The Republic of Iraq
Ministry of Higher Education and Scientific Research
Northern Technical University
Engineering Technical College / Mosul
Department of Power Mechanics Techniques Engineering

NUMERICAL ANALYSIS

For Third-Year Students

Dr. Qais A. Yousif

<u>Non-linear Equation</u>

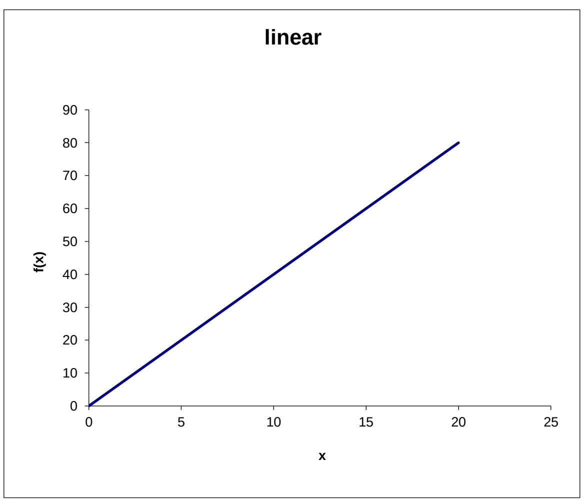
Recall that **a linear equation** can **take the form**

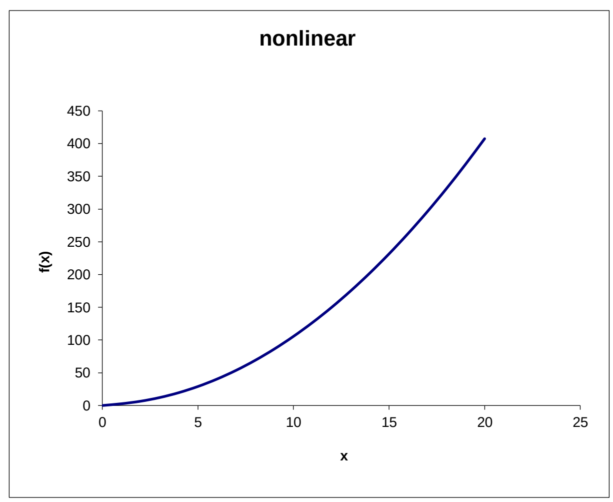
$$ax + by + c = 0$$

any **equation** that **cannot be written** in **this form** and has the **degree larger than one** but **not less than 2** is **nonlinear**. Such as

- **<u>logarithmic</u>**,
- **<u>trigonometric</u>**,
- **<u>exponential</u>** or
- **<u>power function</u>** of **parameters** or **variables**,

It **forms** a **curve** and if the **value of the degree increases**, the **curvature** of the **graph increases**





The **general representation** of **nonlinear equations** is;

$$a\,x^2 + b\,y^2 + c = 0$$

Where ***x*** and ***y*** are **the variables** and ***a, b*** and ***c*** are the **constant values.**

Solution of Nonlinear Equations

Given a **function** $\boldsymbol{f(x)}$, we seek the **value** of $\boldsymbol{x}$ for which $\boldsymbol{f(x) = 0}$. **Solution** $\boldsymbol{x}$ is called a **root of equation**, or **zero of function** $\boldsymbol{f(x)}$. Thus, the **problem** is **known** as **root** or **zero finding**.

Root of Equations

The general form of **quadratic equation** is:

$$\boldsymbol{ax^2 + bx + c = 0}$$

the **value** of $\boldsymbol{x}$, can be find by using the **quadratic formula**

$$\boldsymbol{x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}}$$

These **values** of $\boldsymbol{x}$ are **called** the "**roots**" of **equation**. They **represent** the **values** of $\boldsymbol{x}$ that **make equation equal** to **zero**. Thus, we **can define** the **root** of an **equation** as the **value** of $\boldsymbol{x}$ that **makes** $\boldsymbol{f(x) = 0}$. For this reason, **roots** are **sometimes called** the **zeros** of the **equation**.

Although the **quadratic formula** is **handy** for **solving equation**, there are **many other functions** for which the **root cannot** be **determined** so **easily**. For these cases, the **numerical methods provide efficient means** to **obtain** the **answer**.

Finding for Roots of Nonlinear Equations

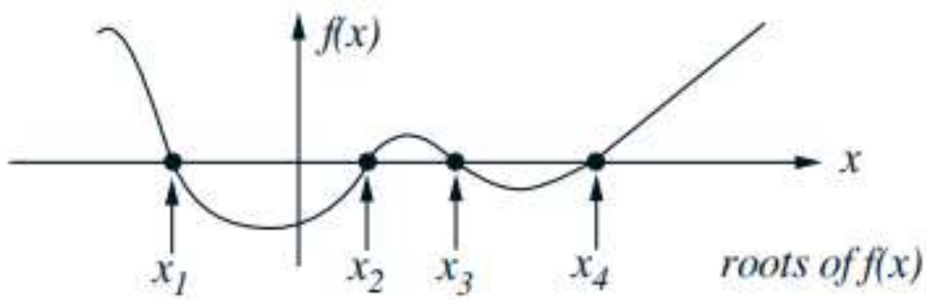
Consider the **problem** of **finding** the **roots** of the **nonlinear equation** $\boldsymbol{f(x) = 0}$, which is a **scalar function** of a **single independent variable**, $\boldsymbol{x}$, using **numerical techniques**. Since the **roots** of a **nonlinear equation**, $\boldsymbol{f(x) = 0}$, **cannot** in general be **expressed** in **closed form**, an **approximate methods** must **use** to **solve** the **problem**.

Usually, **iterative techniques** are **used** to **start from** an **initial approximation** to a **root**, and **produce** a **sequence** of **approximations**, which **converge toward** a **root**. **Convergence** to a **root** is **usually possible provided** that the **function** $\boldsymbol{f(x) = 0}$ is **sufficiently smooth** and the **initial approximation** is **close enough** to the **root**.

Consider the **equation**

$$\boldsymbol{f(x) = 0}$$

Roots of **equation** $\boldsymbol{f(x) = 0}$ are the **values** of $\boldsymbol{x}$ which **satisfy** $\boldsymbol{f(x) = 0}$**,** also **referred** to as the **zeros** of **an equation**



Notes on root finding

- **Roots** of **equations** can be **either real** or **complex** such that $\boldsymbol{x = a}$ is a **real number** and $\boldsymbol{x = a + ib}$ is a **complex number**, where $\boldsymbol{i = \sqrt{-1}}$
- A **large variety** of **root finding algorithms** exist,
- **Each algorithm** has **advantages** and **disadvantages**, **possible restrictions**, etc.

Numerical Methods of Roots Finding of Nonlinear Equations

Numerical methods for **finding Roots** of the **Nonlinear Equations** are:

- **Bracketting Methods**
  - **Bisection Method**
  - **False Position (Regula Falsi**
- **Graphical Method**
- Open Methods
  - **Newton-Raphson**
  - **Secant Method**

Noncomputer Methods for Determining Roots

Before the **advent** of **digital computers**, there were **several ways** to **find** the **roots** of **algebraic** and **transcendental equations**. For some cases, the **quadratic formula**. Although there were equations that could be solved directly, there were many more that could not. **For example**, even an apparently **simple function** such as

$$\boldsymbol{f(x) = e^{-x} - x}$$

cannot be solved analytically. In such instances, the **only alternative** is an **approximate solution technique**.

One method to **obtain** an **approximate solution** is to **plot** the **function** and **determine** where it **crosses** the ***x*** **axis**. This **point**, which **represents** the x **value** for **which** $\boldsymbol{f(x) = 0}$, is the **root**.

Although **graphical methods** are useful for **obtaining rough estimates** of **roots**, they are **limited** because of their **lack** of **precision**. An **alternative approach** is to **use trial** and **error**. This "**technique**" **consists of guessing** a **value** of x and **evaluating whether** $\boldsymbol{f(x) = 0}$. If **not** (as is **almost** always the **case**), **another guess** is **made**, and $\boldsymbol{f(x)}$ is **again**

evaluated to **determine whether** the **new value provides** a **better estimate** of the **root**. The **process** is **repeated until a guess** is **obtained** that **results** in an $\boldsymbol{f(x)}$ that is **close to zero**.

Such **haphazard methods** are **obviously inefficient** and **inadequate** for the **requirements** of **engineering practice**. There are **techniques represent alternatives** that are also **approximate** but **employ systematic strategies** to **home** in on the **true root**. The **combination** of **these systematic methods** and **computers makes** the **solution** of **most applied roots-of-equations problems** a **simple** and **efficient task**.

Bracketing Methods

The **techniques** for **finding** the **roots** of **equations** deals that **exploit** the **fact** that a **function** typically **changes sign** in the **vicinity** of a **root** are called **bracketing methods** because **two initial guesses** for the **root** are **required**. **As** the **name implies**, these **guesses must** "**bracket**," or **be on either side** of, the **root**. These **methods employ different strategies** to **systematically reduce** the **width** of the **bracket** to **find** the **correct answer**.

As a **prelude** to these t**echniques**, **graphical methods** for **depicting functions** and **their roots**. **Beyond their utility** for **providing rough guesses**, **graphical techniques** are also **useful** for **visualizing the properties** of the **functions** and the **behavior** of the **various numerical methods**.

1. Graphical Methods

A **simple method** for **obtaining** an **estimate** of the **root** of the **equation** $\boldsymbol{f(x) = 0}$ is to **make a plot** of the **function** and **observe where** it **crosses** the $\boldsymbol{x}$ **axis**. This **point**, which **represents** the $\boldsymbol{x}$ **value** for which $\boldsymbol{f(x) = 0}$, provides a **rough approximation** of the **root**.

<u>The Graphical Approach</u>

Graphical techniques are of **limited practical value** because they are not precise. However, **graphical methods** can be **utilized** to obtain **rough estimates of roots**. These **estimates** can be **employed** as **starting guesses** for **numerical methods**.

Aside from **providing rough estimates** of the **root**, **graphical interpretations** are **important tools** for **understanding** the **properties** of the **functions** and **anticipating** the **pitfalls** of the **numerical methods**.
For example, **Figures below** show a n**umber** of **ways** in which **roots** can **occur** in an **interval prescribed** by a **lower bound** x_l and an **upper bound** x_u.

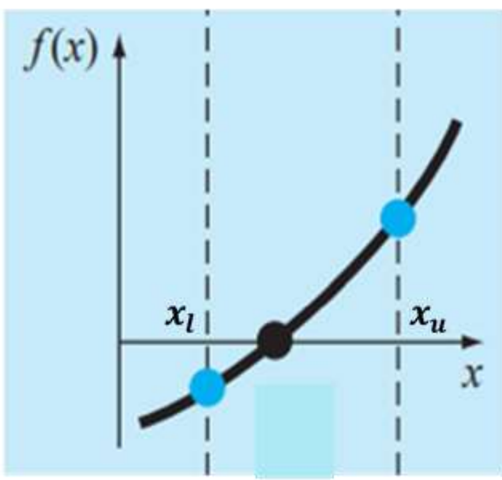


Figure 1.

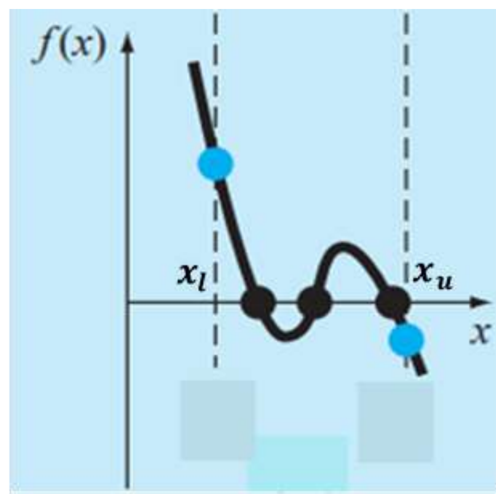


Figure 2.

<u>Figure</u> 1 and **2 depicts** the **case** where a **single root** is **bracketed** by **negative** and **positive values** of **function** $f(x)$. In general,

- **If** $f(x_l)$ and $f(x_u)$ **have** **<u>opposite signs</u>**, then there are **<u>an odd number</u>** of **<u>roots</u>** in the **interval** as **shown** by **<u>Figures</u> 1** and **2**. In **Figures** $f(x_l)$ and $f(x_u)$ are **<u>on opposite sides</u>** of the x **axis**, and **<u>three roots (odd number)</u>** occur **between** the **interval**.

- if $f(x_l)$and $f(x_u)$ **have** the **<u>same sign</u>**, as **shown** in **<u>Figures</u> 3** and **4**, then there are either **<u>no roots</u>** or **<u>an even number</u> (two)** of **roots between** the **values.**

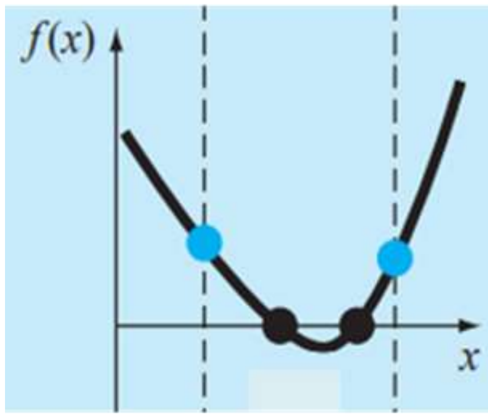


Figure 3.

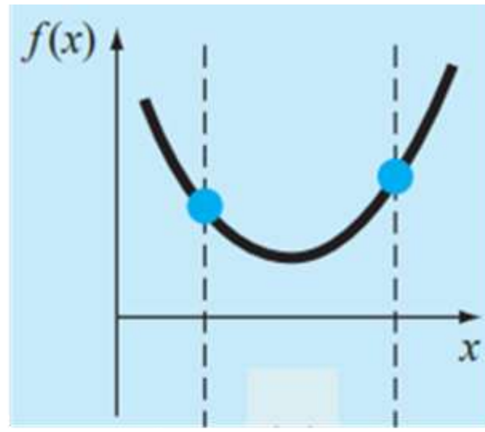


Figure 4.

Although these **generalizations** are **usually true**, there are **cases** where they **do not hold**. For example, **Tangential** and **discontinuous functions** can **<u>violate</u>** these **principles**.

Tangential function in **Figure 5** which is **tangential** to the ***x* axis** has **<u>an even number</u>** of **root** of intersection ***x*-axis** although the **end points** are of **opposite signs**. Also, **discontinuous function** in **Figure 6** has **<u>an even number</u>** of **roots although** the **end points** of **opposite sign**.

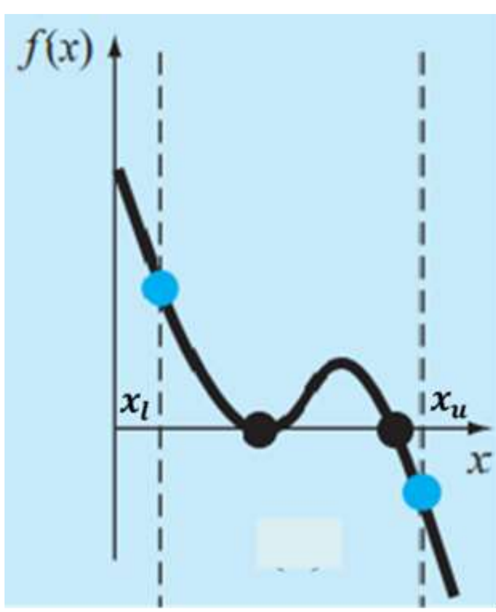


Figure 5.

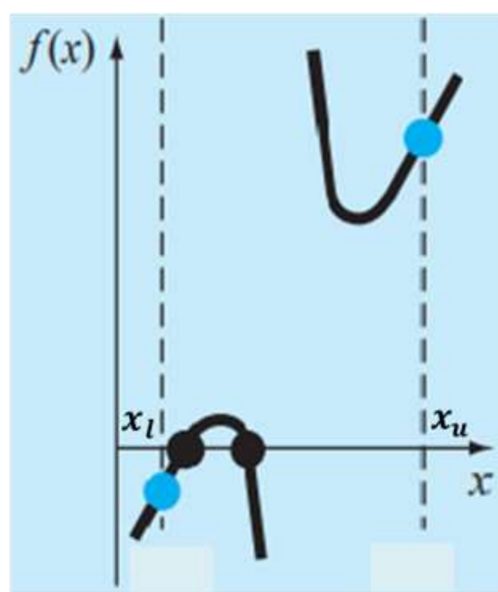


Figure 6.

An **example** of a **function** that is **tangential** to the **axis** is the **cubic equation**:

$$f(x) = (x - 2)(x - 2)(x - 4)$$

Notice that $x = 2$ **makes two terms** in this **polynomial equal** to **zero**. Mathematically, $x = 2$ is **called** a **multiple root**.

The **existence** of **cases illustrated** in **Figures 5** and **6** makes it **difficult** to **develop general computer algorithms guaranteed** to **locate** all the **roots** in **an interval**. However, when **used** in **conjunction** with **graphical approaches**, the **methods** for **solving many roots** of **equations problems** are **extremely useful confronted routinely** by **engineers** and **applied mathematicians**

Example : Use the **graphical approach** to **determine** the **friction factor needed** for a **parachutist** of **mass** $m = 68.1$ **kg** to **have** a **velocity** of **40 m/s** after **freefalling** for **time t =10 s**. Where, the **parachutist's velocity derived** from **Newton's second law** is

$$v = \frac{mg}{c}\left(1 - e^{-\left(\frac{c}{m}\right)t}\right)$$

Solution:

To **solve** this **problem using numerical methods**, it is **conventional** to **re-express velocity equation** by **subtracting** the **dependent variable** v from **both sides** of the **equation** to have

$$f(c) = \frac{mg}{c}\left(1 - e^{-\left(\frac{c}{m}\right)t}\right) - v$$

Therefore, the **value** of **friction factor** c that **makes the function** $f(c) = 0$ is the **root** of the **equation**. This **value** also **represents** the **drag solves** the design **problem**.

Applying **known parameters**:

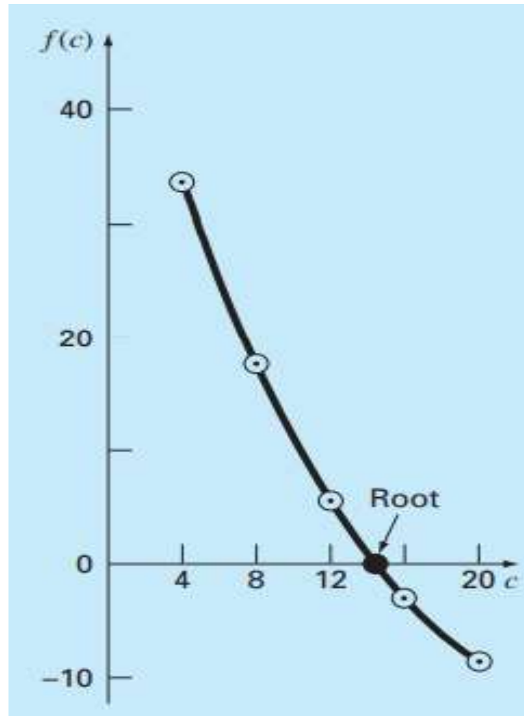
$m = 68.1$ kg, $v = 40\ m/s$, t =10 s, and $g = 9.81\ m/s^2$, yields

$$f(c) = \frac{68.1 \times 9.81}{c}\left(1 - e^{-\left(\frac{c}{68.1}\right)\times 10}\right) - 40$$

$$\Rightarrow\ f(c) = \frac{668.06}{c}\left(1 - e^{-0.15\,c}\right) - 40$$

Various values of c can be **substituted** into the **right-hand side** of this **equation** to compute $f(c)$ and **plot** the **obtained points** as **shown** in **figure below**.

c	$f(c)$
4	34.190
8	17.712
12	6.114
16	-2.23
20	-8.368



The **resulting curve crosses** the c**-axis between 12** and **16**. **Visual inspection** of the **plo**t **provides** a **rough estimate** of the **root** of **14.75**.

