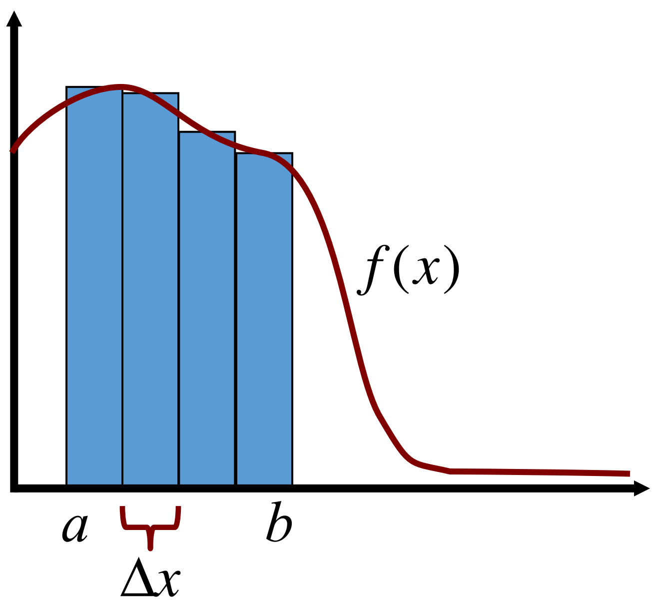
Numerical integration

Recall: some facts on integrals

- Function *f : R --> R*, continuous on an interval *[a, b]*, then its definite integral is

$$\int_a^b f(x)dx = \lim_{n\to\infty}\sum_{k=1}^{n} f(x_k)\Delta x \quad \text{with} \quad \Delta x = \frac{b-a}{n}$$

- Interpretation: it is the area under the curve



- Extends to higher dimensions: a double definite integral -> volume

Why numerical integration?

- Allows us to compute the definite integral without needing to derive the analytical expression!
- Very practical way to compute the definite integrals when the integrands are very complex or for functions whose primitives cannot be expressed in closed form
- When it is needed to compute the integrals (areas, volumes, etc.) for a given set of tabulated data (we have pairs (x_i, y_i) for $i=1,\ldots,n$)

Newton-Cotes integration

- Idea: replace *f(x)* or tabulated data by simple functions that are easy to integrate.
- In the Newton-Cotes scheme *f(x)* is approximated by a n^{th} order polynomial

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

$$p_n(x) = a_0 + a_1 x + ... + a_n x^n$$

- We will discuss at this course:
  - Order 1 (linear approximation), Trapezoid rule, error *$O(h^2)$*
  - Order 2 (quadratic approximation), Simpson's 1/3 rule, error *$O(h^4)$*
  - Order 3 (cubic approximation), Simpson's 3/8 rule, error *$O(h^4)$*

