$\mathcal{E}_{[x_i,x_{i+1}]} = I_{[x_i,x_{i+1}]} - I^{(\mathrm{trap})}_{[x_i,x_{i+1}]}$$

$$I_{[x_i,x_{i+1}]} = \int_{x_i}^{x_{i+1}} f(x)dx = \int_{x_i}^{x_{i+1}} \Big\{ f(x_i) + f'(x_i)(x-x_i) + f''(x_i)\frac{(x-x_i)^2}{2!} + \\ + f'''(x_i)\frac{(x-x_i)^3}{3!} + f^{IV}(x_i)\frac{(x-x_i)^4}{4!} + O[(x-x_i)^4] \Big\} dx$$

- By changing the variable: x-x_i = z; dx = dz,

$$I_{[x_i,x_{i+1}]} = \int_0^h \Big\{ f(x_i) + f'(x_i)z + f''(x_i)\frac{z^2}{2!} + f'''(x_i)\frac{z^3}{3!} + f^{IV}(x_i)\frac{z^4}{4!} + O[z^4] \Big\} dz = \\ = f(x_i)h + f'(x_i)\frac{h^2}{2} + f''(x_i)\frac{h^3}{6} + f'''(x_i)\frac{h^4}{24} + f^{IV}(x_i)\frac{h^5}{120} + O[h^5]$$

Error for the trapezoidal rule

- On the other hand, we have:

$$I^{(\mathrm{trap})}_{[x_i,x_{i+1}]} = \frac{f(x_i)+f(x_{i+1})}{2}\cdot h =$$

$$= \frac{h}{2}\left\{f(x_i)+f(x_i)+f'(x_i)h+f''(x_i)\frac{h^2}{2!}+f'''(x_i)\frac{h^3}{3!}+f^{IV}(x_i)\frac{h^4}{4!}+O[h^4]\right\} =$$

$$= hf(x_i)+\frac{h^2}{2!}f'(x_i)+\frac{h^3}{4}f''(x_i)+\frac{h^4}{12}f'''(x_i)+\frac{h^5}{48}f^{IV}(x_i)+O[h^5]$$

- By subtracting the two formulas:

$$\mathcal{E}_{[x_i,x_{i+1}]} = I_{[x_i,x_{i+1}]} - I^{(\mathrm{trap})}_{[x_i,x_{i+1}]} =$$

$$= h^3 f''(x_i)\left[\frac{1}{6}-\frac{1}{4}\right] + h^4 f'''(x_i)\left[\frac{1}{24}-\frac{1}{12}\right] + h^5 f^{IV}(x_i)\left[\frac{1}{120}-\frac{1}{48}\right] + O(h^5) =$$

$$= -\frac{1}{12}h^3 f''(x_i) - \frac{1}{24}h^4 f'''(x_i) - \frac{1}{80}h^5 f^{IV}(x_i) + O(h^5)$$

Error for the trapezoidal rule

- Supposing that h << 1, the dominant term in the error is:

$$\mathcal{E}_{[x_i,x_{i+1}]} \simeq -\frac{1}{12}h^3 f''(x_i)$$

- The global error on the whole [a,b] interval is found by adding all the errors on each subinterval:

$$\mathcal{E} = \sum_{i=0}^{n-1} \mathcal{E}_{[x_i,x_{i+1}]} \simeq \sum_{i=0}^{n-1} -\frac{1}{12}h^3 f''(x_i) =\simeq -\frac{1}{12}h^3 \sum_{i=0}^{n-1} f''(x_i)$$

Error for the trapezoidal rule

- By definition, the average of n quantities $f(x_0)$, $f(x_1)$, …, $f(x_{n-1})$ is:

$$\bar{f} = \frac{1}{n}\sum_{i=0}^{n-1} f(x_i)$$

that is, previous formula becomes:

$$\mathcal{E} \simeq -\frac{1}{12} h^3 n \bar{f}''$$

- Since: $$h = \frac{(b-a)}{n} \quad \Rightarrow \quad n = \frac{(b-a)}{h}$$

Error for the trapezoidal rule

- We finally obtain:

$$\mathcal{E} \simeq -\frac{1}{12} h^2 (b-a) \bar{f''}$$

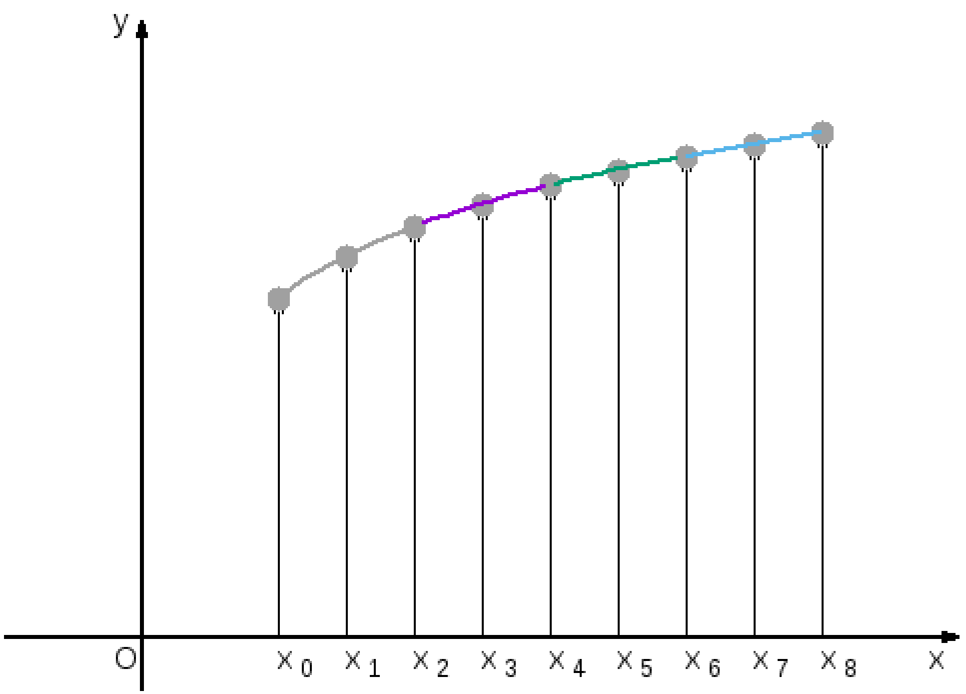
that is, the leading global error for the trapezoidal rule is proportional to h^2!

- In the same manner, one can show that the total global error is given by:

$$\mathcal{E} = -\frac{1}{12} h^2 (b-a) \bar{f''} - \frac{1}{24} h^3 (b-a) \bar{f'''} - \frac{1}{80} h^4 (b-a) f^{\bar{I}V} + O(h^4)$$

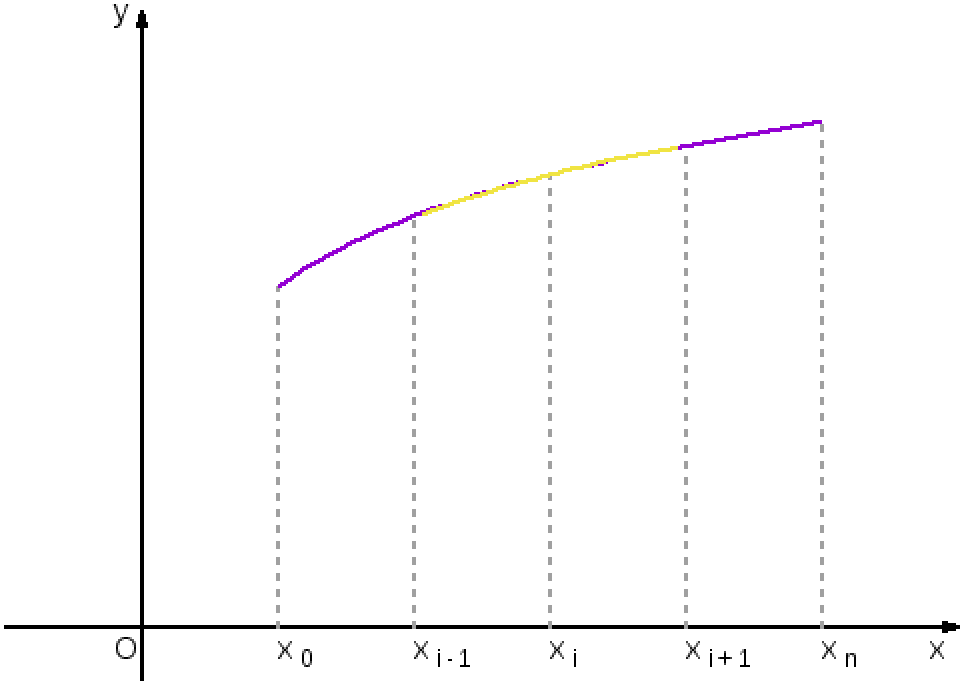
Cavalieri-Simpson's rule

- Is it possible to decrease the error? One could try to improve the situation by approximating the function with a second order polynomial on two consecutive intervals:



Cavalieri-Simpson’s rule

- In this case, since we have to consider couples of intervals, the total number n of intervals has to be **even**!



Cavalieri-Simpson's rule

- The equation of the second order polynomial passing through $f(x_{i-1})$, $f(x_i)$ and $f(x_{i+1})$ is:

$$\Pi_2(x) = ax^2 + bx + c$$

where the coefficients are still to be determined!