The **validity** of the **graphical estimate** can be **checked** by **substituting** $c = 14.75$ into **equation** of $f(c)$ to give

$$f(14.75) = \frac{668.06}{14.75}\left(1 - e^{-0.146843\times 14.75}\right) - 40 = 0.059$$

which is **close** to **zero**.

It **can also** be **checked** by **substituting** ($c = 14.75$) **into equation** of the **parachutist's velocity** using the **known parameter values** to have:

$$v = \frac{68.1 \times 9.81}{14.75}\left(1 - e^{-\left(\frac{14.75}{68.1}\right)\times 10}\right) = 40.059\ m/s$$

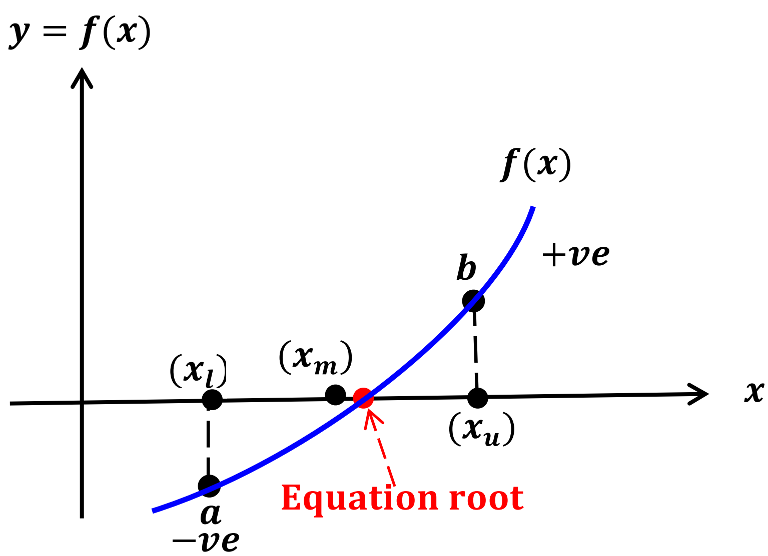
which is **very close** to the **desired fall velocity** of **40 m/s**.

➔ Bisection Method for Finding Roots of Nonlinear Equations

The **bisection method** is a **straightforward technique** to **find numerical solutions** of **an equation** with **one unknown**. **Among all** the **numerical methods**, the **bisection method** is the **simplest one** to **solve** the **transcendental equation**.

The **bisection method**, is **an approximation method** which is also called **binary chopping**, **binary search method**, **interval halving**, or **Bolzano's method**. It is **one type** of **incremental search method** in which the **interval** is always **divided** in **half**. If a **function changes sign** over an **interval**, the **function value** at the **midpoint** is **evaluated**. The **location** of the **root** is then **determined** as **lying** at the **midpoint** of the **subinterval within** which the **sign change occurs**. The **process** is **repeated** to **obtain refined estimates**.

Consider the **function** $f(x)$ **shown** in **figure**, which is **defined** on the **closed interval** $[a, b]$ **given** with $f(a)$ and $f(b)$ of **different signs**.



By the **intermediate theorem**, there is **a point** (x_m) **belong** to the **interval** $[a, b]$ for which $f(x) = 0$.

From the **figures,** it **has observed** that $f(x)$ **changed sign** on **opposite sides** of the **root**. In general, if $f(x)$ is **real** and **continuous** in the **interval from** a as **lower bound** (x_l) to b as **upper bound** (x_u)**,** and $f(x_l)$and $f(x_u)$ have **opposite signs**, that is,

$$f(x_l)\, f(x_u) < 0$$

then **there** is **at least one real root between** x_l and x_u.

Bisection Method Algorithm

A **simple algorithm** for the **bisection calculation** is:

1. **Choose two points**, (x_l) as a **lower bound** and (x_u) as **a upper bound** such that $x_l < x_u$ as **guesses** for the **root** such that the **function changes sign over** the **interval**. This **can be checked** by **ensuring** that

$$f(x_l) \cdot f(x_u) < 0$$

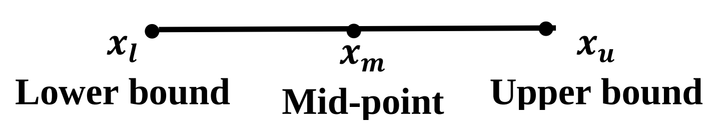
   then **bracket** the **root because of** it is **calling bracket method**

2. **Find** the **midpoint** x_m between **these two points** such that

$$x_m = \frac{x_l + x_u}{2}$$

   as **an estimate** of the **root.**

3. **Evaluate** $f(x_l)$, $f(x_m)$, and $f(x_u)$ to **determine** in **which subinterval** the **root lies**, and **consider** the **product:**



i. **If** $f(x_l)f(x_m) < 0$**; Then** $x_l = x_l$ and $x_u = x_m$, and the **root must lie** in the **interval** $[x_l, x_m]$.

ii. **If the product** $f(x_l)f(x_m) > 0$: **Then** $x_l = x_m$ and $x_u = x_u$, and the **root must lie** in the **interval** $[x_m, x_u]$.

iii. **If** $f(x_l)f(x_m) = 0$; **Then** x_m is the **root** of the **equation** and **stop** the **process**

4. **Repeat** the **process** and **find** the **new estimate** of x_m **by**

$$x_m = \frac{x_l + x_u}{2}$$

5. **Find** the **absolute relative approximate error by**

$$|\epsilon_a| = \left|\frac{x_m^{new} - x_m^{old}}{x_m^{new}}\right| \times 100$$

Where, x_m^{new} – **estimated root** from present iteration
x_m^{old} – **estimated root** from **previous iteration**

6. **Check if** $|\epsilon_a| < \epsilon_s$: **compare** the **absolute relative approximate error** $[\epsilon_a]$ with the **pre-specified relative error tolerance** ϵ_s.

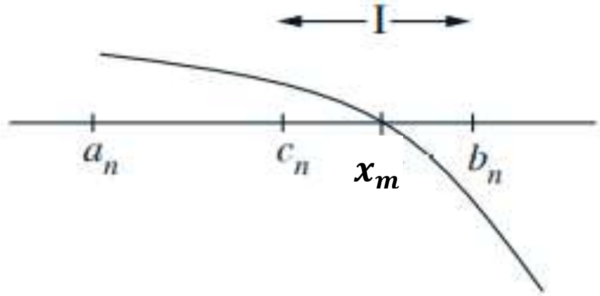
   If $[\epsilon_a] > \epsilon_s$, then **go back** to **Step 2 to repeat the process**, **else stop** the **algorithm**.

7. **Go back** to **step 3** and **repeat** the **process until** a **certain level** of **convergence** has been **achieved**.

Note one **should also check** whether the **number of iterations** is **more than** the **maximum number** of **iterations allowed**. If so, **one needs** to **stop** the **algorithm**.

- **Interval size**, I, after n steps

$$I = \frac{b_n - a_n}{2} = \frac{b_1 - a_1}{2^n}$$



- The **interval size** at any **given iteration** also **corresponds to** the **maximum error** in x_m , therefore

$$E_n \leq \frac{b_1 - a_1}{2^n}$$

- If **wishing** to **limit** the **error** to ε

$$\varepsilon \leq \frac{b_1 - a_1}{2^n} \quad \Rightarrow \quad 2^n \leq \frac{b_1 - a_1}{\varepsilon} \quad \Rightarrow \quad n \geq \frac{ln\left(\frac{b_1 - a_1}{\varepsilon}\right)}{ln2}$$

$$\Rightarrow \quad n \geq 1.443\ ln\left(\frac{b_1 - a_1}{\varepsilon}\right)$$

Iterations must be **applied** to **ensure** the **level of convergence**

<u>Example</u>: Use the **bisection method** to **determine** the **root** of the **equation** $x^2 - 3 = 0$ for $x \in [1, 2]$

<u>Solution</u>:

Let $\quad f(x) = x^2 - 3 = 0$

First of all defining the **lower bound** x_l and **upper bound** x_u **guesses** for the **root,** so $x_l = 1$ and $x_u = 2$.

Then **check whether** the **function** $f(x)$ is **changing sign**, by **evaluating** its **value** at $x_l = 1$ and $x_u = 2$

$$f(x_l = 1) = 1^2 - 3 = -2 < 0$$
$$f(x_2 = 2) = 2^2 - 3 = \quad 1 > 0$$

The **given function** is **continuous** because it **has value** in the **interval [1, 2]**, and is changing sign over the interval, so the **root lies** in this **interval** and **can be checked** by **ensuring** that

$$f(x_l)\, f(x_u) < 0$$

$$f(x_l)\, f(x_u) = -2 \times 1 = -2 < 0$$

Since the **function** is **continuous**, and is **changing** the **sign between 1** and **2,** then **there should** be **at least one root between 1** and **2.** So, we have to **established** that the **lower** and **upper limit** which are **given** for the **root** of the **equation** are **valid**. That is the **first important thing** which we **have to do whenever** we are **going** to **apply bisection method**. So **let's go ahead** and **do iteration one:**

Iteration 1:

Let "x_m" be the **midpoint** of the **interval** i.e.

$$x_m = \frac{1 + 2}{2} = 1.5$$

To **determine** in **which subinterval** the **root lies**, **make** the **following evaluation:**

- **If** $f(x_l)\, f(x_m) < 0$ the **root** **lies** in the **lower subinterval**. Therefore, set $x_l = x_l$ and $x_u = x_m$ and **return** to **step 2**.
- **If** $f(x_l)\, f(x_m) > 0$ the **root** **lies** in the **upper subinterval**. Therefore, set $x_l = x_m$ and $x_u = x_u$ and **return** to **step 2**.
- **If** $f(x_l)\, f(x_m) = 0$ the **root is** x_m **stop** the **computation**

Therefore, the **value** of the **function** at “x_m” is:

$$f(x_m) = (1.5)^2 - 3 = 2.25 - 3 = -0.75 < 0$$

$$f(x_l)\, f(x_m) = -2 \times -0.75 = 1.5 > 0$$

The **function** is **not changing** the **sign** **between** x_l which is **1** and x_m which is **1.5,** so the **root cannot** be **between 1** and **1.5,** thus it **has to** be **between 1.5** and **2** because that is the **upper limit** is **2,** the **lower limit** is **1 and** x_m **is 1.5.** The **new interval** is $[1.5, 2]$and that **reduce** the **interval** which is **happening** in the **bisection method**. So this is the **end** of the **first iteration**. So $x_l = 1$ is **replaced** with $x_m = 1.5$ for the **next iterations**, and the **new interval** will be **[1.5, 2]**.

Iteration 2:

$$x_m = \frac{1.5 + 2}{2} = 1.75$$

$$f(x_l = 1.5) = (1.5)^2 - 3 = 2.25 - 3 = -0.75 < 0$$

$$f(x_m = 1.75) = (1.75)^2 - 3 = 0.0625 - 3 = 0.0625 > 0$$

$$f(x_l)\, f(x_m) = -0.75 \times 0.0625 = -0.0029 < 0$$

The **function** is **changing** the **sign between** x_l which is **1.5** and x_m which is **1.75,** so the **root has to** be **between 1.5** and **1.75.** So, the **new interval is** $[1.5, 1.75]$

The **relative approximate error** between the **old value** and **new value** is:

$$|\epsilon_a| = \left|\frac{x_m^{new} - x_m^{old}}{x_m^{new}}\right| \times 100$$

$$|\epsilon_a| = \left|\frac{1.75 - 1.5}{1.75}\right| \times 100 \cong 14\%$$

Compare the **absolute relative approximate error** $|\epsilon_a|$ with the **pre-specified relative error tolerance** ϵ_s which is:

$$x = \sqrt{3} = 1.7320508 \ldots \text{ and } -1.7320508 \ldots$$

If $\lceil \epsilon_a \rceil > \epsilon_s$, then go to **Step 3**, **else stop** the **algorithm**
The **iterations** for the **given functions** are **listed** in the **following table**:

Iterations	x_l	x_u	x_m	$f(x_l)$	$f(x_u)$	$f(x_m)$
1	**1**	**2**	**1.5**	**-2**	**1**	**-0.75**
2	**1.5**	**2**	**1.75**	**-0.75**	**1**	**0.0625**
3	**1.5**	**1.75**	**1.625**	**-0.75**	**0.0625**	**-0.359**
4	**1.625**	**1.75**	**1.6875**	**-0.3594**	**0.0625**	**-0.1523**
5	**1.6875**	**1.75**	**1.7188**	**-01523**	**0.0625**	**-0.0457**
6	**1.7188**	**1.75**	**1.7344**	**-0.0457**	**0.0625**	**0.0081**
7	**1.7188**	**1.7344**	**1.7266**	**-0.0457**	**0.0081**	**-0.0189**

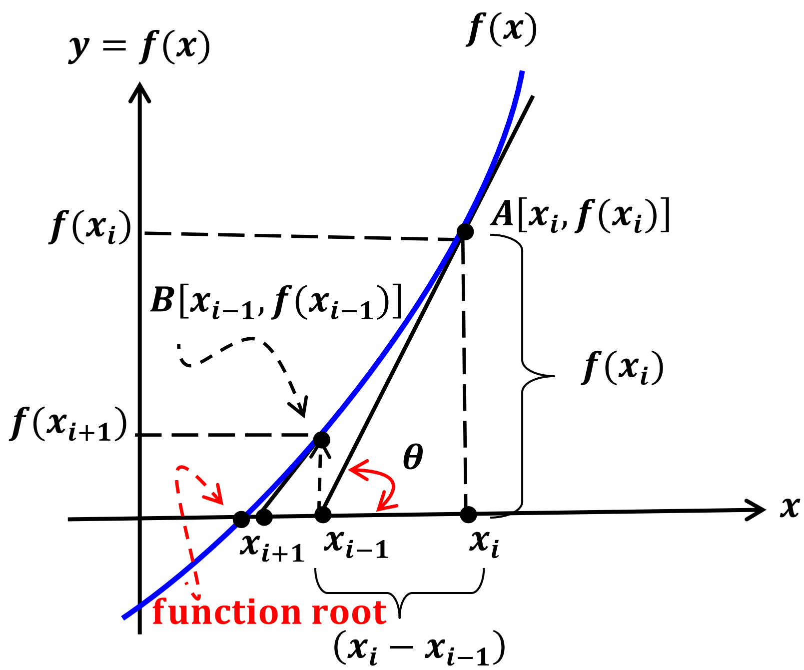
Note one **should also check** whether the **number of iterations** is **more than** the **maximum number** of **iterations allowed**. If so, **one needs** to **stop** the **algorithm**.

So, at the **seventh iteration**, we get the **final interval [1.7266, 1.7344]**
Hence, **1.7344** is the **approximated solution**.

➔ Newton-Raphson Method of Solving Nonlinear Equations

Newton-Raphson method is a **method** of **finding** the **root** of a **nonlinear equation**.

To **develop** or **derive** the **formula** of **Newton-Raphson method** for **finding** the **root** of the **nonlinear equation** $\boldsymbol{f(x) = 0}$, **suppose** we have the **function** $\boldsymbol{f(x)}$ **shown** in **figure**



Starting with **some initial guess point** $\boldsymbol{A[x_i, f(x_i)]}$ and **draw** the **tangent** at this **particular point. Extending** this **tangent** we will **find out** that is **crossing** the ***x*-axis** at **point** $\boldsymbol{x_{i-1}}$. Then **go back** to the **function** to have **new guess point** $\boldsymbol{B[x_{i+1}, f(x_{i+1})]}$ and **again draw tangent** at **this point**. **Keeping on drawing tangents**, it is seen that **this tangent** is **eventually** going to **pass very close** to the **zero** of the **function** $\boldsymbol{f(x)}$**,** or to the **root** where $\boldsymbol{f(x) = 0}$, and this is what the **basis** of the **Newton-Raphson method** is. **Continually draw** the **tangents** and see **where it crosses** the ***x*-axis** and **use it** as **new estimate** of the **root**.

Suppose that the **angle** of the **tangent** to **the function** $\boldsymbol{f(x)}$ is $\boldsymbol{\theta}$, then from the figure we have,

$$\tan\theta = \frac{f(x_i)}{x_i - x_{i+1}}$$

Since the **slope** of the **tangent line** at a **point on the function** is **equal** to the **derivative** of the **function** at the **same point**, thus we have

$$\tan\theta = f'(x_i)$$

so,

$$\Rightarrow \quad f'(x_i) = \tan\theta = \frac{f(x_i)}{x_i - x_{i+1}}$$

From this equation we can write

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

This equation is called the **Newton-Raphson formula** for finding the root of an equation.