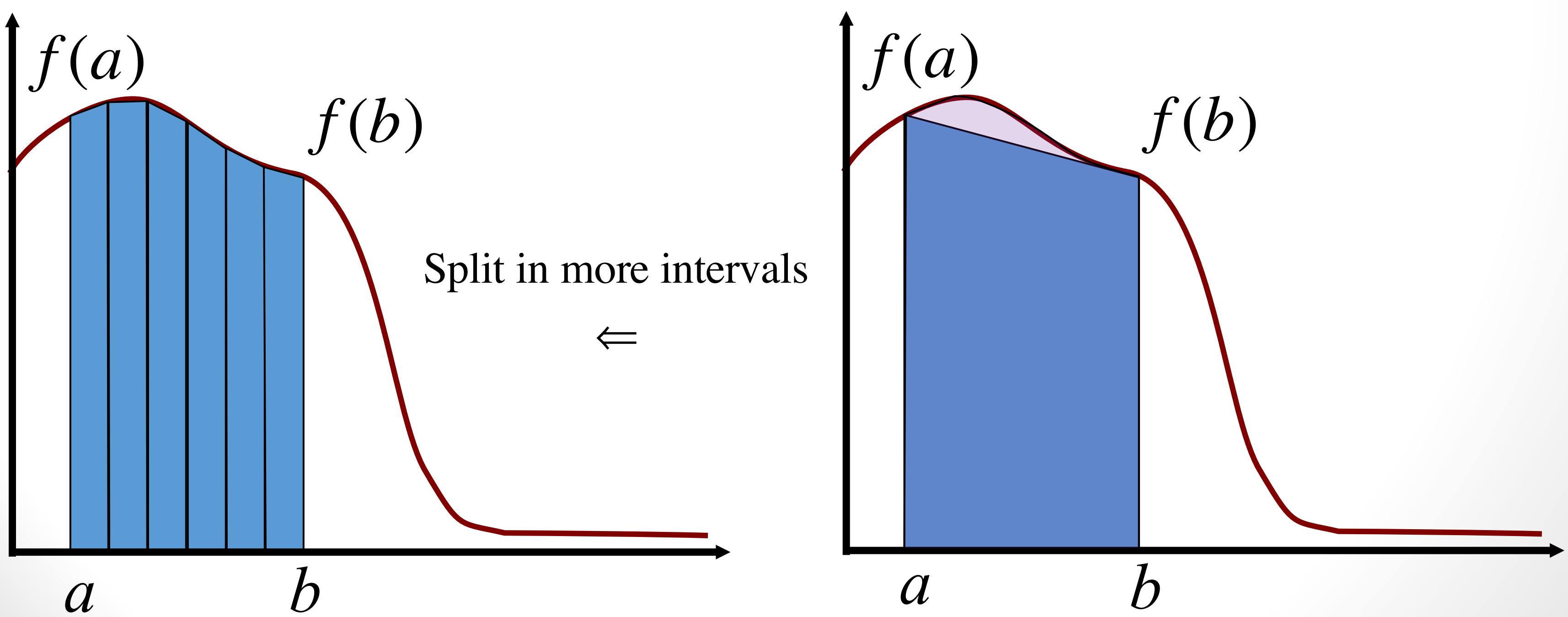
Trapezoid rule

- The area under the curve *f(x)* is approximated by a *1st order polynomial*

$$I = \int_a^b f(x)dx \approx \int_a^b (a_0 + a_1 x)dx$$

$$a_0 + a_1 x = f(a) + \frac{f(b) - f(a)}{b-a}(x-a)$$

$$I \approx \frac{(b-a)}{2}\left[f(a) + f(b)\right]$$



Composite trapezoid rule

- Improve the accuracy by splitting the integration interval in *n* equal subintervals, i.e., $h = \dfrac{(b-a)}{n}$
- From

$$I = \int_a^{a+h} f(x)dx + \int_{a+h}^{a+2h} f(x)dx + \ldots + \int_{b-h}^{b} f(x)dx = \sum_{i=1}^{n} \int_{a+(i-1)h}^{a+ih} f(x)dx$$

- and

$$\int_{a+(i-1)h}^{a+ih} f(x)dx \approx \underbrace{\frac{(a+ih-a-(i-1)h)}{2}}_{h/2}\left[f(a+ih)+f(a+(i-1)h)\right]$$

- we obtain

$$I \approx I_h = \frac{h}{2}\left[f(a+h)+f(a)+f(a+2h)+f(a+h)+\cdots+f(b)\right]$$

$$\boxed{I_h = \frac{h}{2}\left[f(a)+2\sum_{i=1}^{n-1} f(a+ih)+f(b)\right]}$$

Composite trapezoid rule error

- It can be shown that the error of this rule is

$$\int_a^b f(x)dx - I_h = -\frac{h^3}{12}\sum_{i=1}^{n} f''(\eta_i)$$

for some η_i in the intervals *[a+ih, a+(i-1)h]*. This can be expressed as

$$\int_a^b f(x)dx - I_h = -\frac{h^2(b-a)}{12\,n}\sum_{i=1}^{n} f''(\eta_i) = -\frac{h^2(b-a)}{12} f''(\mu)$$

- If *f(x)* is linear, then this error is zero!
- $O(h^2) \rightarrow$ if *h* is halved, then this error is quartered!
- It depends on the width of the integrated area

Romberg integration

- Recall Richardson extrapolation: use two estimates (in this case of an integral) for two *h* to compute a third more accurate approximation
- We have seen that

$$\int_a^b f(x)\,dx - I_h = Ch^2 + O(h^4)$$

- If we compute an approximation for *h/2* we obtain

$$\int_a^b f(x)\,dx - I_{h/2} = C\left(h/2\right)^2 + O(h^4)$$

- Cancelling h^2 terms we have

$$\int_a^b f(x)\,dx = \frac{4}{3} I_{h/2} - \frac{1}{3} I_h + O(h^4)$$

Romberg integration

- So, we have obtained an *$O(h^4)$* estimate

$$\int_a^b f(x)dx \approx \frac{4}{3} I_{h/2} - \frac{1}{3} I_h = \frac{4 I_{h/2} - I_h}{4^1 - 1} = I_{h/2} + \frac{I_{h/2} - I_h}{4^1 - 1}$$