- By letting: $h = x_{i+1} - x_i = x_i - x_{i-1}$ one has:

$$I_{[x_{i-1},x_{i+1}]} \simeq \int_{x_{i-1}}^{x_{i+1}} \Pi_2(x)dx = \int_{x_i-h}^{x_i+h} (ax^2 + bx + c)dx =$$

$$= \left[\frac{ax^3}{3} + \frac{bx^2}{2} + cx\right]_{x_i-h}^{x_i+h} = \frac{a}{3}\left(2h^3 + 6x_i^2 h\right) + \frac{b}{2}4x_i h + 2ch$$

Cavalieri-Simpson's rule

- Now we have to determine a, b and c by imposing that the polynomials passes through the same points as the initial function $f(x)$:

$$\begin{aligned}
\Pi_2(x_{i-1}) &= a(x_i - h)^2 + b(x_i - h) + c = \\
&= ax_i^2 - 2ahx_i + ah^2 + bx_i - bh + c = f(x_{i-1}) \\
\Pi_2(x_i) &= ax_i^2 + bx_i + c = f(x_i) \\
\Pi_2(x_{i+1}) &= a(x_i + h)^2 + b(x_i + h) + c = \\
&= ax_i^2 + 2ahx_i + ah^2 + bx_i + bh + c = f(x_{i+1})
\end{aligned}$$

Cavalieri-Simpson's rule

- If we add up the 1st and 3rd equations and subtract the 2nd one multiplied by 2:

$$f(x_{i+1}) + f(x_{i-1}) - 2f(x_i) = 2ah^2 \quad \Rightarrow$$

$$\Rightarrow \quad a = \frac{f(x_{i+1}) + f(x_{i-1}) - 2f(x_i)}{2h^2} = \frac{D_2}{2}$$

where we let: $D_2 = \dfrac{f(x_{i+1}) + f(x_{i-1}) - 2f(x_i)}{h^2}$

- In same way, if we subtract the 1st equation from the 3rd one:

Cavalieri-Simpson's rule

$$f(x_{i+1}) - f(x_{i-1}) = 4ahx_i + 2bh \quad \Rightarrow$$

$$\Rightarrow \quad b = \frac{f(x_{i+1}) - f(x_{i-1})}{2h} - 2ax_i$$

- Again, if we let:

$$D_1 = \frac{f(x_{i+1}) - f(x_{i-1})}{2h}$$

we get: $b = D_1 - x_i D_2$

- Finally, from the 2nd equation:

$$c = f(x_i) - ax_i^2 - bx_i = f(x_i) - x_i^2 \frac{D_2}{2} - x_i D_1 + x_i^2 D_2 =$$

$$= f(x_i) - x_i D_1 + \frac{1}{2} x_i^2 D_2$$

Cavalieri-Simpson's rule

- By substituting these quantities in the approximation for the integral:

$$
\begin{aligned}
I_{[x_{i-1},x_{i+1}]} &\simeq \frac{a}{3}\left(2h^3+6x_i^2h\right)+\frac{b}{2}4x_ih+2ch= \\
&= \frac{1}{3}\frac{D_2}{2}\left(2h^3+6x_i^2h\right)+2hx_i\left(D_1-x_iD_2\right)+ \\
&+2h\left(f\left(x_i\right)-x_iD_1+\frac{1}{2}x_i^2D_2\right)= \\
&= \frac{h}{3}\left[f\left(x_{i+1}\right)+4f\left(x_i\right)+f\left(x_{i-1}\right)\right]
\end{aligned}
$$

Cavalieri-Simpson's rule

- In order to obtain the value of the integral I we have to sum all contribution for each couple of subintervals:

$$
\begin{aligned}
I &= I_{[x_0,x_2]} + I_{[x_2,x_4]} + I_{[x_4,x_6]} + \ldots + I_{[x_{n-2},x_n]} = \\
&= \frac{h}{3}\left[f(x_0) + 4f(x_1) + f(x_2)\right] + \frac{h}{3}\left[f(x_2) + 4f(x_3) + f(x_4)\right] + \\
&+ \frac{h}{3}\left[f(x_4) + 4f(x_5) + f(x_6)\right] + \ldots + \frac{h}{3}\left[f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)\right] = \\
&= \frac{h}{3}\left[f(x_0) + f(x_n) + 2\sum_{\substack{i=2 \\ i\ \text{even}}}^{n-2} f(x_i) + 4\sum_{\substack{i=1 \\ i\ \text{odd}}}^{n-1} f(x_i)\right]
\end{aligned}
$$

Cavalieri-Simpson’s rule

- This is the final form of the Cavalieri-Simpson’s rule:

The value of the integral is given by one third of h multiplied by the sum of the first and last value of the function plus 2 times the sum of all the values of the function on the even internal points, plus 4 times the sum of all the values of the function on the odd internal points.

Cavalieri-Simpson's rule

- With an approach similar to that used for the trapezoidal rule, it is possible to show that the leading term of the error, for the Cavalieri-Simpson's rule is given by:

$$\mathcal{E}_{[x_{i-1},x_{i+1}]} \simeq -\frac{1}{180} f^{IV}(x_i)h^5$$

- Finally, by adding up all the errors on the couples of subintervals, we get, by remembering that $n = (b\text{-}a)/h$, the global error:

$$\mathcal{E} \propto h^4$$