Newton-Raphson Method Algorithm

The **algorithm** for **finding** the **root** of the **equation** $f(x) = 0$ is:

i. Calculate $f'(x)$ symbolically.
ii. Choose an initial guess (x_0).
iii. Applying Newton-Raphson formula

$$x_1 = x_0 - \frac{f(x_o)}{f'(x_0)}$$

iv. Find the absolute value of relative approximate error,

$$|\varepsilon_a| = \left|\frac{x_1 - x_0}{x_1}\right| \times 100$$

v. Check if relative approximate error, $|\varepsilon_a|$ less than or equal to pre-specified tolerance, ε_s

$$|\varepsilon_a| \le \varepsilon_s$$

- If $|\varepsilon_a| \le \varepsilon_s$, then the process has to be stopped

- If not, then repeat this process again

<u>Example</u>: Use the **Newton-Raphson method** to **find** the **root** of the **equation** $x^3 = 20$ with **initial guess** $x_0 = 3$ and **conduct three iterations.**

<u>Solution</u>:

The first thing which have to understand that the **Newton-Raphson method** is **written** for the **case** of $f(x) = 0$, so re-write the function in the form

$$f(x) = x^3 - 20 = 0$$

The general form of the **Newton-Raphson method** is:

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

<u>Iteration</u> 1:

Start with $i = 0$, and $x_0 = 3,$ we have

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = x_0 - \frac{x_0^3 - 20}{3x_0^2}$$

$x_0 = 3$:

$$\Rightarrow \quad f(x_0) = x_0^3 - 20 \Rightarrow \quad f(3) = 3^3 - 20 = \ 3^3 - 20 = 7$$

$$\Rightarrow \quad f'(x_0) = 3x^2 \qquad \Rightarrow \quad f'(x) = 3x^2 = 3 \times 3^2 = 27$$

$$\Rightarrow \ x_1 = 3 - \frac{f(x_0)}{f'(x_0)} = 3 - \frac{7}{27} = 2.741$$

The **absolute relative approximate error** is:

$$|\varepsilon_a| = \left|\frac{x_1 - x_0}{x_1}\right| \times 100 = \left|\frac{2.741 - 3}{2.741}\right| \times 100 = 9.45\ \%$$

<u>Iteration 2</u>: $i = 1$

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = x_1 - \frac{x_1^3 - 20}{3x_1^2} = 2.741 - \frac{(2.741)^3 - 20}{3 \cdot (2.741)^2} = 2.715$$

$$|\varepsilon_a| = \left|\frac{x_2 - x_1}{x_2}\right| \times 100 = \left|\frac{2.715 - 2.741}{2.715}\right| \times 100 = 0.96\ \%$$

Because the $|\boldsymbol{\varepsilon_a}| < \mathbf{5}$**, the number 2 is the correct significant digit in answer since it is less than 5.**

<u>Iteration 3</u>: $\boldsymbol{i = 2}$

$$\boldsymbol{x_3 = x_2 - \frac{f(x_2)}{f'(x_2)} = x_2 - \frac{x_2^3 - 20}{3x_2^2} = 2.715 - \frac{(2.715)^3 - 20}{3 \cdot (2.715)^2} = 2.714}$$

$$|\boldsymbol{\varepsilon_a}| = \left|\frac{\boldsymbol{x_3 - x_2}}{\boldsymbol{x_3}}\right| \times \mathbf{100} = \left|\frac{\mathbf{2.714 - 2.715}}{\mathbf{2.714}}\right| \times \mathbf{100 = 0.009} < \mathbf{0.05}\,\%$$

If we look at the **absolute relative approximate error**, we find out that is going from

$$\mathbf{9.45}\,\% \;\Rightarrow\; \mathbf{0.96}\,\% \;\Rightarrow\; \mathbf{0.009}$$

So the **Newton-Raphson method** is **converging very fast** to the **root**, and it is **less than 5 %** and **less than 0.5 %** and **less than 0.05 %** which means that **at least three significant digits** are **correct** in the **solution** because for **one significant** to **be correct** it **needs at least 5 %,** and for **two 0.5 %** and for **three 0.05 %.** So, we can **trust** in **2** and **trust in 7** and **trust in 1.** So we get **three significant digits** which are **at least correct** in the **root** of the **equation:**

$$\boldsymbol{x^3 - 20 = 0}$$

➔ <u>Secant Method of Solving Nonlinear Equations</u>

The **Newton-Raphson method** of **solving** a **nonlinear equation** $\boldsymbol{f(x) = 0}$ is given by the **iterative formula**

$$\boldsymbol{x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}}$$

One of the **drawbacks** of the **Newton-Raphson method** is that you **have to evaluate** the **derivative** of the **function**. However, this **process still** a **laborious process**, and **even intractable** if the **function** is **derived** as **part** of a **numerical scheme**. To **overcome these drawbacks**, the **derivative** of the **function**, $f(x)$ is **approximated** as

$$f'(x_i) = \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}}$$

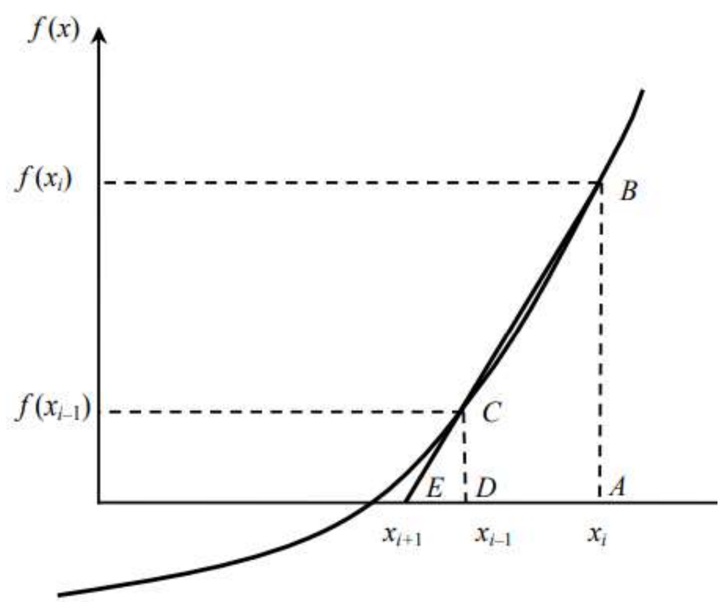
Substituting this equation **into Newton-Raphson formula** gives

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

This equation is called the **secant method**. This **method** now **requires two initial guesses**, **but unlike** the **bisection method**, the **two initial guesses** do **not need** to **bracket** the **root** of the **equation**. The **secant method** is **an open method** and **may** or **may not converge**.

However, when **secant method converges**, it will **typically converge faster** than the **bisection method**. However, since the **derivative** is **approximated** as given by above equation, it **typically converges slower** than the **Newton-Raphson method**.

The **secant method** can also be **derived** from **geometry**, as shown in figure below.



Taking two initial guesses, x_{i-1} and x_i , **one draws a straight line between** $f(x_i)$ and $f(x_{i-1})$ **passing through** the **x-axis** at x_{i+1} . ***ABE*** and ***DCE*** are **similar triangles**. Hence

$$\frac{AB}{AE} = \frac{DC}{DE} \quad \Rightarrow \quad \frac{f(x_i)}{x_i - x_{i+1}} = \frac{f(x_{i-1})}{x_{i-1} - x_{i+1}}$$

On rearranging, the **secant method** is given as

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

Secant Method Algorithm

The **algorithm** for **finding** the **root** of the **equation** $f(x) = 0$ is:

- **i.** Choose $i = 1$.
- **ii.** Start with guesses (x_{i-1}, x_i)**.**
- **iii.** Use secant method formula.
- **iv.** Find the absolute value of relative approximate error,

$$|\varepsilon_a| = \left|\frac{x_{i+1} - x_i}{x_{i+1}}\right| \times 100$$

- **v.** Check if relative approximate error, $|\varepsilon_a|$ less than or equal to pre-specified tolerance, ε_s

$$|\varepsilon_a| \leq \varepsilon_s$$

- If $|\varepsilon_a| \leq \varepsilon_s$, then the **process** has to be **stopped.**
- If **not**, then **repeat** this **process again**

<u>Example</u>: Use the **secant method** to **find** the **estimate** of the **root** of the **equation** $x^3 = 20$ with **initial guesses** $x_0 = 4$ **and** $x_1 = 5.5$, after **two iterations**.

<u>Solution</u>:

Re-write the **equation** in the **form** $f(x) = 0$ as follows:

$$f(x) = x^3 - 20 = 0$$

Applying the **formula** for the **secant method**

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

<u>Iteration 1</u>:

$$x_2 = x_1 - \frac{f(x_1)(x_1 - x_0)}{f(x_1) - f(x_0)}$$

where $x_1 = 5.5$ and $x_0 = 4$

$$\Rightarrow \quad x_2 = 5.5 - \frac{[(5.5)^3 - 20](5.5 - 4)}{[(5.5)^3 - 20] - [(4)^3 - 20]} = 3.353$$

$$|\varepsilon_a| = \left|\frac{x_2 - x_1}{x_2}\right| \times 100 = \left|\frac{3.353 - 5.5}{3.353}\right| \times 100 = 63.92\ \%$$

<u>Iteration 2</u>:

$$x_3 = x_2 - \frac{f(x_2)(x_2 - x_1)}{f(x_2) - f(x_1)}$$

where $x_1 = 5.5$ and $x_2 = 3.353$

$$\Rightarrow x_3 = 3.353 - \frac{[(3.353)^3 - 20](3.353 - 5.5)}{[(3.353)^3 - 20] - [(5.5)^3 - 20]} = 3.059$$

$$|\varepsilon_a| = \left|\frac{x_3 - x_2}{x_3}\right| \times 100 = \left|\frac{3.059 - 3.353}{3.059}\right| \times 100 = 9.691\ \%$$

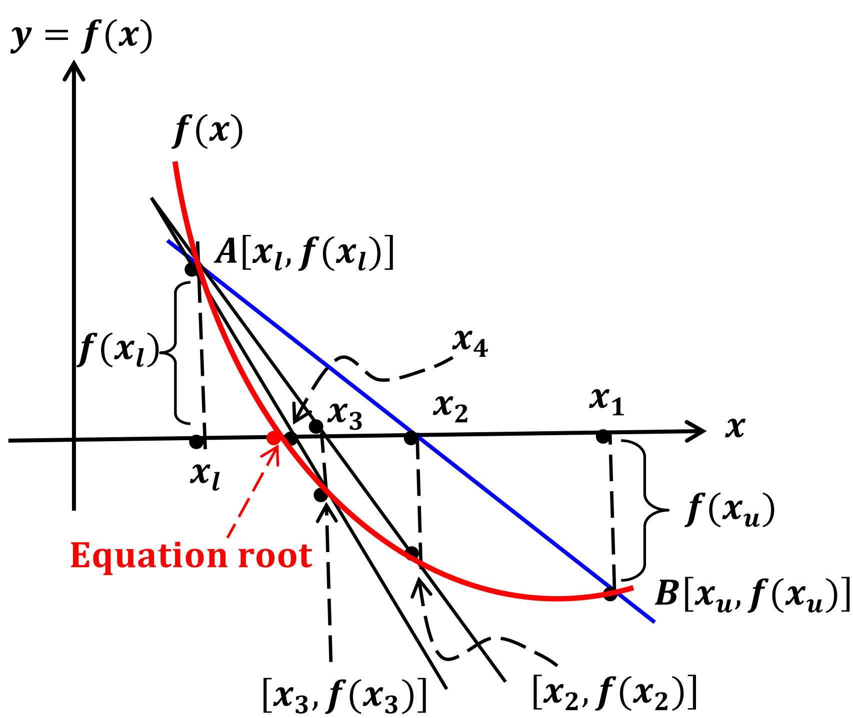
➔ False-Position Method of Solving a Nonlinear Equation

The **false-position method** is very similar to **bisection method**. However, the **next iteration point** is **not midpoint** of the **interval** $[\boldsymbol{A}, \boldsymbol{B}]$ but the **intersection** of $\boldsymbol{x}$**-axis** with the **line secant through** these **two points**; $\boldsymbol{A}[\boldsymbol{x_l}, \boldsymbol{f}(\boldsymbol{x_l})]$ and $\boldsymbol{B}[\boldsymbol{x_u}, \boldsymbol{f}(\boldsymbol{x_u})]$**.**

The **false-position method** is **used** to **estimate** the **root** by **finding** the **point** at which a **line drawn** between $\boldsymbol{x_l}$ and $\boldsymbol{x_u}$ **crosses** the $\boldsymbol{x}$**-axis**. That is, performing **linear interpolation between** $\boldsymbol{x_l}$ and $\boldsymbol{x_u}$ to **find** the **approximate root**.

Like the **bisection method**, the **false-position method** will always **converge**, and generally it will **converge faster than** the **bisection method**. Unfortunately, this **general rule** is **not always true**, and sometimes the **bisection method converges faster than** the **false-position method**.

To **get** the **formula** for the **false-position method**, consider the **function** $\boldsymbol{f}(\boldsymbol{x})$ **drawn** in the **figure.**



Recall that the **equation** of the **line** between the **two points** $\boldsymbol{A}[\boldsymbol{x_l}, \boldsymbol{f}(\boldsymbol{x_l})]$ and $\boldsymbol{B}[\boldsymbol{x_u}, \boldsymbol{f}(\boldsymbol{x_u})]$**.** which is **given by**

$$\frac{y - y_1}{x - x_1} = \frac{y_2 - y_1}{x_2 - x_1} \quad \text{or} \quad y - y_1 = \frac{y_2 - y_1}{x_2 - x_1}(x - x_1)$$

$$\Rightarrow \quad y_r - f(x_l) = \frac{f(x_u) - f(x_l)}{x_u - x_l}\,(x_r - x_l)$$

$$\Rightarrow \quad y_r = f(x_l) + \frac{f(x_u) - f(x_l)}{x_u - x_l}\,(x_r - x_l)$$

Hence, **the intersection** of **this line** with the ***x*-axis**, where $\boldsymbol{y_r = 0}$**,** then the equation will be as

$$0 = f(x_l) + \frac{f(x_u) - f(x_l)}{x_u - x_l}\,(x_r - x_l)$$

$$\Rightarrow \quad (x_r - x_l)\frac{f(x_u) - f(x_l)}{x_u - x_l} = -f(x_l)$$

$$\Rightarrow \quad (x_r - x_l) = \frac{-f(x_l)}{\frac{f(x_u) - f(x_l)}{x_u - x_l}} = -f(x_l)\frac{x_u - x_l}{f(x_u) - f(x_l)}$$

This **leads directly** to

$$x_r = x_l - f(x_l)\frac{x_u - x_l}{f(x_u) - f(x_l)}$$

This is the **formula** for the **false-position method**.

The **next predicted root** $\boldsymbol{x_r}$ would be **closer** to the **equation root** as shown in the figure. The **false-position method** takes **advantage** of this **observation mathematically** by **drawing** a **secant from** the **function value** at $\boldsymbol{x_l}$ to the **function value** at $\boldsymbol{x_u}$, and **estimates** the **root** as where it **crosses** the ***x*-axis**. Thus, the **new upper** and **lower bounds** are then **established**, and the **procedure** is **repeated until** the **convergence** is **achieved such that** the **new lower** and **upper bounds** are **sufficiently close** to **each other**.

False-Position Algorithm

The **steps** to **apply** the **false-position method** to **find** the **root** of the **equation** $\boldsymbol{f(x) = 0}$ are as follows.

1. **Choose** $\boldsymbol{x_l}$ and $\boldsymbol{x_u}$ as **two guesses** for the **root** such that

$$\boldsymbol{f(x_l) \cdot f(x_u) < 0}$$

   or in other words, $\boldsymbol{f(x)}$ **changes sign between** $\boldsymbol{x_l}$ and $\boldsymbol{x_u}$.

2. **Estimate** the **root** $\boldsymbol{x_r}$ of the **equation** $\boldsymbol{f(x) = 0}$ as

$$\boldsymbol{x_r = x_l - f(x_l)\frac{x_u - x_l}{f(x_u) - f(x_l)}}$$

$$\boldsymbol{x_2 = \frac{x_0 f(x_l) - x_1 f(x_0)}{f(x_1) - f(x_0)}}$$

3. **Check** the **following**:

   i. **If the product** $\boldsymbol{f(x_r)f(x_u) < 0}$**; Then** $\boldsymbol{x_l = x_r}$ and $\boldsymbol{x_u = x_u}$**,** the **root must lie** in the **interval** $\boldsymbol{[x_l, x_r]}$**.**
   ii. **If the product** $\boldsymbol{f(x_r)f(x_u) > 0}$: **Then** $\boldsymbol{x_l = x_l}$ and $\boldsymbol{x_u = x_r}$**,** the **root must lie** in the **interval** $\boldsymbol{[x_r, x_u]}$**.**
   iii. **If** $\boldsymbol{f(x_r)f(x_u) = 0}$; **Then** $\boldsymbol{x_r}$ is the **root** of the **equation** and **stop** the **process**

4. **Repeat** the **process** to **find** the **new estimate** of the **root** $\boldsymbol{x_{r1}}$ by

$$\boldsymbol{x_{r1} = x_r - f(x_r)\frac{x_u - x_r}{f(x_u) - f(x_r)}}$$

$$\boldsymbol{x_2 = \frac{x_0 f(x_l) - x_1 f(x_0)}{f(x_1) - f(x_0)}}$$

5. **Find** the **step size** ϵ_s and **absolute relative approximate error** $|\epsilon_a|$ by

$$|\epsilon_s| = \left|x_r^{new} - x_r^{old}\right|$$

$$|\epsilon_a| = \left|\frac{x_r^{new} - x_r^{old}}{x_r^{new}}\right|$$

Where

x_r^{new} – **estimated root** from present iteration

x_r^{old} **– estimated root** from **previous iteration**

Compare the **absolute relative approximate error** $[\epsilon_a]$ with the **pre-specified relative error tolerance** ϵ_s**:**

- **If** $|\epsilon_a| < \epsilon_s$**; stop the process**
- **If** $[\epsilon_a] > \epsilon_s$, then **go back** to **Step 2** and **repeat** the **process**.

6. **Go back** to **step 2** and **repeat** the **process until** a **certain level** of **convergence** has been **achieved**.

- **Note** one should also **check whether** the **number of iterations** is **more than** the **maximum number** of **iterations allowed**. If so, one needs to **terminate** the **algorithm**.
- **Note** that the **false-position** and **bisection algorithms** are **quite similar**. The **only difference** is the **formula used** to **calculate** the **new estimate** of the **root** x_r as **shown** in **steps 2** and **4.**

Example 1: Use a False-Position Method (Regula Falsi Method) to **find** the **root** of the **nonlinear equation**

$$f(x) = x^2 - 3,$$

if $\varepsilon_s = 0.01, \varepsilon_a = 0.01,$ and **start** with the **interval** $[1, 2]$**.**

Solution:

Rewrite the equation in the form

$$f(x) = x^2 - 3 = 0$$

Let $x_l = 1$ and $x_u = 2$

Check if the **function changes sign** to ensure that the **root lies** between the **lower** and **upper bounds** by applying the **rule**

$$f(x_l) \cdot f(x_u) < 0$$

$$f(x_l) = f(1) = 1^2 - 3 = -2 \quad \text{and} \quad f(x_u) = f(2) = 2^2 - 3 = 1$$

$$f(x_l) \cdot f(x_u) = f(1) \cdot f(2) = -2 \times 1 = -2 < 0$$

Thus, the root lies between $x_l = 1$ and **$x_u = 2$.**

<u>**Iteration 1**</u>**:**

Applying False-Position Method by **using** its **formula**

$$x_r = x_l - f(x_l) \cdot \frac{x_{u-}x_l}{f(x_u) - f(x_l)} = 1 - f(1) \cdot \frac{2 - 1}{f(2) - f(1)}$$

$$\Rightarrow \quad x_r = 1 - (-2) \cdot \frac{2 - 1}{1 - (-2)} = 1.6667$$

$$f(x_l) \cdot f(x_u) = f(x_r) \cdot f(x_u) < 0$$

$$f(1.6667) = (1.6667)^2 - 3 = -0.2221$$

$$f(x_r) \cdot f(x_u) = f(1.6667) \cdot f(2) = -0.2221 \times 1 = -0.2221 < 0$$

Thus, the root lies between $x_l = x_r = 1.6667$ and **$x_u = 2$.**

$$\varepsilon_{step} = |x_r - x_l| = |0.6667 - 1| = 0.6667 > 0.01$$

$$|\varepsilon_a| = \left|\frac{x_r^{new} - x_r^{old}}{x_r^{new}}\right| = \left|\frac{1.6667 - 1}{1.6667}\right| = 0.40 > 0.01$$

<u>Iteration 2</u>:

$$x_l = x_r = 1.6667 \text{ and } x_u = 2$$

$$x_{r1} = x_r - f(x_r) \cdot \frac{x_u - x_r}{f(x_u) - f(x_r)}$$

$$= 1.6667 - (-0.2221) \cdot \frac{2 - 1.6667}{1 - (-0.2221)}$$

$$\Longrightarrow \quad x_{r1} = 1.7273$$

$$f(x_{r1}) \cdot f(x_u) < 0$$

$$f(x_{r1}) = f(1.7273) = (1.7273)^2 - 3 = -0.0164$$

$$f(x_{r1}) \cdot f(x_u) = -0.0164 \times 1 = -0.0164 < 0$$

Thus, the root lies between $x_l = x_{r1} = 1.7273$ and $x_u = 2$.

$$\varepsilon_{step} = |x_{r1} - x_r| = |1.7273 - 1.6667| = 0.0606 > 0.01$$

$$|\varepsilon_a| = \left|\frac{x_r^{new} - x_r^{old}}{x_r^{new}}\right| = \left|\frac{1.7273 - 1.6667}{1.7273}\right| = 0.035 > 0.01$$

<u>Iteration 3</u>:

$$x_l = x_{r1} = 1.7273 \quad \text{and} \quad x_u = 2$$

$$x_{r2} = x_{r1} - f(x_{r1}) \cdot \frac{x_u - x_{r1}}{f(x_u) - f(x_{r1})}$$

$$= 1.7273 - (-0.0164) \cdot \frac{2 - 1.7273}{1 - (-0.0164)} = 1.7317$$

$$f(x_{r2}) \cdot f(x_u) < 0$$

$$f(1.7317) = (1.7317)^2 - 3 = -0.0012$$

$$f(x_{r2}) \cdot f(x_u) = -0.0012 \times 1 = -0.0012 < 0$$

Thus, the root lies between $x_l = x_{r2} = 1.7317$ and **$x_u = 2.$**

$$\varepsilon_{step} = |x_{r2} - x_{r1}| = |1.7317 - 1.7273| = 0.0044 < 0.01$$

$$|\varepsilon_a| == \left|\frac{x_r^{new} - x_r^{old}}{x_r^{new}}\right| = \left|\frac{1.7317 - 1.7273}{1.7317}\right| = 0.0025 < 0.01$$

Thus, with the **third iteration**, the **last step $1.7273 \implies 1.7317$** has step size and **$|\varepsilon_a| = 0.0025$ less than 0.01**, and therefore **$x_{r2} = 1.7317$** is an **approximation** of the **root**.