- Notation: $I(1,1)=I_h$, $I(2,1)=I_{h/2}$

$$I_{h/2} + \frac{I_{h/2} - I_h}{4^1 - 1} \Rightarrow \boxed{I(2,1) + \frac{I(2,1) - I(1,1)}{4^1 - 1} = I(2,2)}$$

		k=1	k=2		
j=1	h	$I(1,1)=I_h$			
j=2	$h/2$	$I(2,1)=I_{h/2}$	→ $I(2,2)$		
j=3	$h/4$	$I(3,1)=I_{h/4}$			
j=4	$h/8$	$I(4,1)=I_{h/8}$			
	Error order	$O(h^2)$			

Romberg integration

- So, we have obtained an *O(h⁴)* estimate

$$\int_a^b f(x)dx \approx \frac{4}{3} I_{h/2} - \frac{1}{3} I_h = \frac{4I_{h/2} - I_h}{4^1 - 1} = I_{h/2} + \frac{I_{h/2} - I_h}{4^1 - 1}$$

- Notation: $I(1,1)=I_h$, $I(2,1)=I_{h/2}$

$$I_{h/2} + \frac{I_{h/2} - I_h}{4^1 - 1} \Rightarrow I(2,1) + \frac{I(2,1) - I(1,1)}{4^1 - 1} = I(2,2)$$

		k=1	k=2		
j=1	*h*	$I(1,1)=I_h$			
j=2	*h/2*	$I(2,1)=I_{h/2}$	$I(2,2)$		
j=3	*h/4*	$I(3,1)=I_{h/4}$	$I(3,2)$		
j=4	*h/8*	$I(4,1)=I_{h/8}$	$I(4,2)$		
	Error order	$O(h^2)$	$O(h^4)$		

Romberg integration

- Following a similar logic, we can combine two *$O(h^4)$* estimates to obtain an *$O(h^6)$* estimate:

$$I(j,3) = I(j,2) + \frac{I(j,2) - I(j-1,2)}{4^2 - 1}, \quad j = 3,4,\ldots$$

		k=1	k=2	k=3	
j=1	h	$I(1,1)=I_h$			
j=2	$h/2$	$I(2,1)=I_{h/2}$	$I(2,2)$		
j=3	$h/4$	$I(3,1)=I_{h/4}$	$I(3,2)$	$I(3,3)$	
j=4	$h/8$	$I(4,1)=I_{h/8}$	$I(4,2)$	$I(4,3)$	
	Error order	$O(h^2)$	$O(h^4)$	$O(h^6)$	

Romberg integration

- To obtain an $O(h^8)$ estimate:

$$I(j,4) = I(j,3) + \frac{I(j,3) - I(j-1,3)}{4^3 - 1}, \quad j = 4,5,\ldots$$

		k=1	k=2	k=3	k=4
j=1	h	$I(1,1)=I_h$			
j=2	$h/2$	$I(2,1)=I_{h/2}$	$I(2,2)$		
j=3	$h/4$	$I(3,1)=I_{h/4}$	$I(3,2)$	$I(3,3)$	
j=4	$h/8$	$I(4,1)=I_{h/8}$	$I(4,2)$	$I(4,3)$	$I(4,4)$
	Error order	$O(h^2)$	$O(h^4)$	$O(h^6)$	$O(h^8)$

Romberg integration

- From this (and previous) estimate(s)

$$I(j,4) = I(j,3) + \frac{I(j,3) - I(j-1,3)}{4^3 - 1}, \quad j = 4,5,\ldots$$

- we deduce a general recursive scheme:

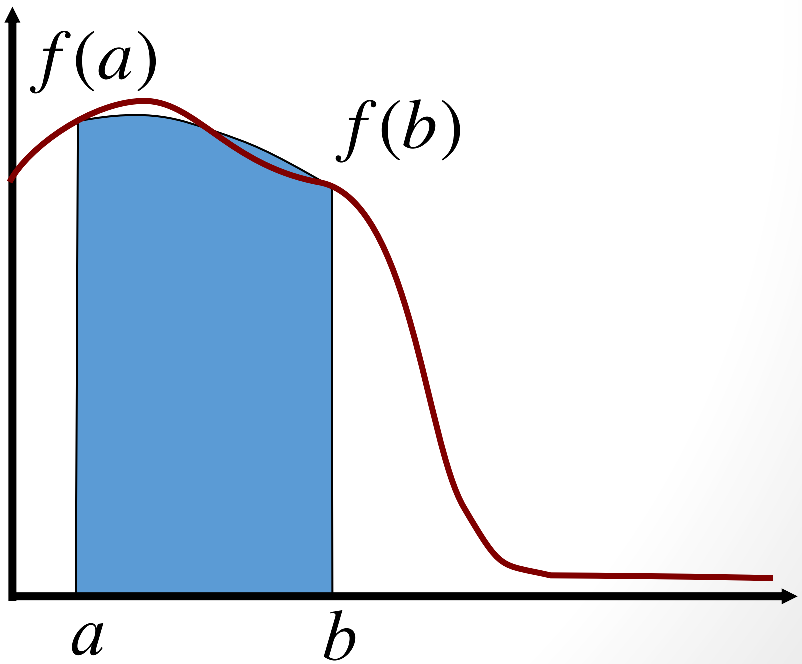
$$I(j,k+1) = I(j,k) + \frac{I(j,k) - I(j-1,k)}{4^k - 1}$$

- Stopping condition
  - Stop if $| I(j,j) - I(j,j-1) | < \varepsilon$
  - Stop after pre-specified *k* steps

Increasing the order of polynomial

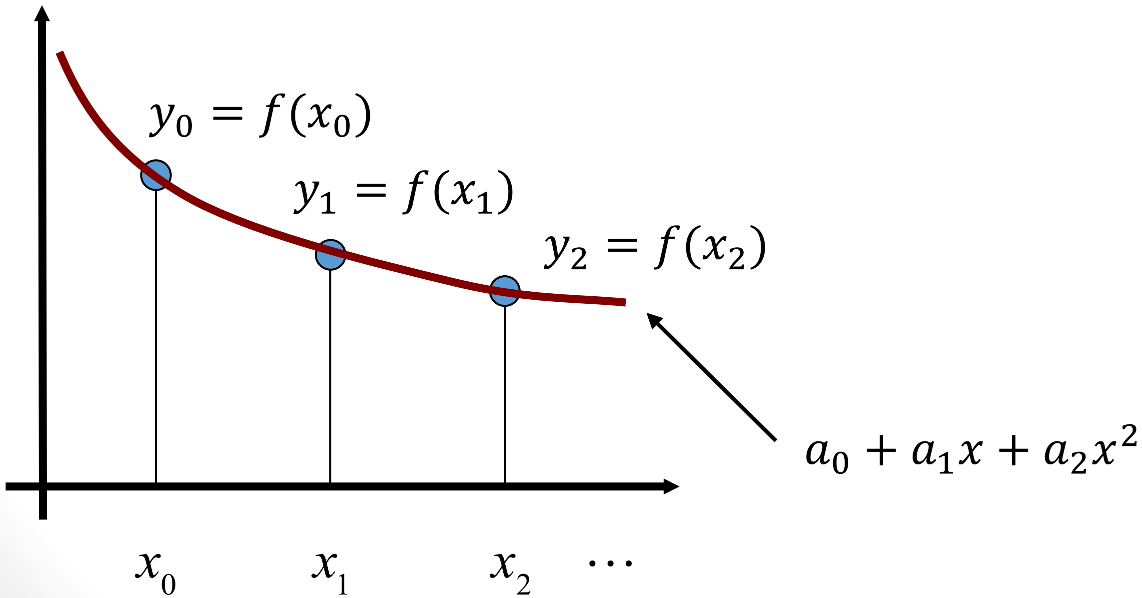
- Romberg integration: to increase the accuracy we have applied recursively the Richardson extrapolation over the 1st order polynomial approximations
- Instead, we might increase the order of polynomial

$$I = \int_a^b f(x)\,dx \approx \int_a^b a_0 + a_1 x + a_2 x^2 \, dx$$



How to determine coefficients a_0, a_1, and a_2?

- For each interval, we have to find the coefficients of the 2nd-order polynomial (a_0, a_1, and a_2) that passes through the points $y_0 = f(x_0)$, $y_1 = f(x_1)$, and $y_2 = f(x_2)$…
- A solution -> Lagrange interpolating polynomials

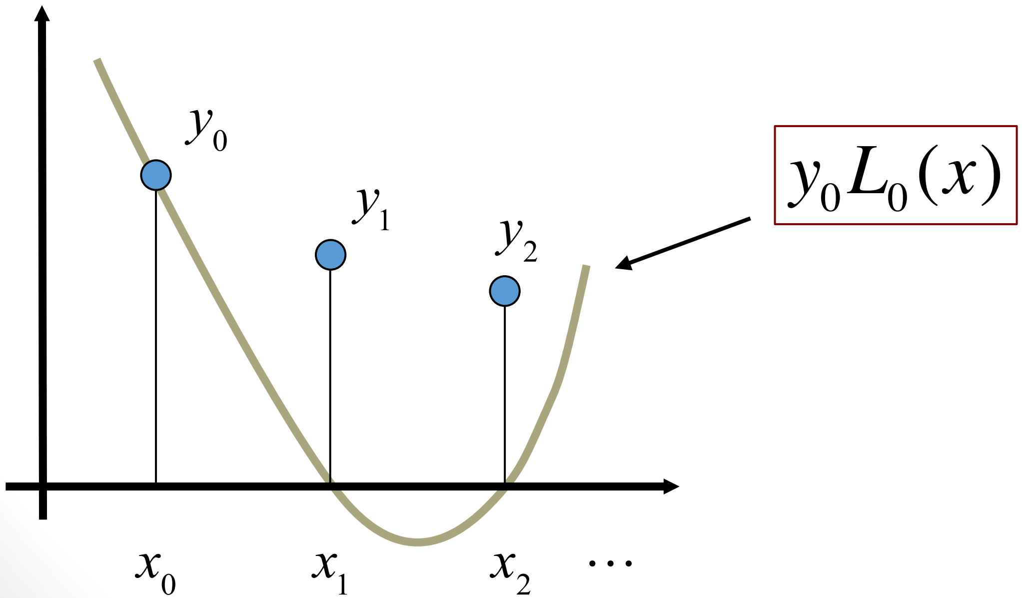


Numerical Methods

Lagrange interpolating polynomial

- Consider the 2nd-order Lagrange polynomial

$$L_0(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} \Rightarrow \begin{array}{l} L_0(x_0) = 1; \\ L_0(x_1) = 0; \quad L_0(x_2) = 0; \end{array}$$

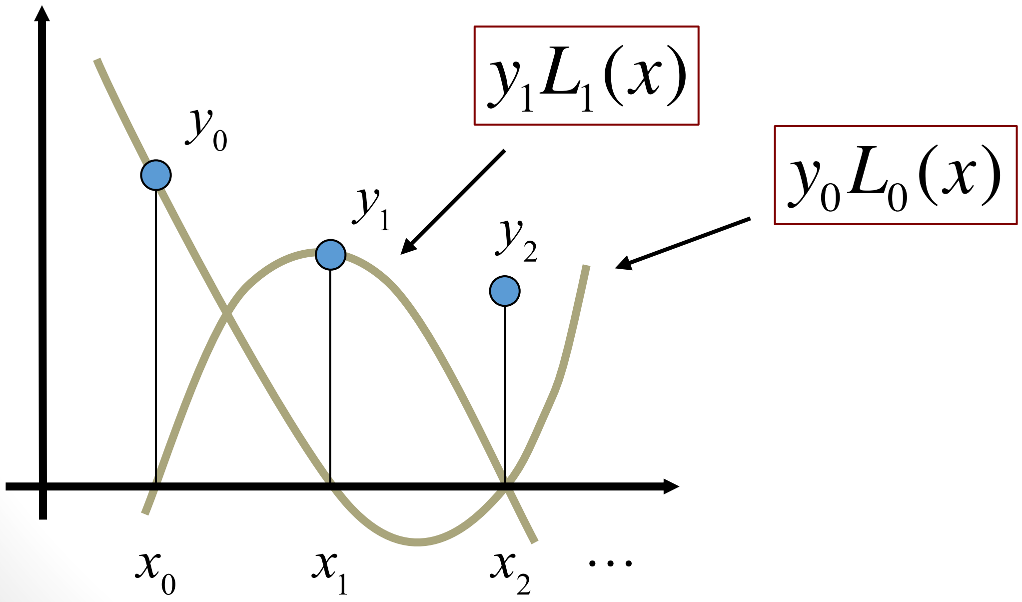


Lagrange interpolating polynomial

- Consider the 2nd-order Lagrange polynomial

$$L_0(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} \Rightarrow \begin{array}{l} L_0(x_0) = 1; \\ L_0(x_1) = 0; \quad L_0(x_2) = 0; \end{array}$$

$$L_1(x) = \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} \Rightarrow \begin{array}{l} L_1(x_0) = 0; \\ L_1(x_1) = 1; \\ L_1(x_2) = 0; \end{array}$$