Or **finding** the **solution** by **making a table** and **insert all** the **obtained results** such as

x_l	x_u	$f(x_l)$	$f(x_u)$	x_r	$f(x_r)$	Update	Step Size
1	**2**	-2	**1**	**1.6667**	-0.2221	$x_l = x_r$	0.6667
1.6667	**2**	-0.2221	**1**	**1.7273**	-0.0164	$x_l = x_{r1}$	0.0606
1.7273	**2**	-0.0164	**1**	**1.7317**	-0.0012	$x_l = x_{r2}$	0.0044

<u>Example</u> 2: Use a False-Position Method to **find** the **root** of the **equation**

$$f(x) = e^x(3.2\, sin\, x - 0.5\, cosx),$$

on the **interval** $[3, 4]$, the **time** with $\varepsilon_s = 0.001, \varepsilon_a = 0.001$

<u>Solution:</u>

Let $x_l = 3$ and **$x_u = 5$**

Check if the **function changes sign** and the **root lies** between the **lower** and **upper bounds** by applying the **rule**

$$f(x_l) \cdot f(x_u) < 0$$

$$f(x_l) = f(3) = e^3(3.2 \sin 3 - 0.5 \cos 3) = 0.0471$$ and

$$f(x_u) = f(5) = e^5(3.2 \sin 5 - 0.5 \cos 5) = -0.0384$$

$$f(x_l) \cdot f(x_u) = f(3) \cdot f(5) = 0.0471 \times -0.0384 = -2 < 0$$

Thus, the root lies between $x_l = 3$ and **$x_u = 5$.**

i	x_l	x_u	$f(x_l)$	$f(x_u)$	x_r	$f(x_r)$	Update	Step Size
1	**3**	**4**	0.0471	-0.0384	**3.5513**	-0.0234	$x_u = x_r$	0.4487
2	**3**	**3.5513**	0.0471	-0.0234	**3.3683**	-0.008	$x_u = x_{r1}$	0.1830
3	**3**	**3.3683**	0.0471	-0.008	**3.3149**	-0.0022	$x_u = x_{r2}$	0.0534
4	**3**	**3.3149**	0.0471	-0.0022	**3.3010**	-0.0005	$x_u = x_{r3}$	0.0139
5	**3**	**3.3010**	0.0471	-0.0005	**3.2978**	-0.0002	$x_u = x_{r4}$	0.0032
6	**3**	**3.2978**	0.0471	-0.0002	**3.2969**	-0.00004	$x_u = x_{r5}$	0.0009

Thus, **after** the **sixth iteration**, the **final step**, $3.2978 \Longrightarrow 2.2969$ has **step size 0.0009** and $|\varepsilon_a| = 0.0003$ **<u>less than</u>** the given **0.001** and therefore $x_{r5} = 3.2969$ is the **approximation** of the **root**.

➔ Fixed Point Iteration Method

To **find** the **approximate solutions** of the **equation**

$$f(x) = 0$$

First **rewrite** the **equation** in the **form**

$$x = g(x)$$

in such a way that any solution of this equation, which is a **fixed point** of **g**, is a **solution** of the **equation** $f(x) = 0$.

The **algorithm starts from any point** x_0 and then to **use** the **iteration** with an **initial guess** x_0 chosen, compute a sequence

$$x_{n+1} = g(x_n) \quad n = 0, 1, 2, 3, \ldots.$$

If the **function** $f(x)$ is **continuous** and (x_n) **converges** to some l_0 then it is clear that l_0 is a **fixed point** of **g** and hence it is a **solution** of the **equation** $f(x)$. Moreover, x_n (for a **large** n) can be **considered** as an **approximate solution** of the **equation** $f(x) = 0$.

Example: Find the root of the equation $x^3 - 7x + 2 = 0$ using **Fixed Point Method.** if $\varepsilon_s = 0.0001$.

Solution:

$$f(x) = x^3 - 7x + 2 = 0$$

To **determine** the **interval between** which the **root is laying**, **guess** the **lower** and **upper bound** as follows:

x	$f(x)$
-2	$+8$
-1	$+8$
0	$+2$
$+1$	-4

Thus, $[0, 1]$ is the **interval between** which the **root** is **laying**

Rewrite equation in the **form**

$$7x = x^3 + 2 \quad \Longrightarrow \quad x = \frac{x^3 + 2}{7}$$

Next, **define** the **process** by

$$x_{n+1} = \frac{x_n^3 + 2}{7}$$

$$x_n = 0 \quad \Longrightarrow \quad x_{0+1} = \frac{x_0^3 + 2}{7} \quad \Longrightarrow \quad x_1 = \frac{0^3 + 2}{7} = \frac{2}{7} = 0.286$$

$$x_n = 1 \quad \Longrightarrow \quad x_{1+1} = \frac{x_1^3 + 2}{7} \Longrightarrow x_2 = \frac{(0.286)^3 + 2}{7} = \frac{2}{7} = 0.2891$$

$$\varepsilon_s = \left|x_n^{new} - x_n^{old}\right| = |0.289 - 0.286| = 0.003$$

$$x_n = 2 \quad \Longrightarrow \quad x_{2+1} = \frac{x_2^3 + 2}{7} \Longrightarrow x_3 = \frac{(0.2891)^3 + 2}{7} = \frac{2}{7} = 0.2892$$

$$\varepsilon_s = \left|x_n^{new} - x_n^{old}\right| = |0.2892 - 0.2891| = 0.0001$$

Thus, $x_3 = 0.2892$ is **an approximation** of the **root**.

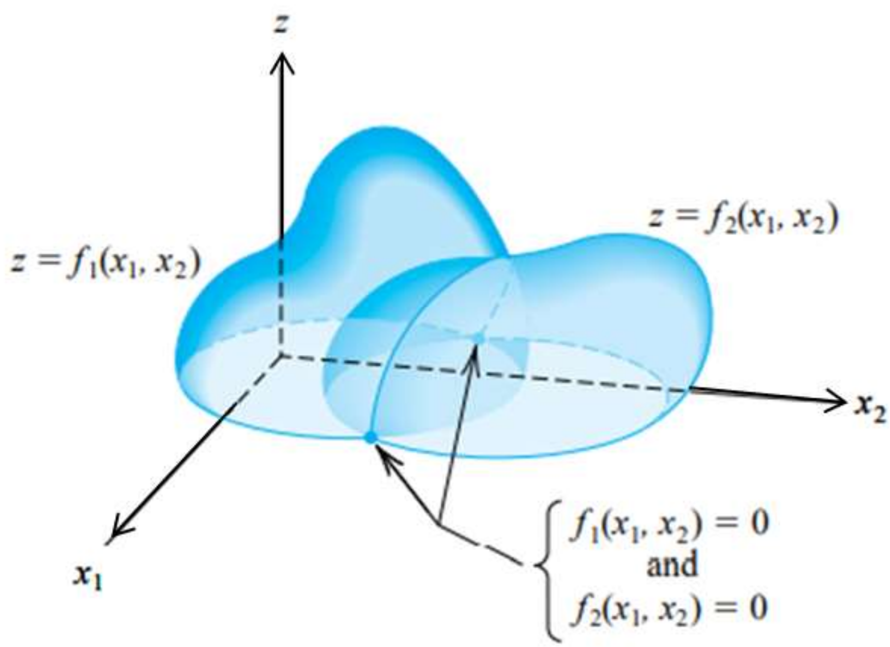
System of nonlinear equations

A system of nonlinear equations is a **system** of **equations containing** at least **one equation** that is of **degree larger than one**. It is a **system** of **two** or **more equations** in **two** or **more variables** containing **at least** one **equation** that is **not linear**.

A **system** of **nonlinear equations** has the form

$$\begin{aligned} f_1(x_1, x_2, x_3, \cdots, x_n) &= 0 \\ f_2(x_1, x_2, x_3, \cdots, x_n) &= 0 \\ f_3(x_1, x_2, x_3, \cdots, x_n) &= 0 \\ \vdots \qquad\qquad &\quad \vdots \\ f_n(x_1, x_2, x_3, \cdots, x_n) &= 0 \end{aligned}$$

A **geometric representation** of a **nonlinear system** when **n = 2** is **given** in the **figure**.



A **general system** of ***n*** **nonlinear equations** in ***n*** **unknowns** can be **alternatively represented** by **defining a function *F***, by

$$F(x_1, x_2, x_3, \cdots, x_n) = [f_1(x_1, x_2, x_3, \cdots, x_n), f_2(x_1, x_2, x_3, \cdots, x_n), f_3(x_1, x_2, x_3, \cdots, x_n), \cdots, f_n(x_1, x_2, x_3, \cdots, x_n)]^t$$

the **nonlinear system assumes** the **form**

$$F(x_1, x_2, x_3, \cdots, x_n) = 0$$

The **functions** $f_1, f_2, f_3, \cdots, f_n$ are the **coordinate functions** of F.

The **substitution method used** for **linear systems** is the **same method** used for **nonlinear systems**. **Solve one equation** for **one variable** and then **substitute** the **result** into the **second equation** to **solve** for **another variable**, and **so on**. There is, however, a **variation** in the **possible outcomes**.

An example is:

$$x^2 - 6x + 9 = 0$$

$$x - cosx = 25$$

$$e^x + lnx^2 - xcosx = 0$$

The **techniques** of **solution** that have been **used** were **analytic** and the obtained **results** were **precise** and **tailored** to the **type** of **equation**

There is a **crucial difference between** the **analytic** and **numerical approach**. With the **analytic** approach which **starts** with **discussion** of **domain** of f, and then **solve** the **equation** using a **type-tailored method**. Finally, **check whether** the **solution belongs** to the **domain** of f.

However, in the **numerical approach**, first of all **must roughly localize** the **solution** and then either **narrow** the **respective interval** or **approach** the **solution using iteration**. Thus, the **methods** of **calculation** in **numerical approach** are **divided** into **two groups**:

i. **Narrowing** the **interval** or

ii. **Iteration** from a **given estimate**.

Methods for Solving a System of Nonlinear Equations

- **Solve** a **system** that **represents** the **intersection** of a **parabola** and a **line using substitution**.
- **Solve** a **system** that **represents** the **intersection** of a **circle** and a **line using substitution**.
- **Solve** a **system** that **represents** the **intersection** of a **circle** and an **ellipse using elimination**.

➔ Intersection of a Parabola and a Line

There are **three possible types** of **solutions** for a **system of nonlinear equations** involving a **parabola** and a **line**.

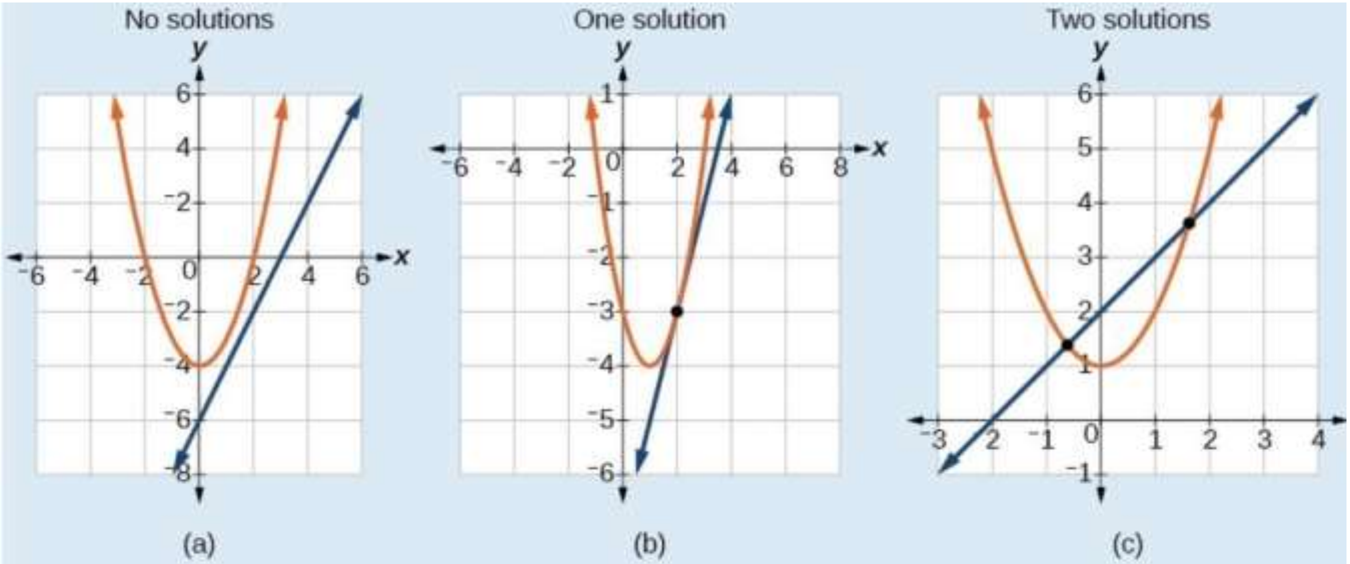
Possible Types of Solutions for the Points of Intersection of a Parabola and a Line

The **graphs** below **illustrate possible solution sets** for a **system** of **equations involving** a **parabola** and a **line**.

(a) **No solution**: The **line** will **never intersect** the **parabola (Figure a)**.

(b) **One solution**: The **line** is **tangent** to the **parabola** and **intersects** the **parabola** at **exactly** **one point** **(Figure b)**.

(c) **Two solutions**: The **line crosses** on the **inside** of the **parabola** and **intersects** the **parabola** at **two points** **(Figure c)**.



- **Procedure of Finding the Solution of a System of Equations Containing a Line and a Parabola**

1. **Solve** the **linear equation** for **one** of the **variables.**
2. **Substitute** the **expression obtained** in **step one into** the **parabola equation**.
3. **Solve** for the **remaining variable**.
4. **Check obtained solutions** in **both equations**.

Example 1: **Solve** the **system** of **equations**

$$\boldsymbol{x - y = -1}$$
$$\boldsymbol{y = x^2 + 1}$$

Solution:

Solve the **first equation** for ***x***:

$$\boldsymbol{x - y = -1 \quad \Rightarrow \quad x = y - 1}$$

substitute the resulting expression **into** the **second equation:**

$$y = x^2 + 1 = (y-1)^2 + 1 = y^2 - 2y + 1 + 1$$

$$\Rightarrow \quad y^2 - 3y + 2 = 0 \quad \Rightarrow \quad (y-2)(y-1) = 0$$

Solving for **y** gives: $y = 2$ and $y = 1$

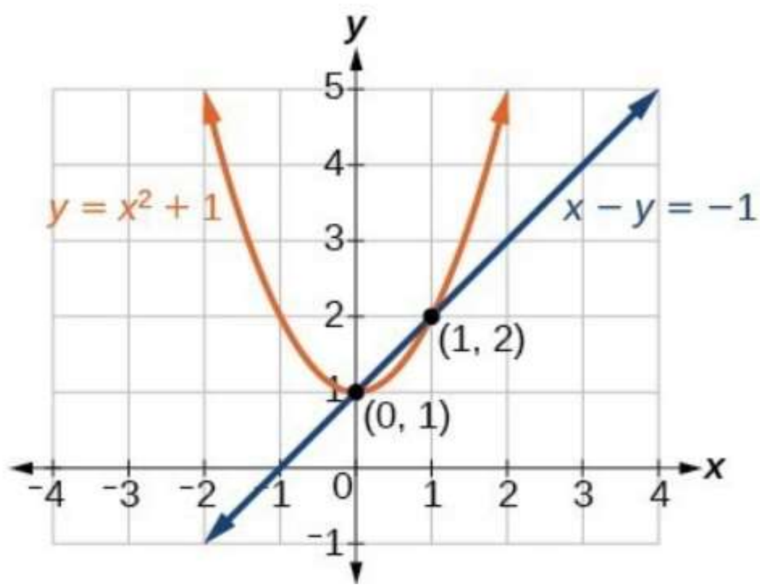
Substitute each value for **y into** the **first equation** to **solve** for x**:**

$$y = 2 \quad \Rightarrow \quad x = y - 1 = 2 - 1 = 1$$

$$y = 1 \quad \Rightarrow \quad x = y - 1 = 1 - 1 = 0$$

The **solutions** are: $(1, 2)$ and $(0, 1)$

which can be **verified** by **substituting** these (x, y) **values into both** of the **original equations**



__Example__ **2: Solve** the **given system** of **equations** by **substitution**.

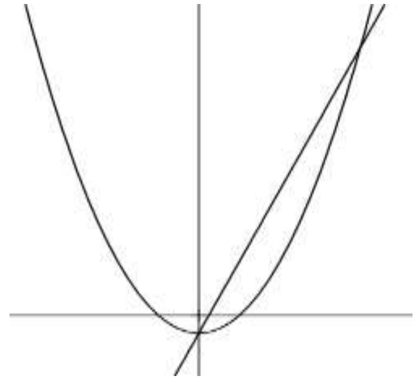
$$3x - y = -2$$
$$2x^2 - y = 0$$

__Solution__**:**

$$\left(-\frac{1}{2}, \frac{1}{2}\right) \text{ and } (2, 8)$$

<u>Example</u> 3: Find the solution to the **system** of linear **equations** given by

$$y = 4x - 1$$
$$y = x^2 - 1$$



<u>Solution</u>:

$(4, 15)$ and $(0, -1)$

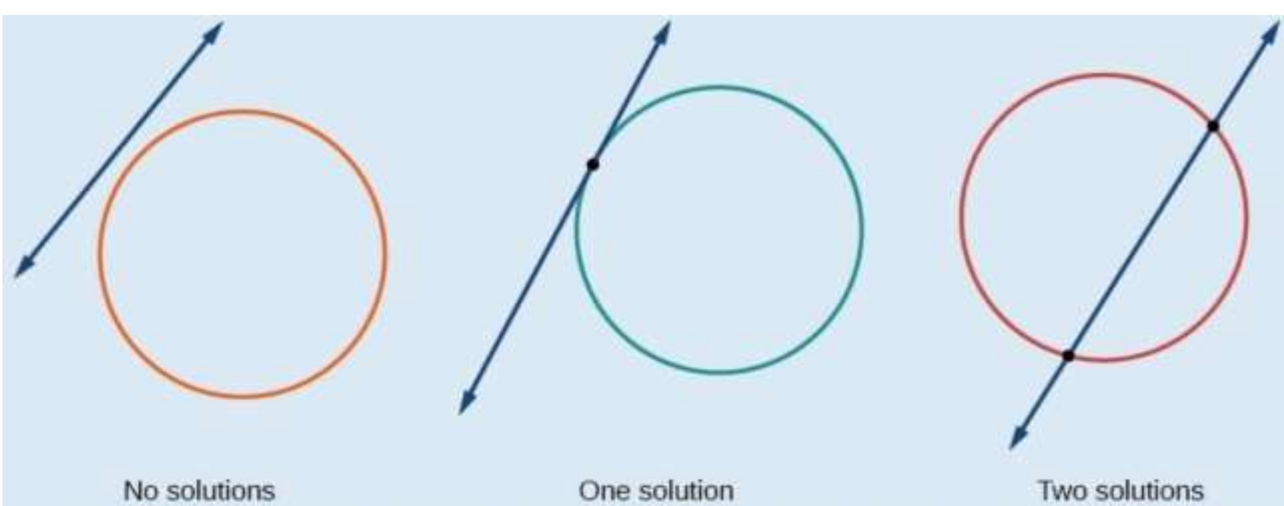
➔ <u>Intersection of a Circle and a Line</u>

Just as with a **parabola** and **a line**, there are **three possible outcomes** when solving a system of equations representing a circle and a line.