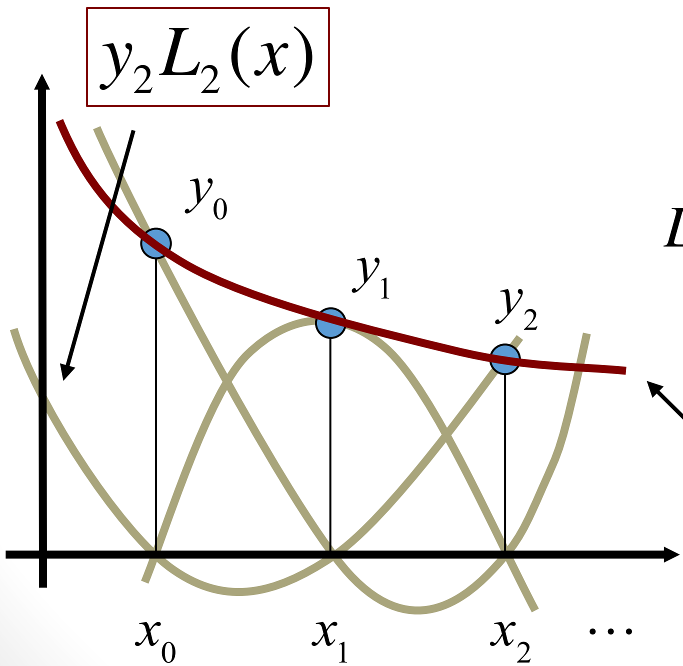
Lagrange interpolating polynomial

- Consider the 2nd-order Lagrange polynomial

$$L_0(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} \Rightarrow \begin{array}{l} L_0(x_0) = 1; \\ L_0(x_1) = 0; \quad L_0(x_2) = 0; \end{array}$$

$$L_1(x) = \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} \Rightarrow \begin{array}{l} L_1(x_0) = 0; \\ L_1(x_1) = 1; \\ L_1(x_2) = 0; \end{array}$$

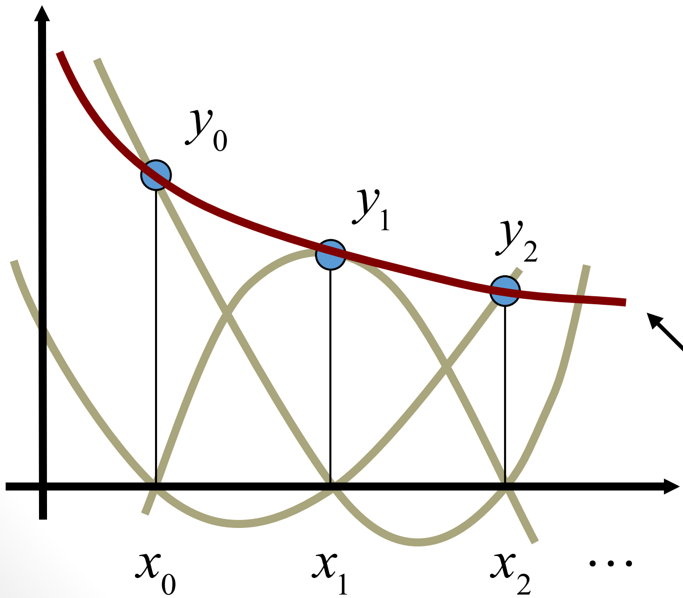
$$L_2(x) = \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} \Rightarrow \begin{array}{l} L_2(x_0) = 0; \\ L_2(x_1) = 0; \\ L_2(x_2) = 1; \end{array}$$



$$f(x) = y_0 L_0(x) + y_1 L_1(x) + y_2 L_2(x)$$

Lagrange interpolating polynomial

2nd-order polynomial passing through $y_0, y_1,$ and y_2
is computed as
a weighted sum of three Lagrange polynomials!



$$f(x) = y_0 L_0(x) + y_1 L_1(x) + y_2 L_2(x)$$

Simpson's 1/3 Rule

- Simpson's 1/3 rule uses a 2nd-order Lagrange polynomial

$$I \approx \int_{x_0}^{x_2} \left[\sum_{j=0}^{2} f(x_j) \prod_{\substack{i=0 \\ i \neq j}}^{2} \frac{(x - x_i)}{x_j - x_i} \right] dx = \sum_{j=0}^{2} f(x_j) \int_{x_0}^{x_2} \left[\prod_{\substack{i=0 \\ i \neq j}}^{2} \frac{(x - x_i)}{x_j - x_i} \right] dx$$