<u>Possible Types of Solutions for the Points of Intersection of a Circle and a Line</u>

The **graphs** below **illustrate possible solution sets** for a **system** of **equations involving** a **circle** and a **line**.

- **(a)** **<u>No solution</u>**: The **line does not intersect** the **circle**.
- **(b)** **<u>One solution</u>**: The **line** is **tangent** to the **circle** and **intersects** the **circle** at **exactly one point**.
- **(c)** **<u>Two solutions</u>**: The **line crosses** the **circle** and **intersects** the **circle** at **two points**.



➢ <u>Procedure of Finding the Solution of a System of Equations Containing a Line and a Circle</u>

1. **<u>Solve</u>** the **<u>linear equation</u>** for **one** of the **variables.**
2. **Substitute** the **expression obtained** in **<u>step one</u> into** the **<u>circle equation</u>**.
3. **<u>Solve</u>** for the **remaining variable**.
4. **<u>Check</u> obtained solutions** in **both equations**.

<u>Example</u> 4: Find the **intersection** of the **given circle** and the **given line** by **substitution**

$$\boldsymbol{x^2 + y^2 = 5}$$
$$\boldsymbol{y = 3x - 5}$$

<u>Solution</u>:

Substitute $\boldsymbol{y = 3x - 5}$ **into** the **equation** for the **circle**, yields

$$\boldsymbol{x^2 + y^2 = x^2 + (3x - 5)^2 = 5}$$

$$\Longrightarrow \quad \boldsymbol{x^2 + 9x^2 - 30x + 25 = 5} \quad \Longrightarrow \quad \boldsymbol{10x^2 - 30x + 20 = 0}$$

$$\Longrightarrow \quad \boldsymbol{10(x^2 - 3x + 2) = 0}$$

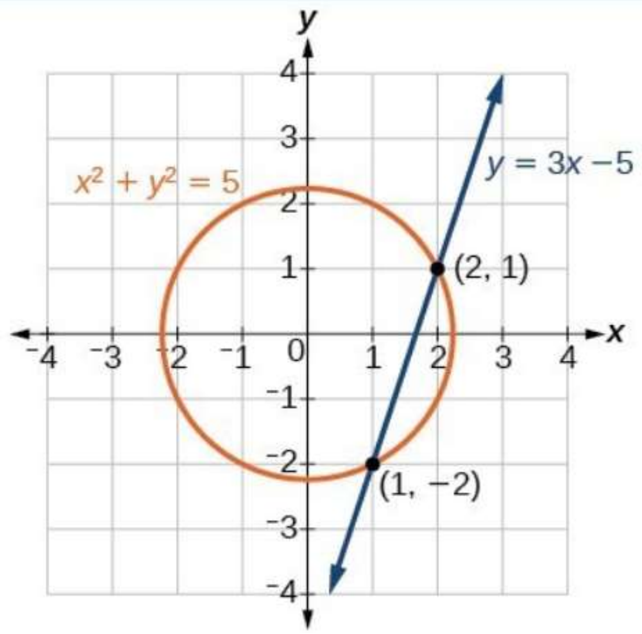
Factor and **solve** for ***x***, yields

$$\boldsymbol{(x - 2)\,(x - 1) = 0} \quad \Longrightarrow \quad \boldsymbol{x = 2} \quad \text{and} \quad \boldsymbol{x = 1}$$

Substitute the **values** of ***x*** into the **equation** for **y**, we have

$$\boldsymbol{x = 2 \quad \Longrightarrow \quad y = 3(2) - 5 = 1}$$
$$\boldsymbol{x = 1 \quad \Longrightarrow \quad y = 3(1) - 5 = -2}$$

The **line <u>intersects</u>** the **circle** at two points: $\boldsymbol{(2, 1)}$ and $\boldsymbol{(1, -2)}$, which can be **verified** by **substituting these values into both** of the **original equations.**



Example 5: Solve the system of nonlinear equations

$$\boldsymbol{x^2 + y^2 = 10}$$
$$\boldsymbol{x - 3y = -10}$$

Solution:

$$(\boldsymbol{-1, 3})$$

Example 6: Solve the following **system** of **non-linear equations**

$$(\boldsymbol{x - 10})^2 + (\boldsymbol{y - 8})^2 = \boldsymbol{100}$$
$$\boldsymbol{-3x + y = 8}$$

Solution:

$$(\boldsymbol{0, 8}) \text{ and } (\boldsymbol{2, 14})$$

Solving a System of Nonlinear Equations Using Elimination

It has seen that **substitution** is often the **preferred method** when a **system** of **equations** includes a **linear equation** and a **nonlinear equation**. However, when **both equations** in the system **have like variables** of the **second degree**, **solving them** using **elimination** by **addition** is **often easier**

than substitution. Generally, **elimination** is a **far simpler method** when the **system involves** only **two equations** in **two variables** (a **two-by-two system**), **rather than** a **three-by-three system**, as there are **fewer steps**. As **an example**, it will be **investigated** the **possible types** of **solutions** when **solving a system** of **equations representing** a **circle** and **an ellipse**.

Possible Types of Solutions for the Points of Intersection of a Circle and a Line

The **figures** below **illustrate possible solution sets** for a **system** of **equations involving** a **circle** and an **ellipse**.

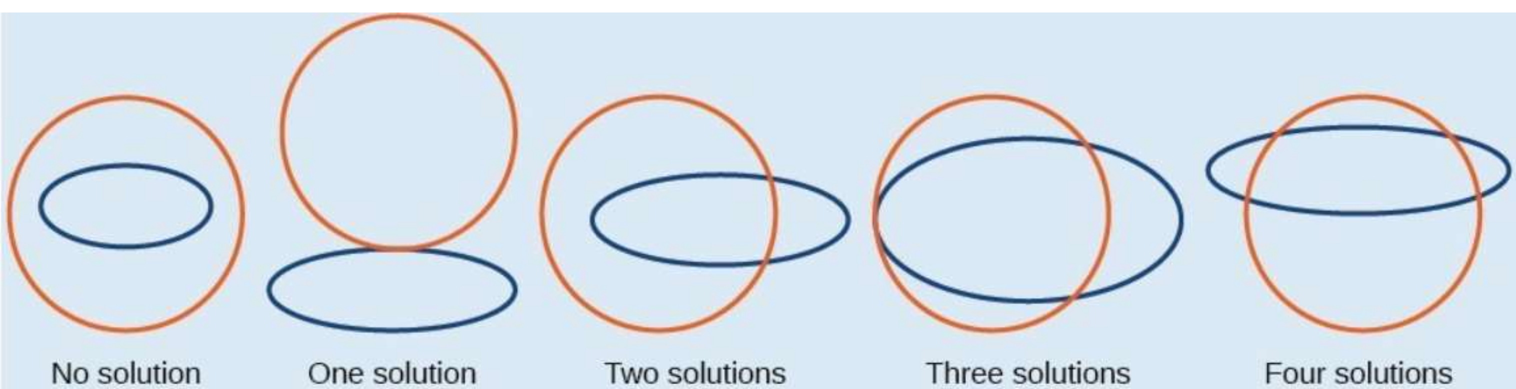
(a) **No solution**: The **circle** and **ellipse do not intersect.** One **shape** is **inside** the **other**, or the **circle** and the **ellipse** are a distance away from the other

(b) **One solution**: The **circle** and **ellipse** are **tangent** to each other and **intersect** at **exactly** **one point**.

(c) **Two solutions**: The **circle** and **ellipse intersect** at **two points**.

(d) **Three solution:** The **circle** and **ellipse intersect** at **three points**.

(e) **Four solution:** The **circle** and **ellipse intersect** at **four points**.



<u>Example</u> 7: Find the **solution** set of the given **system** of **nonlinear equations**

$$x^2 + y^2 = 26 \qquad (1)$$

$$3x^2 + 25y^2 = 100 \qquad (2)$$

<u>Solution</u>:

<u>Multiply</u> Eq. (1) by (-3) and **<u>add</u>** to the **Eq. (2),** yields

$$-3x^2 + -3y^2 = -78 \qquad (1)$$

$$3x^2 + 25y^2 = 100 \qquad (2)$$

$$22y^2 = 22$$

Solve for **y**, we have

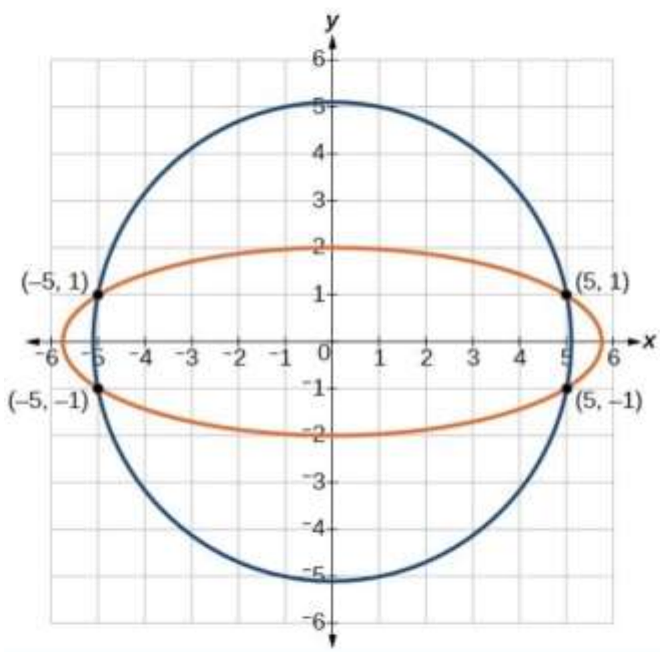
$$22y^2 = 22 \quad \Rightarrow \quad y^2 = 1 \quad \Rightarrow \quad y = \pm 1$$

Substitute $y = \pm 1$ **into one** of the **equations** and **solve** for x, we get

$$x^2 + 1^2 = 26 \;\Rightarrow\; x^2 = 25 \;\Rightarrow\; x = \pm 5$$

$$x^2 + (-1)^2 = 26 \;\Rightarrow\; x^2 = 25 \;\Rightarrow\; x = \pm 5$$

Thus, there are **<u>four solutions</u>**



<u>**Example**</u> **5: Solve** the following **system** of **nonlinear equations**

$$\begin{aligned} 4x^2 + y^2 &= 13 \\ x^2 + y^2 &= 10 \end{aligned}$$

<u>**Solution**</u>**:**

$$(1,3), (1,-3), (-1,3), \text{and } (-1,-3)$$

<u>Iterative Methods for Solving System of nonlinear equations</u>

A system of nonlinear equations is a **system** of **equations containing** at least **one equation** that is of **degree larger than one**. It is a **system** of **two** or **more equations** in **two** or **more variables** containing **at least** one **equation** that is **not linear**.

A **system** of **nonlinear equations** has the form

$$\left.\begin{aligned} f_1(x_1, x_2, x_3, \cdots, x_n) &= 0 \\ f_2(x_1, x_2, x_3, \cdots, x_n) &= 0 \\ f_3(x_1, x_2, x_3, \cdots, x_n) &= 0 \\ \vdots \qquad\qquad & \quad \vdots \\ f_n(x_1, x_2, x_3, \cdots, x_n) &= 0 \end{aligned}\right\} \quad \ldots\ldots\ldots (1)$$

Algebraic and **transcendental equations** that **do not fit** this **format** are, For example,

$$\begin{aligned} x^2 + xy &= 10 \\ y + 3xy^2 &= 57 \end{aligned}$$

are **two simultaneous nonlinear equations** with **two unknowns**, x and **y**. They **can be expressed** in the **form** of **Equation (1):**

$$\begin{aligned} u(x,y) &= x^2 + xy - 10 = 0 \\ v(x,y) &= y + 3xy^2 - 57 = 0 \end{aligned}$$

Thus, the **solution** would be the **values** of x and **y** that **make** the **functions** $u(x, y)$ and $v(x, y)$ **equal** to **zero**. **Most methods** for **determining** such **solutions** are **extensions** of the **open methods** for **solving single equations** such as **fixed-point iteration** and **Newton-Raphson iteration method**.

➔ Fixed-Point Method for Solving a Nonlinear System

The **fixed-point-iteration method** can be **modified** to **solve two simultaneous, nonlinear equations**. This **method** is **illustrated** in the **following example**.

Example: **Use fixed-point iteration** to **determine** the **roots** of the **following two nonlinear equations**

$$x^2 + xy = 10$$
$$y + 3xy^2 = 57$$

Note that a **correct pair** of **roots** is $x = 2$ and **y = 3**. **Initiate** the **computation** with **guesses** of $x = 1.5$ and **y = 3.5**.

Solution:

Rewritten the **equations** in the **form:**

$$x = \frac{10 - x^2}{y}$$
$$y = 57 - 3xy^2$$

Using the **initial guesses**, $x = 1.5$ and **y = 3.5**, yields

$$x_1 = \frac{10 - x_0^2}{y_0} = \frac{10 - (1.5)^2}{3.5} = 2.214$$

$$y_1 = 57 - 3x_0y_0^2 = 57 - 3(1.5)(3.5)^2 = -24.375$$

Thus, the **approach seems** to be **diverging**.

$$\boldsymbol{x_2 = \frac{10 - x_1^2}{y_1} = \frac{10 - (2.214)^2}{-24.375} = -0.209}$$

$$\boldsymbol{y_2 = 57 - 3x_1y_1^2 = 57 - 3(2.214)(-24.375)^2}$$

$$\boldsymbol{\Longrightarrow \quad y_2 = 429.709}$$

Obviously, the **approach** is **deteriorating**.

Now **repeating** the **computation** but **with** the **original equations set up** in a **different format** as follows:

$$\boldsymbol{x^2 + \;\; xy \;\; = 10 \quad \Longrightarrow \quad x = \sqrt{10 - xy}}$$

$$\boldsymbol{y + 3xy^2 = 57 \quad \Longrightarrow \quad y = \sqrt{\frac{57 - y}{3x}}}$$

Applying the **given initial guesses**, ***x*** **= 1.5** and **y = 3.5**, yields

$$\boldsymbol{x_1 = \sqrt{10 - x_0y_0} = \sqrt{10 - (1.5)(3.5)} = 2.179}$$

$$\boldsymbol{y_1 = \sqrt{\frac{57 - y_0}{3x_0}} = \sqrt{\frac{57 - (3.5)}{3(1.5)}} = 2.861}$$

Now the **results** are **more satisfactory**

$$\boldsymbol{x_2 = \sqrt{10 - x_1y_1} = \sqrt{10 - (2.179)(2.861)} = 1.941}$$

$$\boldsymbol{y_2 = \sqrt{\frac{57 - y_1}{3x_1}} = \sqrt{\frac{57 - (2.861)}{3(2.179)}} = 3.049}$$

Thus, the **method** is **converging on** the **true values** of ***x*** **= 2** and **y = 3.**

Thus, **even** where the **convergence** is **possible**, **divergence** can **occur** if the initial guesses are **insufficiently close** to the **true solution**. So, it **can be concluded** that the **sufficient conditions** for **convergence** for the **two-equation case** are:

$$\left|\frac{\partial u}{\partial x}\right| + \left|\frac{\partial u}{\partial y}\right| < 1 \quad \text{and} \quad \left|\frac{\partial v}{\partial x}\right| + \left|\frac{\partial v}{\partial y}\right| < 1$$

These **criteria** are **so restrictive** that **fixed-point iteration** has **limited utility** for **solving nonlinear systems**.

$$u(x, y) = x^2 + \ xy - 10$$

$$v(x, y) = y + 3xy^2 - 57$$

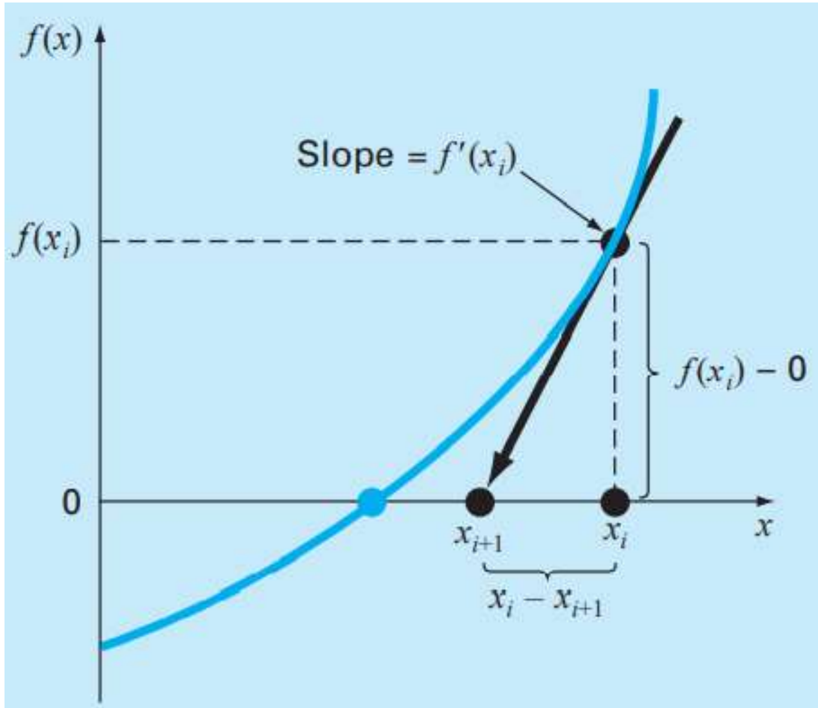
$$\Rightarrow \quad \left|\frac{\partial u}{\partial x}\right| + \left|\frac{\partial u}{\partial y}\right| = 2x + y + x = 3x + y = 3(1.5) + 3.5 = 8 > 1$$

$$\Rightarrow \quad \left|\frac{\partial v}{\partial x}\right| + \left|\frac{\partial v}{\partial y}\right| = 3y^2 + 1 + 6xy = 3(3.5)^2 + 1 + 6(1.5)(3.5)$$

$$= 69.25 \ > 1$$

➔ Newton-Raphson Method for Solving a Nonlinear System

Recall that the **Newton-Raphson method** was **predicated** on **employing** the **derivative** (that is, the **slope**) of a **function** to **estimate** its **intercept with** the **axis** of the **independent variable** (that is, the **root** as shown in figure.



The **estimate** was **based on** a **first-order Taylor series expansion**

$$\boldsymbol{f(x_{i+1}) = f(x_i) + (x_{i+1} - x_i)\, f'(x_i)}$$

where $\boldsymbol{x_i}$ is the **initial guess** at the **root** and $\boldsymbol{x_{i+1}}$ is the **point** at which the **slope intercepts** the $\boldsymbol{x}$ **axis**. At this **intercept**, $\boldsymbol{f(x_{i+1})}$ by definition **equals <u>zero</u>** and **equation** can be **rearranged** to **yield**

$$\boldsymbol{x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}}$$

which is the **single-equation form** of the **Newton-Raphson method**. The **multiequation form** is **derived** in an **identical manner**.

For the **two-variable case**, a **first-order Taylor series** can be **written** for **each nonlinear equation** as:

$$u_{i+1} = u_i + (x_{i+1} - x_i)\frac{\partial u_i}{\partial x} + (y_{i+1} - y_i)\frac{\partial u_i}{\partial y}$$

$$v_{i+1} = v_i + (x_{i+1} - x_i)\frac{\partial v_i}{\partial x} + (y_{i+1} - y_i)\frac{\partial v_i}{\partial y}$$

Just **as for** the **single-equation**, the **root estimate corresponds to** the **values** of x and **y**, where u_{i+1} and v_{i+1} **equal zero**. For this **situation**, **equations** can be **rearranged** to **give**

$$\frac{\partial u_i}{\partial x}x_{i+1} + \frac{\partial u_i}{\partial y}y_{i+1} = -u_i + x_i\frac{\partial u_i}{\partial x} + y_i\frac{\partial u_i}{\partial y}$$

$$\frac{\partial v_i}{\partial x}x_{i+1} + \frac{\partial v_i}{\partial y}y_{i+1} = -v_i + x_i\frac{\partial v_i}{\partial x} + y_i\frac{\partial v_i}{\partial y}$$

Because **all values subscripted** with ***i*'s** are **known** (they **correspond to** the **latest guess** or **approximation**), the **only unknowns** are x_{i+1} and y. Thus, **equations** are of **two linear equations with two unknowns.** Consequently, **algebraic manipulations** (for example, **Cramer's rule**) can be **employed** to **solve for**

$$x_{i+1} = x_i - \frac{u_i\frac{\partial v_i}{\partial y} - v_i\frac{\partial u_i}{\partial y}}{\frac{\partial u_i}{\partial x}\frac{\partial v_i}{\partial y} - \frac{\partial u_i}{\partial y}\frac{\partial v_i}{\partial x}}$$

$$y_{i+1} = y_i - \frac{v_i\frac{\partial u_i}{\partial x} - u_i\frac{\partial v_i}{\partial x}}{\frac{\partial u_i}{\partial x}\frac{\partial v_i}{\partial y} - \frac{\partial u_i}{\partial y}\frac{\partial v_i}{\partial x}}$$

The **denominator** of **each** of **these equations** is **formally referred** to as the **determinant** of the **Jacobian** of the **system**. **Equations** are the **two-equation** of the **Newton-Raphson method**. As in the **following example**, it can be **employed iteratively** to **home** in on the **roots** of **two simultaneous equations**.

<u>Example</u>: **Use** the **multiple-equation Newton-Raphson method** to **determine roots** of equations

$$x^2 + \ xy \ = 10$$
$$y + 3xy^2 = 57$$

Note that a **correct pair** of **roots** is ***x* = 2** and **y = 3**. **Initiate** the **computation** with **guesses** of ***x* = 1.5** and **y = 3.5**.

<u>Solution</u>:

First compute the **partial derivatives** and **evaluate them** at the **initial guesses** of $x_0 = 1.5$ and $y_0 = 3.5$:

$$\frac{\partial u_0}{\partial x} = 2x_0 + y_0 = 2(1.5) + 3.5 = 6.5$$

$$\frac{\partial u_0}{\partial y} = x_0 = 1.5$$

$$\frac{\partial v_0}{\partial x} = 2y_0^2 = 3(3.5) = 36.75$$

$$\frac{\partial v_0}{\partial y} = 1 + 6x_0 y_0 = 1 + 6(1.5)(3.5) = 32.5$$

Thus, the **determinant** of the **Jacobian** for the first **iteration** is:

$$\frac{\partial u_i}{\partial x}\frac{\partial v_i}{\partial y} - \frac{\partial u_i}{\partial y}\frac{\partial v_i}{\partial x} = 6.5(32.5) - 1.5(36.75) = 156.125$$

Evaluate the **functions** at the **initial guesses** as:

$$u_0 = x_0^2 + x_0 y_0 - 10 = (1.5)^2 + (1.5)(3.5) - 10 = -2.5$$

$$v_0 = y_0 + 3x_0 y_0^2 - 57 = (3.5) + 3(1.5)(3.5)^2 - 57 = 1.625$$

Substitute these **values** into **equations** of **Newton-Raphson method** to give

$$x_1 = x_0 - \frac{u_0 \frac{\partial v_0}{\partial y} - v_0 \frac{\partial u_0}{\partial y}}{\frac{\partial u_0}{\partial x}\frac{\partial v_0}{\partial y} - \frac{\partial u_0}{\partial y}\frac{\partial v_0}{\partial x}}$$

$$\Rightarrow \quad x_1 = 1.5 - \frac{(-2.5)(32.5) - (1.625)(1.5)}{156.125} = 2.036$$

$$y_1 = y_0 - \frac{v_0 \frac{\partial u_0}{\partial x} - u_0 \frac{\partial v_0}{\partial x}}{\frac{\partial u_0}{\partial x}\frac{\partial v_0}{\partial y} - \frac{\partial u_0}{\partial y}\frac{\partial v_0}{\partial x}}$$

$$\Rightarrow \quad y_1 = 3.5 - \frac{(1.625)(6.5) - (-2.5)(36.75)}{156.125} = 2.844$$

Thus, the results are converging to the true values of $x = 2$ and $y = 3$. The **computation** can be **repeated until** an **acceptable accuracy** is **obtained**.

Just **as with fixed-point iteration**, the **Newton-Raphson method** will **often diverge** if the **initial guesses** are not sufficiently close to the **true roots**. Whereas **graphical methods** could be **employed** to **derive good guesses** for the **single-equation case**, **no such simple procedure** is **available** for the **multiequation method**. **Although** there are some **advanced approaches** for **obtaining acceptable first estimates**, often the **initial guesses** must be **obtained** on the **basis** of **trial** and **error** and **knowledge** of the **physical system being modeled**. The **two-equation Newton-Raphson approach** can be **generalized** to **solve *n* simultaneous equations**.

Numerical Differentiation

In calculus the **derivative** of a function is **defined** as the **limit** of a **difference quotient**

$$f'(x) = \frac{df}{dx} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

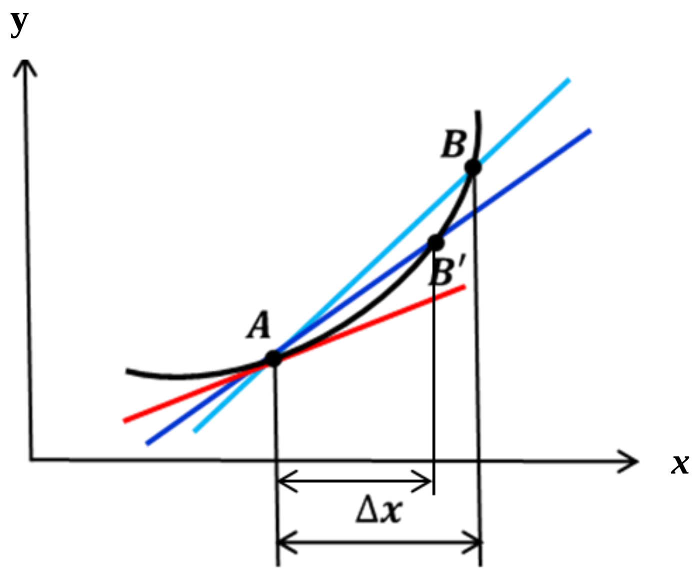
then **proceed** in **evaluating** a **few** of **these limits** as **examples**:

If $f(x) = 4x^3 + 3x^2 + 6x - 7$ then $f'(x) = \frac{df}{dx} = 12x^2 + 6x + 6$

If $f(x) = 5\cos(3x)$ then $f'(x) = \frac{df}{dx} = -15\sin(3x)$

If $f(x) = 7e^{2.1x}$ then $f'(x) = \frac{df}{dx} = 14.7e^{2.1x}$

In calculus, the **value** of the **derivative** for **some value** of the **independent variable**, say $x = x_0$, **gives** the **slope** of the **line** which is **tangent** to the **curve** represented by the function itself at the **point** $x = x_0$. This is **geometrically shown** in the **graph below**.



Notice that the **process** of the **limit** is also **shown** with Δx **representing** the **horizontal distance between** the **two points** ***A*** and ***B*** of the **secant line** ***AB***.

- The **light blue line** is a **secant** ***AB*** of the **curve** with a **large** Δx,
- The **darker blue line** is **another secant** ***AB***' with **smaller** Δx and
- The **red line** is the **tangent** (a **secant** with Δx **equal** to **zero**).

There are **many instances** when the **analytic calculation** of the **derivative** of a **given function** is **<u>not possible</u>**. This **may occur** when the **function under consideration** is **not one** of the **standard functions**. Trying **to evaluate** the **derivative** by **mimicking** the **process** of the **limit numerically**. **Use** the **secant** to **represent** the **tangent** (and **its slope produces** the **value** of the **derivative**). **Taking** progressively **closer points** (**smaller** Δx) **until** the **value** of the **slope stops changing significantly.**

It **can be seen** that as the **two intersection points get closer**, the **secant** will **look more** like a **tangent**.

How supposed to **move the points** so that they **get close together**?

Depending on the **answer** to **this question** there are **three different formulas** for the **numerical calculation** of **derivative**.

- If **fix** the **point *A*** and keep **bringing** the **point *B*** **closer** to the **left**, then **will obtain** the **forward derivative** at **point *A***.
- If **fix** the **point *B*** and **keep moving point *A*** **closer** then **will obtain** the **backward derivative** at **point *B***.
- If **fix** the **middle point** of the **interval between point *A*** and **point *B*** and **keep bringing both point *A*** and **B** **closer** to the **center**, then **will obtain** the **central derivative** at the **midpoint between point *A*** and **point *B***.

Taylor Series Review

If **expand function** f around x_i and f is $(n+1)$ – **times continuously differentiable** on an **open interval** containing x_i, **Taylor's theorem** with the **remainder term** says that if x_{i+1} is **another point** in this interval, then:

$$f(x_{i+1}) \cong f(x_i) + f'(x_i)\frac{\Delta x}{1!} + f''(x_i)\frac{\Delta x^2}{2!} + f'''(x_0)\frac{\Delta x^3}{3!} + \cdots + f^{(n)}(x_i)\frac{\Delta x^n}{n!} + R_n$$

Where:

$$
\begin{aligned}
\mathbf{1}! &= \mathbf{1} \times \mathbf{1} = \mathbf{1} \\
\mathbf{2}! &= \mathbf{1} \times \mathbf{2} = \mathbf{2} \\
\mathbf{3}! &= \mathbf{1} \times \mathbf{2} \times \mathbf{3} = \mathbf{6} \\
&\vdots \\
\boldsymbol{n}! &= \mathbf{1} \times \mathbf{2} \times \mathbf{3} \times \cdots \times \boldsymbol{n}
\end{aligned}
$$

and

$$\boldsymbol{R_n} = \boldsymbol{f}^{(\boldsymbol{n}+\mathbf{1})}(\boldsymbol{\xi})\,\frac{(\Delta \boldsymbol{x})^{\boldsymbol{n}+\mathbf{1}}}{(\boldsymbol{n}+\mathbf{1})\,!}$$

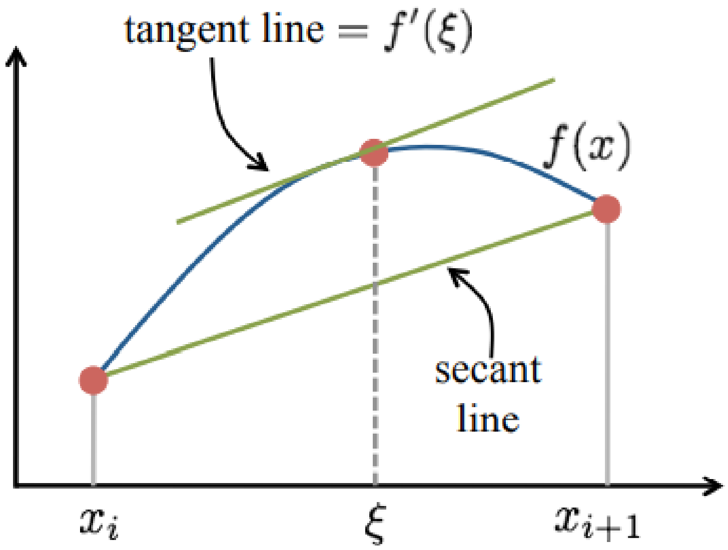
$\boldsymbol{\xi}$: is a **number** in the **open interval between** $\boldsymbol{x_i}$ and $\boldsymbol{x_{i+1}}$.

$\Delta \boldsymbol{x} = \boldsymbol{h} = \boldsymbol{x_{i+1}} - \boldsymbol{x_i}$

$\boldsymbol{f}(\boldsymbol{x_{i+1}}) = \boldsymbol{f}(\boldsymbol{x_i} + \Delta \boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x_i} + \boldsymbol{h})$

Mean Value Theorem

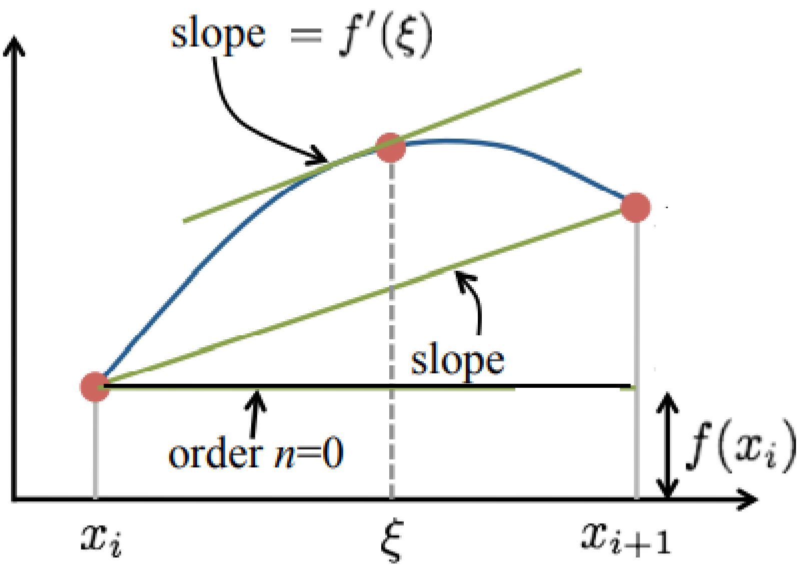
The **appearance** of $\boldsymbol{\xi}$, a **point between** $\boldsymbol{x_i}$ and $\boldsymbol{x_{i+1}}$, suggests that **there** is a **connection between** this **result** and **the Mean Value Theorem**, which **states** that given a **planar arc between two endpoints**, there is **at least one point** at which the **tangent** to the **arc** is **parallel** to the **secant** through **its endpoints(see figure below)**.



If **a function** $\boldsymbol{f}$ is **continuous** on $[\boldsymbol{x_i}, \boldsymbol{x_{i+1}}]$ and **differentiable** on $(\boldsymbol{x_i}, \boldsymbol{x_{i+1}})$, then there **exists a point** $\boldsymbol{\xi}$ such that:

$$\boldsymbol{f}'(\boldsymbol{\xi}) = \frac{\boldsymbol{f}(\boldsymbol{x_{i+1}}) - \boldsymbol{f}(\boldsymbol{x_i})}{\boldsymbol{x_{i+1}} - \boldsymbol{x_i}}$$

Mean Value Theorem and Taylor's Theorem



Back to the **Taylor series**, for $\boldsymbol{n} = \mathbf{0}$:

$$f(x_{i+1}) \cong f(x_i) + R_0$$

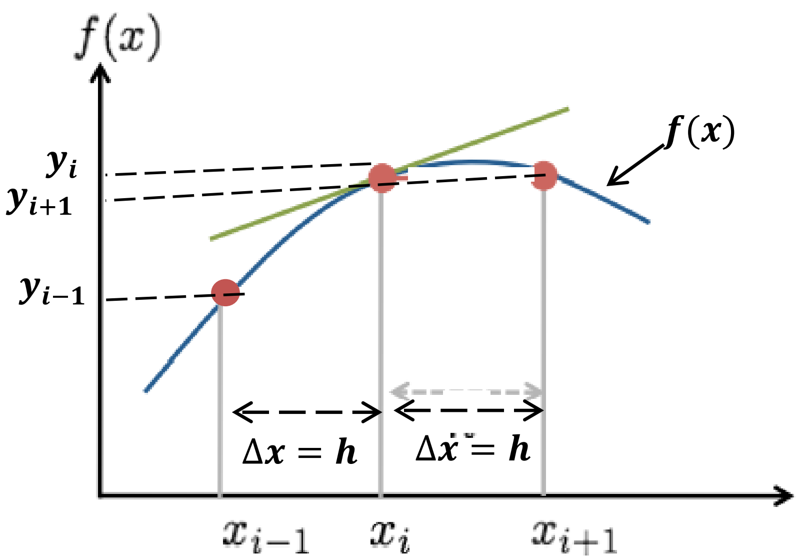
where: $R_0 = f'(\xi)\Delta x$ and $\Delta x = h = x_{i+1} - x_i$

Then $$f(x_{i+1}) \cong f(x_i) + f'(\xi)\,(x_{i+1} - x_i)$$

Where ξ is **between** x_i and x_{i+1}. This is the **Mean Value Theorem**, which is **used** to **prove Taylor's theorem**. Also **can regard a Taylor expansion** as an **extension** of the **Mean Value Theorem**.

Approximation of First Order Derivative

Let choose **three points** on the **curve** of the **function** $f(x)$ as shown in the figure



point i $: x_i \rightarrow f_i = f(x_i)$

point $i+1$ $: x_{i+1} \rightarrow f_{i+1} = f(x_{i+1})$

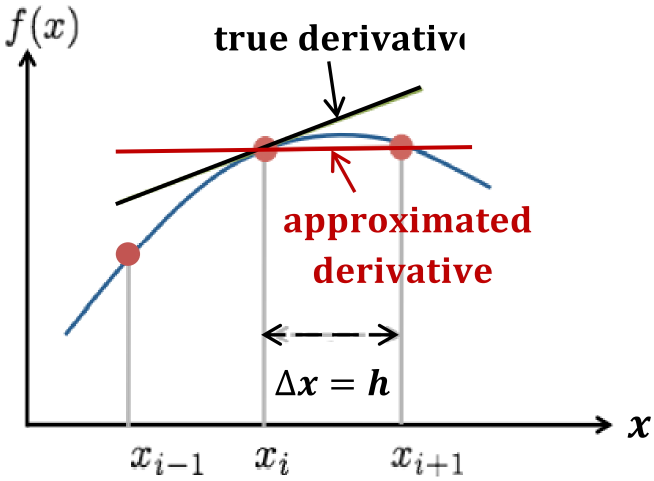
point $i-1$ $: x_{i-1} \rightarrow f_{i-1} = f(x_{i-1})$

1. By Forward Difference

Truncating the **Taylor series after** the **first derivative** yields

$$f(x_{i+1}) = f(x_i) + f'(x_i)\Delta x$$

Rearranging equation gives

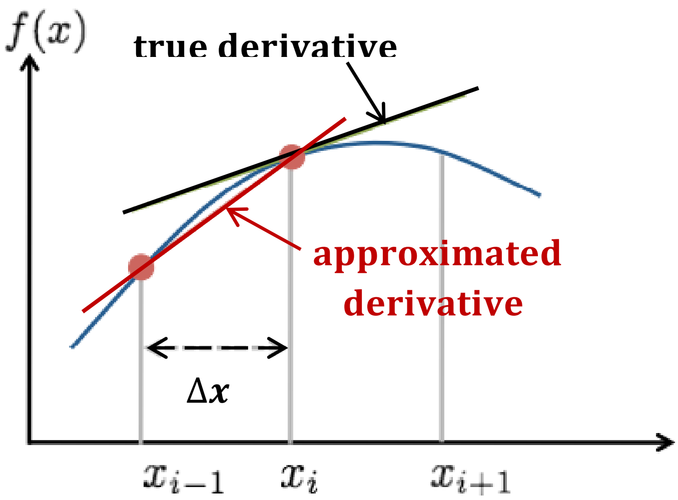


$$f'(x_i) = \frac{f(x_{i+1}) - f(x_i)}{\Delta x}$$

Forward difference is the **slope** of the **line** that **connects points** $[x_i, f(x_i)]$ and $[x_{i+1}, f(x_{i+1})]$

2. By Backward Difference

The **Taylor series** can be **expanded backward** to **calculate a previous value** on the **basis** of a **present value**.