with $x_0 = a$, $x_2 = b$, $x_1 = (a+b)/2$. Since $x_2 - x_1 = x_1 - x_0 = h$, we get

$$I \approx f(x_0) \int_{x_0}^{x_2} \frac{(x - x_0 - h)(x - x_0 - 2h)}{2h^2} dx - f(x_1) \int_{x_0}^{x_2} \frac{(x - x_0)(x - x_0 - 2h)}{h^2} dx$$
$$+ f(x_2) \int_{x_0}^{x_2} \frac{(x - x_0)(x - x_0 - h)}{2h^2} dx$$

substitute $\mathrm{t} = x - x_0$ and integrate to finally obtain

$$\boxed{I \approx \frac{h}{3}\left[f(x_0) + 4f(x_1) + f(x_2) \right]}$$

Composite Simpson's 1/3 Rule

- Improve the accuracy by splitting the integration interval in *n* equal subintervals, i.e., $h = \frac{(b-a)}{n}$
- Since Simpson's 1/3 rule needs 3 points (2 intervals) -> divide the integration interval in an even number of parts:

$$I \approx \int_{x_0}^{x_2} f(x)dx + \int_{x_2}^{x_4} f(x)dx + \ldots + \int_{x_{n-2}}^{x_n} f(x)dx$$

$$I \approx (b-a)\frac{f(x_0) + 4\sum_{i=1,3,5}^{n-1} f(x_i) + 2\sum_{j=2,4,6}^{n-2} f(x_j) + f(x_n)}{3n}$$

Composite Simpson's 1/3 rule error

- It can be shown that the error of this rule is

$$\int_a^b f(x)dx - I = -\frac{h^4(b-a)}{180}\frac{\sum_{i=1}^{n} f^{(4)}(\eta_i)}{n} = O(h^4)$$

for some η_i in the intervals *[a+ih, a+(i-1)h]*

- If *f(x)* is cubic, then this error is zero!
- If *h* is halved, then the error is reduced by factor of 16!
- It depends on the width of the integrated area

Simpson's 3/8 rule

- 3rd-order Lagrange polynomial is used -> Simpson's 3/8 rule

$$I = \frac{3}{8}h\left[f\left(x_0\right) + 3\left(x_1\right) + 3f\left(x_2\right) + f\left(x_3\right)\right]$$

- Distance between points *h=(b-a)/3* (equidistant)

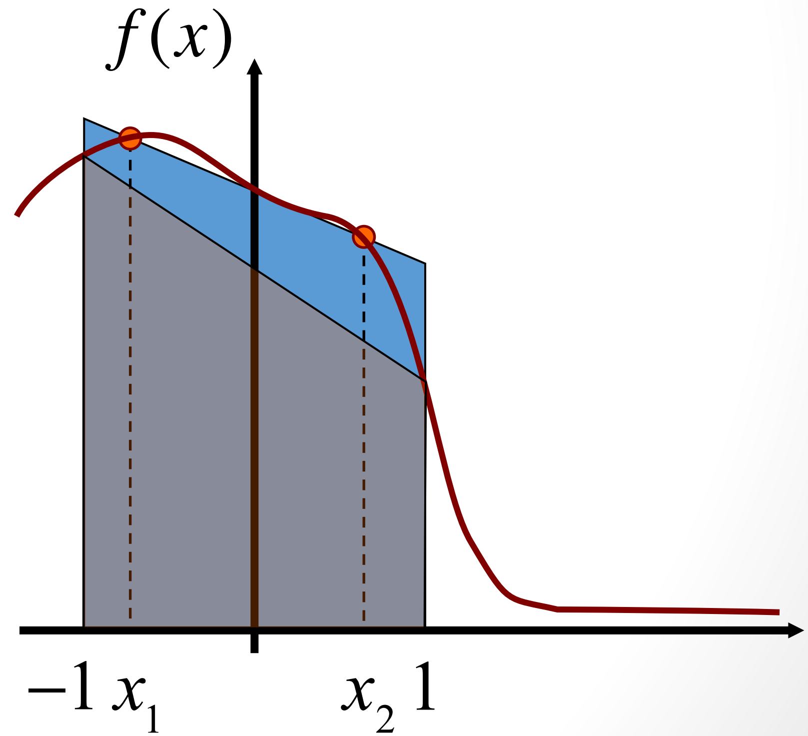
- Error is of the same order as 1/3 -> $O(h^4)$