$\Delta x = x_i - x_{i-1}$

$\Rightarrow x_{i-1} = x_i - \Delta x$

$f(x_{i-1}) = f(x_i - \Delta x)$

$f(x_{i-1}) \cong f(x_i) - f'(x_i)\Delta x +$

Truncating the **expansion** afte

$$f(x_{i-1}) \cong f(x_i) - f'(x_i)\Delta x$$

Rearranging the **equation** gives

$$f'(x_i) \cong \frac{f(x_i) - f(x_{i-1})}{\Delta x}$$

Backward difference is the **slope** of the **line** that **connects points** $[x_{i-1}, f(x_{i-1})]$ and $[x_i, f(x_i)]$

3. By Central Difference

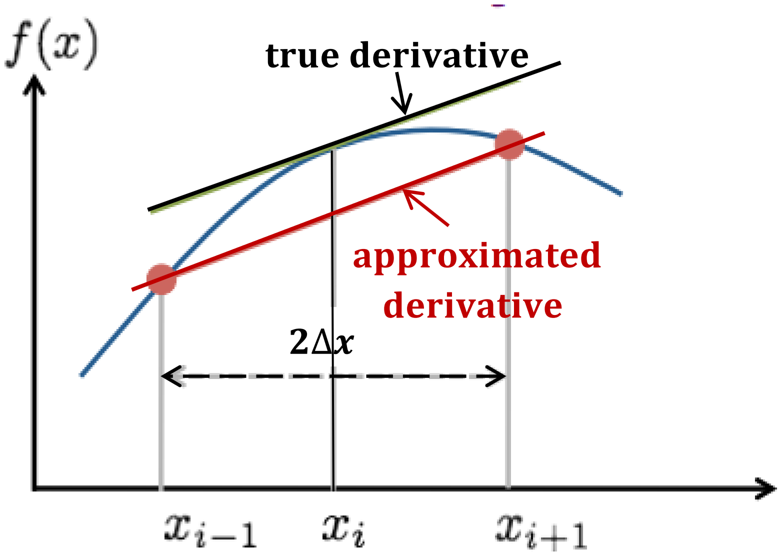
The **third way** to **approximate** the **first derivative** is **by** **subtracting backward difference** from **forward difference**

$$f(x_{i+1}) \cong f(x_i) + f'(x_i)\Delta x + f''(x_i)\frac{\Delta x^2}{2} + f'''(x_i)\frac{\Delta x^3}{6} + \cdots$$

$$-$$

$$f(x_{i-1}) \cong f(x_i) - f'(x_i)\Delta x + f''(x_i)\frac{\Delta x^2}{2} - f'''(x_i)\frac{\Delta x^3}{6} + \cdots$$

$$f(x_{i+1}) - f(x_{i-1}) = 2f'(x_i)\Delta x + 2f'''(x_i)\frac{\Delta x^3}{6} + \cdots$$



Truncating the **expansion after** the **first derivative** yields

$$f(x_{i+1}) - f(x_{i-1}) \cong 2f'(x_i)\Delta x$$

Rearranging the **equation** gives

$$f'(x_i) \cong \frac{f(x_{i+1}) - f(x_{i-1})}{2\Delta x}$$

Central difference is the **slope** of the **line** that **connects points** $[x_{i-1}, f(x_{i-1})]$ and $[x_{i+1}, f(x_{i+1})]$

The **following table** lists **difference formulas**, of **various accuracy**, that **can be used** for **approximating first derivatives**:

Method	Formula
Two-point forward difference	$f'(x_i) = \frac{1}{\Delta x}[f(x_{i+1}) - f(x_i)]$
Three-point forward difference	$f'(x) = \frac{1}{2\,\Delta x}[-3f(x_i) + 4f(x_{i+1}) - f(x_{i+2})]$
Two-point backward difference	$f'(x_i) = \frac{1}{\Delta x}[f(x_i) - f(x_{i-1})]$
Three-point backward difference	$f'(x) = \frac{1}{2\,\Delta x}[f(x_{i-2}) - 4f(x_{i-1}) + 3f(x_i)]$
Two-point central difference	$f'(x_i) = \frac{1}{2\Delta x}[f(x_{i+1}) - f(x_{i-1})]$
Four-point central difference	$f'(x) = \frac{f(x_{i-2}) - 8f(x_{i-1}) + 8f(x_{i+1}) - f(x_{i+2})}{12\,\Delta x}$

All these **three formulas give approximations** to the **derivative** at the **point x_i**. It **can be shown** that the **third formula** is **more accurate** than the others as follows:

Example 9: Approximate the **derivative** of the **function**

$$f(x) = x^2 + 2x \text{ at } x = 3$$

using the **Forward, Backward,** and **Central Difference method** and **step size 1.**

Solution:

$x_i = 3$ and $\Delta x = 1$, so

$x_{i+1} = x_i + \Delta x = 3 + 1 = 4$

$x_{i-1} = x_i - \Delta x = 3 - 1 = 2$

The values of the corresponding functions are:

$f(x_{i-1}) = f(2) = 2^2 + 2(2) = 8$

$f(x_i) \quad = f(3) = 3^2 + 2(3) = 15$

$f(x_{i+1}) = f(4) = 4^2 + 2(4) = 24$

FDA: $f'(x_i) \approx \frac{f(x_{i+1}) - f(x_i)}{\Delta x}$

$$f'(3) \approx \frac{f(4) - f(3)}{1} = \frac{24 - 15}{1} = 9$$

$$BDA\text{: } f'(x_i) \approx \frac{f(x_i) - f(x_{i-1})}{\Delta x}$$

$$f'(3) \approx \frac{f(3) - f(2)}{1} = \frac{15 - 8}{1} = 7$$

$$CDA\text{: } f'(x_i) \approx \frac{f(x_{i+1}) - f(x_{i-1})}{2\Delta x}$$

$$f'(3) \approx \frac{f(4) - f(2)}{2 \times 1} = \frac{24 - 8}{2} = 8$$

The **actual value** of the **function** is:

$$f(x) = x^2 + 2x \implies f'(x) = 2x + 2 \implies f'(3) = 2 \times 3 + 2 = 8$$

<u>Example</u> 10: Consider the **table given** below which **lists the values** of a **function** $f(x)$ for **various values** of x **ranging** from $x = -1.0$ to $x = +2.0$ at **uniform increments** of **0.25**.

x	$f(x)$
-1.0	-0.841
-0.75	-0.682
-0.50	-0.479
-0.25	-0.247
-0.00	-0.00
$+0.25$	$+0.247$
$+0.50$	$+0.479$ (x_{i-1})
$+0.75$	$+0.682$ (x_i)
$+1.00$	$+0.841$ (x_{i+1})
$+1.25$	$+0.949$
$+1.50$	$+0.997$
$+1.75$	$+0.984$
$+2.00$	$+0.909$

Use F<u>orward</u>, <u>Backward</u>, and **<u>Central Difference</u> <u>method</u>** to **find** the **approximate derivative** at **point** ***x* = 0.5**.

<u>Solution</u>:

- The **<u>Forward Difference Approximation (FDA)</u>** at **point** $x_i = 0.5$ is:

$$f'(x) = \frac{f(x_{i+1}) - f(x_i)}{\Delta x} = \frac{0.682 - 0.479}{0.25} = 0.812$$

$$f'(x) = \frac{f(x + \Delta x) - f(x)}{\Delta x} = \frac{0.682 - 0.479}{0.25} = 0.812$$

- The **<u>Backward Difference Approximation (BDA)</u>** at **point** $\boldsymbol{x_i = 0.5}$ is:

$$f'(x) = \frac{f(x_i) - f(x_{i-1})}{\Delta x} = \frac{0.479 - 0.247}{0.25} = 0.928$$

$$f'(x) = \frac{f(x) - f(x - \Delta x)}{\Delta x} = \frac{0.479 - 0.247}{0.25} = 0.928$$

- The **<u>Central Difference Approximation (CDA)</u>** at **point** $\boldsymbol{x_i = 0.5}$ is:

$$f'(x) = \frac{f(x_{i+1}) - f(x_{i-1})}{2\Delta x} = \frac{0.682 - 0.247}{2 \times 0.25} = 0.870$$

$$f'(x) = \frac{f(x + \Delta x) - f(x - \Delta x)}{2\Delta x} = \frac{0.682 - 0.247}{2 \times 0.25} = 0.870$$

➢ **Notice** that **all three** are **different**. Also, the **<u>central difference</u>** is the **<u>average</u>** of the **<u>other two</u>**.

$$\frac{0.812 + 0.928}{2} = 0.87$$

This is **<u>NOT a coincidence</u>**.

➢ **Also notice** that at the **points** $\boldsymbol{x = -1.0}$ and $\boldsymbol{x = +2.0}$ (the **end points** of **interval**) **cannot be calculated more than** a **single approximation**.

If the **calculation** of the **derivative** at **all points** of the **table**, then a **spreadsheet** is **very appropriate** for this type of problem. **Enter** the table to a **sheet** and then can use the **formulas** to **generate three more columns**, which will **contain** the **approximations** to the **derivative**.

	x	$f(x)$	Forward	Backward	Central
	-1.0	-0.841	0.636		
	-0.75	-0.682	0.812	0.636	0.724
	-0.50	-0.479	0.928	0.812	0.870
	-0.25	-0.247	0.988	0.928	0.958
	-0.00	-0.00	0.988	0.988	0.988
x_{i-1}	$+0.25$	$+0.247$	0.928	0.988	0.958
x_i	$+0.50$	$+0.479$	0.812	0.928	0.870
x_{i+1}	$+0.75$	$+0.682$	0.636	0.812	0.724
	$+1.00$	$+0.841$	0.432	0.636	0.534
	$+1.25$	$+0.949$	0.192	0.432	0.312
	$+1.50$	$+0.997$	-0.052	0.192	0.070
	$+1.75$	$+0.984$	-0.300	-0.052	-0.176
	$+2.00$	$+0.909$		-0.300	

<u>Example</u> 10: Using Three – Point Forward Difference, Backward Difference, and Central Difference Formula to find solution $f'(1.10)$.

x	1	1.05	1.10	1.15	1.20	1.25	1.30
$f(x)$	1	1.02470	1.04881	1.07238	1.09545	1.11803	1.14018

<u>Solution</u>:

<u>Three – Point Forward Difference Formula</u>:

$$f'(x) = \frac{1}{2\,\Delta x}[-3f(x_i) + 4f(x_{i+1}) - f(x_{i+2})]$$

From the **table** it is **evident** that the **values ranging** from 1.0 to $x = 1.30$ at **uniform increments** of 0.05. **Thus,**

$$\Delta x = 0.05$$

Then,

$$f'(1.10) = \frac{1}{2(0.05)}[-3f(1.10) + 4f(1.15) - f\{1,10 + 2(0.05)\}]$$

$$f'(1.10) = \frac{1}{0.10}[-3f(1.10) + 4f(1.15) - f(1,20)]$$

$$f'(1.10) = \frac{1}{0.10}[-3(1.04881) + 4(1.07238) - (1,09545)]$$

$$\Rightarrow \quad f'(1.10) = 0.4764$$

<u>Three – Point Backward Difference Formula</u>:

$$f'(x) = \frac{1}{2\,\Delta x}[f(x_{i-2}) - 4f(x_{i-1}) + 3f(x_i)]$$

$$f'(1.10) = \frac{1}{2\,(0.05)}[f\{1.10 - 2(0.05)\} - 4f(1.10 - 0.05) + 3f(1.10)]$$

$$f'(1.10) = \frac{1}{0.10}[f(1) - 4f(1.05) + 3f(1.10)]$$

$$f'(1.10) = \frac{1}{0.10}[1 - 4(1.0247) + 3f(1.04881)] = 0.4763$$

<u>Three – Point Central Difference Formula</u>:

$$f'(x) = \frac{1}{2\,\Delta x}[f(x_{i+1}) - f(x_{i-1})]$$

$$f'(x) = \frac{1}{2\,(0.05)}[f(1.10 + 0.05) - f(1.10 - 0.05)]$$

$$f'(x) = \frac{1}{0.10}[f(1.15) - f(1.05)] = \frac{1}{0.10}[1.07238 - 1.0247]$$

$$\Rightarrow \quad f'(1.10) = 0.4768$$

<u>Example</u> 12: Use Three – Point Forward Difference, Backward Difference, and Central Difference Formula to find solution of $f'(2.5)$ if **$h = 0.5$ of the function**

$$f(x) = 2x^3 + x^2 - 4$$

Also **find the exact value** of $f'(x)$ and **absolute error** for **each estimation.**

<u>Solution</u>:

$$f(x) = 2x^3 + x^2 - 4 \quad \Rightarrow \quad f'(x) = 6x^2 + 2x$$

Exact value of $f'(2.5)$is:

$$f'(2.5) = 6(2.5)^2 + 2(2.5) = 42.5$$

<u>Three – Point Forward Difference Formula</u>:

$$f'(x) = \frac{1}{2\,\Delta x}[-3f(x_i) + 4f(x_{i+1}) - f(x_{i+2})]$$

<u>Three – Point Backward Difference Formula</u>:

$$f'(x) = \frac{1}{2\,\Delta x}[f(x_{i-2}) - 4f(x_{i-1}) + 3f(x_i)]$$

<u>Three – Point Central Difference Formula</u>:

$$f'(x) = \frac{1}{2\,\Delta x}[f(x_{i+1}) - f(x_{i-1})]$$

From the formula for each method it is clear to find $f'(2.5)$, the points that needed are:

$x_i = 2.5$

$x_{i+1} = 2.5 + 0.5 = 3$

$x_{i+2} = 2.5 + 2(0.5) = 3.5$

$x_{i-1} = 2.5 - 0.5 = 2$

$x_{i-2} = 2.5 - 2(0.5) = 1.5$

Thus, it is needed to find:

$f(x_i) = f(2.5) = 2(2.5)^3 + (2.5)^2 - 4 = 33.5$

$f(x_{i+1}) = f(3) = 2(3)^3 + (3)^2 - 4 = 59$

$f(x_{i+2}) = f(3.5) = 2(3.5)^3 + (3.5)^2 - 4 = 94$

$f(x_{i-1}) = f(2) = 2(2)^3 + (2)^2 - 4 = 16$

$f(x_{i-2}) = f(1.5) = 2(1.5)^3 + (1.5)^2 - 4 = 5$

x	1.5	2	2.5	3	3.5
$f(x)$	5	16	33.5	59	94

Three – Point Forward Difference Formula:

$$f'(x) = \frac{1}{2\,\Delta x}[-3f(x_i) + 4f(x_{i+1}) - f(x_{i+2})]$$

$$f'(2.5) = \frac{1}{2\,(0.5)}[-3(33.5) + 4(59) - (94)] = 41.5$$

Absolute error

$$\varepsilon = |\text{Exact value} - \text{Approximated value}|$$

$$\varepsilon = |42.5 - 41.5| = 1$$

Three – Point Backward Difference Formula:

$$f'(x) = \frac{1}{2\,\Delta x}[f(x_{i-2}) - 4f(x_{i-1}) + 3f(x_i)]$$

$$f'(2.5) = \frac{1}{2\,(0.5)}[5 - 4(16) + 3(33.5)] = 42.5$$

Absolute error

$$\varepsilon = |42.5 - 42.5| = 0$$

Three – Point Central Difference Formula:

$$f'(x) = \frac{1}{2\,\Delta x}[f(x_{i+1}) - f(x_{i-1})]$$

$$f'(2.5) = \frac{1}{2\,(0.5)}[59 - 16] = 43$$

Absolute error:

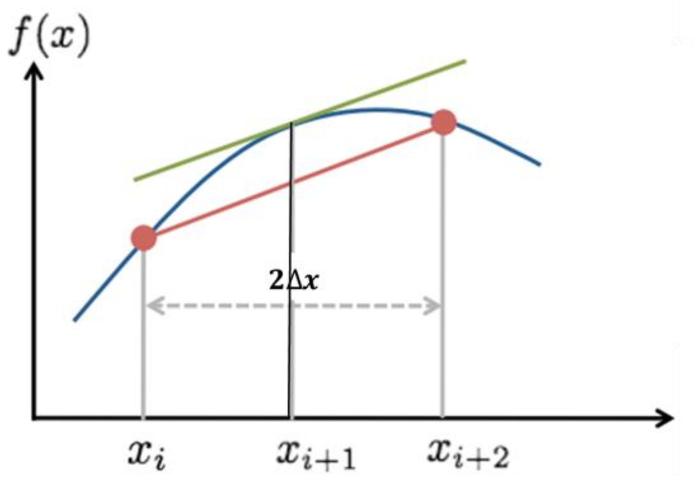
$$\varepsilon = [42.5 - 43] = 0.5$$

The **same approach used** to **develop finite difference formulas** for the first derivative **can be used** to **develop expressions** for **higher-order derivatives.**

Approximation of Second Order Derivative

1. By Forward Difference

To **approximate 2nd order derivatives**, write a **forward Taylor series expansion** for $f(x_{i+2})$ **in terms** of $f(x_i)$:



$$f(x_{i+2}) \cong f(x_i) + f'(x)(2\Delta x) + f''(x)\frac{(2\Delta x)^2}{2} + f'''(x)\frac{(2\Delta x)^3}{6} + \cdots$$

Truncating the **expansion after** the **second derivative** gives

$$f(x_{i+2}) \cong f(x_i) + f'(x)(2\Delta x) + f''(x)\frac{(2\Delta x)^2}{2}$$

Rearranging yields

$$f(x_{i+2}) \cong f(x_i) + 2f'(x)(\Delta x) + 2\, f''(x)(\Delta x)^2$$

Truncating the **forward difference after** the **second derivative** and **multiplying** by **2** yields

$$2\, f(x_{i+1}) \cong 2\, f(x_i) + 2\, f'(x_i)\Delta x + \, f''(x)(\Delta x)^2$$

Subtracting from a**bove equation** yields

$$f(x_{i+2}) - 2\, f(x_{i+1}) \cong -f(x_i) + f''(x_i)(\Delta x)^2$$

Rearranging the **equation** yields

$$f''(x_i) \cong \frac{f(x_i) - 2\, f(x_{i+1}) + f(x_{i+2})}{(\Delta x)^2}$$

2. By Backward Difference

The **same manipulations** can be **employed to derive** a **2nd order backward difference**:

$$f''(x_i) \cong \frac{f(x_i) - 2\, f(x_{i-1}) + f(x_{i-2})}{(\Delta x)^2}$$

3. <u>By Central Difference</u>

$$f''(x_i) \cong \frac{f(x_{i+1}) - 2\,f(x_i\,) + f(x_{i-1})}{(\Delta x)^2}$$

As was the case with the **first derivative approximations**, the **central difference** is **more accurate**.