Gauss quadrature

- Trapezoid rule: $$I \approx \frac{(b-a)}{2}\left[f(a)+f(b)\right]$$
- Idea: what if we "optimize" the points where is the function is evaluated, and the corresponding weights?
- So:

$$I = \int_{-1}^{1} f(x)dx \approx \sum_{i=1}^{n} \omega_i \cdot f(x_i)$$

  - ω_i – weights
  - x_i – points of evaluation



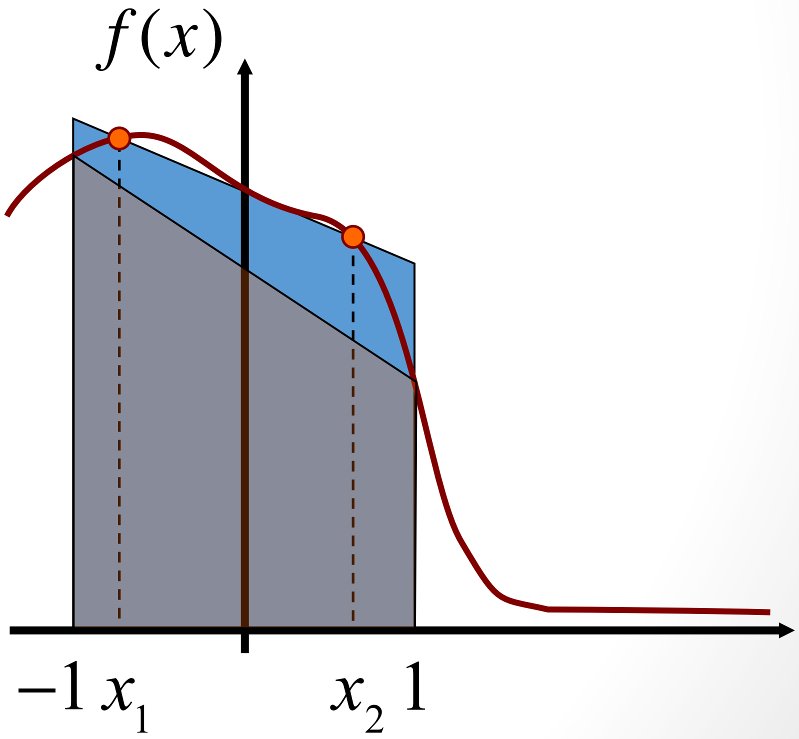
Two points Gauss-Legendre formula

- Consider only two points:

$$I \approx \sum_{i=1}^{2} \omega_i \cdot f(x_i) = \omega_1 \cdot f(x_1) + \omega_2 \cdot f(x_2)$$

- we have to find ω_1, ω_2, x_1 *and* x_2

- Idea: assume that formula fits well the integrals of :
  - constant
  - linear function
  - quadratic function
  - cubic function



Two points Gauss-Legendre formula

- This way we have a nonlinear system of 4 equations and 4 unknowns

$$\int_{-1}^{1} 1\,dx \quad = 2 \quad = \omega_1 \cdot f(x_1) + \omega_2 \cdot f(x_2) \qquad = \omega_1 + \omega_2$$

$$\int_{-1}^{1} x\,dx \quad = 0 \quad = \omega_1 \cdot f(x_1) + \omega_2 \cdot f(x_2) \quad = \omega_1 x_1 + \omega_2 x_2$$

$$\int_{-1}^{1} x^2\,dx \quad = \frac{2}{3} \quad = \omega_1 \cdot f(x_1) + \omega_2 \cdot f(x_2) \quad = \omega_1 x_1^{\,2} + \omega_2 x_2^{\,2}$$

$$\int_{-1}^{1} x^3\,dx \quad = 0 \quad = \omega_1 \cdot f(x_1) + \omega_2 \cdot f(x_2) \quad = \omega_1 x_1^{\,3} + \omega_2 x_2^{\,3}$$

- Solution:

$$\omega_1 = \omega_2 = 1; \quad x_1 = -\frac{1}{\sqrt{3}} = -x_2$$

Two points Gauss-Legendre formula

$$I \approx f\left(-\frac{1}{\sqrt{3}}\right) + f\left(\frac{1}{\sqrt{3}}\right)$$

- This formula gives exact solution for all polynomials up to the ones with a cubic dependency
- Practical issue: it is derived for the interval *(-1,1)*. In a general case, make a change of variables as follows:

$$y = \frac{(b-a)}{2}x + \frac{b+a}{2}$$

- So:

$$\int_a^b f(y)dy = \frac{(b-a)}{2}\int_{-1}^{1} f\left(\frac{(b-a)}{2}x + \frac{b+a}{2}\right)dx$$