Approximations of **higher derivatives** $f'''(x)$, $f^{(4)}(x)$ etc. **can** be **obtained** in a **similar manner**. The **following table** lists **difference formulas**, of **various accuracy**, that **can be used** for **numerical evaluation** of **second derivatives**. The **formulas** can be **used** when the **function** that is **being differentiated** is **specified** as a **set** of **discrete points** with the **<u>independent variable equally spaced</u>**.

Method	Formula
Three-point forward difference	$f''(x_i) \cong \frac{f(x_i) - 2\,f(x_{i+1}\,) + f(x_{i+2})}{(\Delta x)^2}$
Four-point forward difference	$f''(x_i) \cong \frac{2\,f(x_i) - 5\,f(x_{i+1}\,) + 4f(x_{i+2}) - f(x_{i+3}\,)}{(\Delta x)^2}$
Three-point backward difference	$f''(x_i) \cong \frac{f(x_{i-2}) - 2\,f(x_{i-1}\,) + f(x_i)}{(\Delta x)^2}$
Four-point backward difference	$f''(x_i) \cong \frac{-\,f(x_{i-3}) + 4\,f(x_{i-2}\,) - 5f(x_{i-1}) + 2f(x_i\,)}{(\Delta x)^2}$
Three-point central difference	$f''(x_i) \cong \frac{f(x_{i+1}) - 2\,f(x_i\,) + f(x_{i-1})}{(\Delta x)^2}$
Five-point central difference	$f''(x_i) \cong \frac{-\,f(x_{i-2}) + 16\,f(x_{i-1}\,) - 30f(x_i) + 16f(x_{i+1}) - f(x_{i+2})}{12(\Delta x)^2}$

<u>Example</u> 13: Consider the **function** $f(x) = \frac{2^x}{x}$. **Calculate** the **second derivative** at $x = 2$ **numerically** with the **three-point central difference formula** using **points** $x = 1.8$, $x = 2$, and $x = 2.2$. Then **compare** the **results** with the **exact (analytical) derivative.**

<u>Solution</u>:

- **<u>Analytical differentiation</u>**:

  The **second derivative** of the function $f(x) = \frac{2^x}{x}$ is:

$$f'(x) = \frac{2^x \times 1 - x\,(ln2)\,2^x}{x^2} = \frac{2^x}{x^2} - ln\,2\,\frac{2^x}{x}$$

$$f''(x) = \frac{(2x) \times\ 2^x - x^2\,(ln2) 2^x}{x^4} - ln\,2\,\left(\frac{2^x}{x^2} - ln2\,\frac{2^x}{x}\right)$$

$$\Longrightarrow f''(x) = (ln\,2)^2\,\frac{2^x}{x} - 2\,(ln\,2)\,\frac{2^x}{x^2} + 2\,\frac{2^x}{x^3}$$

and the **value** of the **second derivative** at $x = 2$ is:

$$\Longrightarrow f''(2) = 0.57462$$

- **Numerical differentiation**:

The **numerical differentiation** can be **done by** using the **three-point central difference formula**

$$f''(x_i) \cong \frac{f(x_{i-1}) - 2\,f(x_i\,) + f(x_{i+1})}{(\Delta x)^2}$$

where

$x_{i-1} = x = 1.8$,
$x_i \quad = x = 2$, and
$x_{i+1} = x = 2.2$
$\Delta x = x_{i+1} - x_i \ = 2.2 - 2 = 0.2$

and the **values** of the **corresponding functions** are:

$$f(x_{i-1}) = f(1.8) = \frac{2^{1.8}}{1.8} = 1.9346$$

$$f(x_i) \quad = f(2) \quad = \frac{2^2}{2} = 2$$

$$f(x_{i+1}) = f(2.2) = \frac{2^{2.2}}{2.2} = 2.0885$$

substituting the **values** of the function at **points** in the **formula** gives

$$f''(x_i) \cong \frac{1.93456 - 2\,(2\,) + 2.08854}{(0.2)^2} = 0.5775$$

$$\boldsymbol{E_s = 0.5775 - 0.5746 = 0.0029}$$

$$\boldsymbol{E_a = \frac{0.5775 - 0.5746}{0.57462} \times 100 = 0.5\%}$$

The **result shows** that the **three-point central difference formula** gives a **quite accurate approximation** for the **value** of the **second derivative**.

Runge – Kutta Methods

Runge – Kutta method is **an effective** and **widely used method** for **solving** the **initial-value problems** of **differential equations**. **Runge – Kutta method** can be **used** to **construct high order accurate numerical method** by **functions' self without needing** the **high order derivatives** of **functions.**

In **numerical analysis**, the **Runge–Kutta methods** (English: rʊŋəˈkʊtɑ) are a family of **implicit** and **explicit iterative methods**, which **include** the **Euler method**, used in **temporal discretization** for the **approximate solutions** of **simultaneous nonlinear equations**. These **methods** were **developed around 1900** by the **German mathematicians Carl Runge** and **Wilhelm Kutta**.

Euler's Method

If an initial value problem

$$\boldsymbol{y' = \frac{dy}{dt} = f(x, y). \quad y(x_0) = y_0} \qquad (\mathbf{1})$$

cannot be solved **analytically**, it is necessary to resort to numerical methods to obtain useful approximations to a **solution of Equation (1)**.

If interesting in computing approximate values of the solution of the **Equation (1)** at **equally spaced points** $\boldsymbol{x_0, x_1, x_2, \dots, x_n = b}$ in an interval $[\boldsymbol{x_0, b}\ ]$. Thus,

$$\boldsymbol{x_i = x_0 + ih \ , \quad i = 0, 1, 2, \dots, n}$$

where $$\boldsymbol{h = \frac{b - x_0}{n}}$$

Denoting the approximate values of the solution at these points by $\boldsymbol{y_0, y_1, y_2, \dots, y_n}$; thus, $\boldsymbol{y_i}$ is an approximation to $\boldsymbol{y(x_i)}$. calling

$$\boldsymbol{e_i = y(x_i) - y_i}$$

the **error at the ith step**. Because of the **initial condition**

$$\boldsymbol{y(x_0) = y_0}$$

always will have $\boldsymbol{e_0 = 0}$.

However, in general

$$\boldsymbol{e_i \neq 0} \quad \boldsymbol{i} \text{ if } \boldsymbol{i > 0}$$

Encountering **two sources** of **error** in applying a **numerical method** to solve an **initial value problem**:

The **formulas** defining the method are **based on** some sort of **approximation**. **Errors** due to the **inaccuracy** of the **approximation** are called **truncation errors**.

Computers do **arithmetic** with a **fixed number** of **digits**, and therefore **make errors** in **evaluating** the **formulas defining** the **numerical methods**. **Errors** due to the **computer's inability** to do **exact arithmetic** are called **roundoff errors**.

1. Euler method (1st order derivative)

The **simplest numerical method** for solving **Equation (1)** is **Euler's method**. This method is **so crude** that it is **seldom used** in **practice**; however, its simplicity makes it **useful** for **illustrative purposes**. **Euler's method** is **based on** the **assumption** that the **tangent line** to the **integral curve** of **Equation (1)** at $[\boldsymbol{x_i}, \boldsymbol{y(x_i)}]$ **approximates** the **integral curve** over the **interval** $[\boldsymbol{x_i}, \boldsymbol{x_{i+1}}]$.

Since the **slope** of the **integral curve** of **Equation (1)** at $[\boldsymbol{x_i}, \boldsymbol{y(x_i)}]$ is:

$$\boldsymbol{y'(x_i) = f[x_i, y(x_i)]}$$

the **equation** of the **tangent line** to the **integral curve** at $[\boldsymbol{x_i}, \boldsymbol{y(x_i)}]$ is:

$$\boldsymbol{y = y(x_i) + f[x_i, y(x_i)](x - x_i)} \qquad (\mathbf{2})$$

Setting $\boldsymbol{x = x_{i+1} = x_i + h}$, yields

$$\boldsymbol{y_{i+1} = y(x_i) + hf[x_i, y(x_i)]} \qquad (3)$$

as an approximation to $\boldsymbol{y(x_{i+1})}$.

Since $\boldsymbol{y(x_0) = y_0}$ is known, Equation (3) can be used with $\boldsymbol{i = 0}$ to compute

$$\boldsymbol{y_1 = y_0 + hf(x_0, y_0)}$$

However, setting $\boldsymbol{i = 1}$ in Equation (3) yields

$$\boldsymbol{y_2 = y_1 + hf(x_1, y_1)}$$

Having computed $\boldsymbol{y_2}$, can compute

$$\boldsymbol{y_3 = y_2 + hf(x_2, y_2)}$$

In general, Euler's method starts with the known value $\boldsymbol{y(x_0) = y_0}$ and computes $\boldsymbol{y_1, y_2, y_3, \dots, y_n}$ successively by with the formula

$$\boldsymbol{y_{i+1} = y_i + hf(x_i, y_i) \qquad 0 \le i \le n - 1}$$

The next example illustrates the computational procedure indicated in Euler's method.

<u>Example</u> 1: Find **y (0.2)** for

$$\boldsymbol{y' = \frac{x - y}{2} \ , \ \ y(0) = 1}$$

with **step length 0.1** using **Euler method.**

<u>Solution</u>:

Given: $y' = f(x,y) = \frac{x-y}{2}$, $y(0) = 1$, $h = 0.1$, $y(0.2) = ?$

$$y_1 = y_0 + hf(x_0, y_0) = 1 + 0.1f(0,1) = 1 + 0.1\left(\frac{0-1}{2}\right)$$

$$\Longrightarrow \quad y_1 = 1 - 0.05 = 0.95$$

$$y_2 = y_1 + hf(x_1, y_1) = 0.95 + 0.1f(0.1, 0.95)$$

$$= 0.95 + 0.1\left(\frac{0.1 - 0.95}{2}\right) = 0.95 + 0.1 \cdot (-0.425)$$

$$\Longrightarrow \quad y_2 = 0.95 - 0.0425 = 0.9075$$

$$\therefore \quad y(0,2) = 0.9075$$

<u>Example</u> 2: Use **Euler's method** with $h = 0.1$ to **find approximate values** for the **solution** of the **initial value problem**

$$y' + 2y = x^3e^{-2x} \;,\; y(0) = 1 \quad \text{at} \quad x = 0.1, 0.2, 0.3$$

<u>Solution</u>:

Rewriting equation as

$$y' = -2y + x^3e^{-2x} \;,\; y(0) = 1$$

which is the form

$$f(x,y) = -2y + x^3e^{-2x} \;,\; x_0 = 0 \text{ and } y_0 = 1$$

Euler method

$$y_1 = y_0 + hf(x_0, y_0) = 1 + (0.1)f(0,1) = 1 + (0.1)(-2) = 0.8$$

$$y_2 = y_1 + hf(x_1, y_1) = 0.8 + (0.1)f(0.1, 0.8)$$

$$= 0.8 + (0.1)\left[-2(0.8) + (0.1)^3e^{-2(0.1)}\right] = 0.6401$$

$$y_3 = y_2 + hf(x_2, y_2) = 0.6401 + (0.1)f(0.1, 0.6401)$$

$$= 0.6401 + (0.1)\left[-2(0.6401) + (0.2)^3e^{-2(0.2)}\right] = 0.5126$$

$$y_4 = y_3 + hf(x_3, y_3) = 0.5126 + (0.1)f(0.3, 0.5126)$$

$$= 0.5126 + (0.1)\left[-2(0.5126) + (0.3)^3e^{-2(0.3)}\right] = 0.46496$$

2. Euler method (2nd order derivative)

$$\boldsymbol{y_1 = y_0 + hf(x_0 + y_0 + z_0)}$$

<u>Example</u> 1: Find $\boldsymbol{y(0.1)}$ for

$$\boldsymbol{y'' = 1 + 2xy - x^2z}, \quad \boldsymbol{x_0 = 0}, \quad \boldsymbol{y_0 = 1}, \quad \boldsymbol{y'_0 = 0}, \quad \boldsymbol{z_0 = 0}$$

with **step size 0.1** using **Euler method** (**2nd order derivative**)

<u>Solution</u>:

Put $\frac{dy}{dx} = z$ and **differentiate** with respect to x, yields

$$\frac{dz}{dx} = \frac{d^2y}{dx^2}$$

Thus, the **system** of **equations** is:

$$\frac{dy}{dx} = z = f(x, y, z)$$

$$\frac{dz}{dx} = 1 + 2xy - x^2z = g(x, y, z)$$

Euler method for **second order differential equation:**

$$y_1 = y_0 + hf(x_0 + y_0 + z_0) = 1 + 0.1f(0, 1, 0)$$

$$= 1 + 0.1[1 + 2(0)(1) - 0^3(0)] = 1.1$$

$$\therefore \quad y(0.1) = 1.1$$

<u>Runge–Kutta Methods</u>

The **most widely known member** of the **Runge – Kutta** family is generally referred to as "**RK4**", the "**classic Runge – Kutta method**" or simply as "the **Runge – Kutta method**".

Let an **initial value problem** be specified as follows:

$$y' = \frac{dy}{dt} = f(t, y). \quad y(t_0) = y_0$$

Here $\boldsymbol{y}$ is an **unknown function** (**scalar** or **vector**) of **time** $\boldsymbol{t}$, from which **would like** to **approximate**; we are told that $\frac{dy}{dt}$, the **rate** at which $\boldsymbol{y}$ changes, is a **function** of $\boldsymbol{t}$ and of $\boldsymbol{y}$ itself. At the initial time $\boldsymbol{t_0}$ the corresponding $\boldsymbol{y}$ value is $\boldsymbol{y_0}$. The function $\boldsymbol{f}$ and the **initial conditions** $\boldsymbol{t_0}$, $\boldsymbol{y_0}$ are given.

Now picking a **step – size** $\boldsymbol{h > 0}$ and **define**:

$$\boldsymbol{y_{n+1} = y_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)h}$$

$$\boldsymbol{t_{n+1} = t_n + h}$$

for $\boldsymbol{n = 0, 1, 2, 3, \ldots}$, using

$$\boldsymbol{k_1 = f(t_n, y_n)},$$

$$\boldsymbol{k_2 = f\left(t_n + \frac{h}{2}, y_n + h\frac{k_1}{2}\right)},$$

$$\boldsymbol{k_3 = f\left(t_n + \frac{h}{2}, y_n + h\frac{k_2}{2}\right)},$$

$$\boldsymbol{k_4 = f\left(t_n + \frac{h}{2}, y_n + hk_3\right)}.$$

Here $\boldsymbol{y_{n+1}}$ is the "**RK4**" **approximation** of $\boldsymbol{y\ (t_{n+1})}$**,** and the **next value** $(\boldsymbol{y_{n+1}})$ is determined by the **present value** $(\boldsymbol{y_n})$ **plus** the **weighted average** of **four increments**, where **each increment** is the **product** of the **size** of **the interval**, $(\boldsymbol{h})$, and an **estimated slope specified** by **function**$(\boldsymbol{f})$ on the **right – hand side** of the **different equation**

- $\boldsymbol{k_1}$ is the **slope** at the **beginning** of the **interval**, using $\boldsymbol{y}$ (**Euler's method**);
- $\boldsymbol{k_2}$ is the **slope** at the **midpoint** of the **interval**, using $\boldsymbol{y}$ and $\boldsymbol{k_1}$;
- $\boldsymbol{k_3}$ is again the **slope** at the **midpoint,** but now using $\boldsymbol{y}$ and $\boldsymbol{k_2}$;
- $\boldsymbol{k_4}$is the **slope** at the **end** of the **interval**, using $\boldsymbol{y}$ and $\boldsymbol{k_3}$.

In **averaging** the **four slopes**, **greater weight** is **given** to the **slopes** at the **midpoint**. If $\boldsymbol{f}$ is **independent** of $\boldsymbol{y}$, so that the **differential equation** is **equivalent** to a **simple integral**, then **RK4** is **Simpson's rule**.

The **RK4 method** is a **fourth – order method**, meaning that the **local truncation error** is on the **order of** $\boldsymbol{O(h^5)}$, while the **total accumulated error** is on the order of $\boldsymbol{O(h^4)}$.

In **many practical applications** the **function** $\boldsymbol{f}$ is **independent** of $\boldsymbol{t}$ (so **called autonomous system**, or **time – invariant system**, especially in physics), and **their increments** are **not computed** at all and **not passed** to **function** $\boldsymbol{f}$, with only the **final formula** for $\boldsymbol{t_{n+1}}$used.

__Explicit Runge–Kutta methods__

The family of **explicit Runge–Kutta methods** is a **generalization** of the **RK4 method**. It is given by

$$y_{n+1} = y_n + h\sum_{i=1}^{s} b_i k_i$$

where,

$$k_1 = f(t_n, y_n),$$
$$k_2 = f[t_n + c_2 h, y_n + h(a_{21}k_1)],$$
$$k_3 = f[t_n + c_3 h, y_n + h(a_{31}k_1 + a_{32}k_2)],$$
$$\vdots \qquad\qquad \vdots \qquad\qquad \vdots$$
$$k_s = f\left[t_n + c_s h, y_n + h\left(a_{s1}k_1 + a_{s2}k_2 + \cdots + a_{s,s-1}k_{s-1}\right)\right],$$

To **specify a particular method**, one **needs to provide** the **integer** $\boldsymbol{s}$ (the **number** of **stages**), and the **coefficients** $\boldsymbol{a_{ij}}$ (for $\boldsymbol{1 \le j \le S}$), $\boldsymbol{b_i}$ (for $\boldsymbol{i = 1, 2, \cdots, S}$) and $\boldsymbol{c_i}$ (for $\boldsymbol{i = 1, 2, \cdots, S}$). The **matrix** $\boldsymbol{[a_{ij}]}$ is called the **Runge – Kutta matrix**, while the $\boldsymbol{b_i}$ and $\boldsymbol{c_i}$ are **known** as the **weights** and the **nodes**. These **data** are usually **arranged** in a **mnemonic device**, known as a **Butcher tableau** (after **John C. Butcher**):

$$\begin{array}{c|ccccc}
\mathbf{0} & & & & & \\
c_2 & a_{21} & & & & \\
c_3 & a_{31} & a_{32} & & & \\
\vdots & \vdots & & \ddots & & \\
c_S & a_{S1} & a_{S2} & \cdots & a_{S,S-1} & \\
\hline
 & b_1 & b_2 & \cdots & b_{S-1} & b_S
\end{array}$$

A **Taylor series expansion** shows that the **Runge–Kutta method** is **consistent if and only if**

$$\sum_{i=1}^{S} b_i = 1$$

There are also **accompanying requirements** if **one requires** the **method** to have a **certain order** p, meaning that the **local truncation error** is $O(h^{p+1})$. These **can be derived** from the **definition** of the **truncation error itself**. For example, a **two-stage method** has **order 2** if

$$b_1 + b_2 = 1 \quad , \qquad b_2 c_2 = \frac{1}{2} \quad , \qquad \text{and} \qquad b_2 a_{21} = \frac{1}{2}$$

Note that a **popular condition** for **determining coefficients** is:

$$\sum_{j=1}^{i-1} a_{ij} = c_i \quad \text{for} \quad i = 1, 2, \cdots, S$$

This condition alone, however, is neither sufficient nor necessary for consistency.

In general, if an **explicit S – stage Runge–Kutta method** has order p, then it can be proven that the **number of stages** must **satisfy** $S \geq p$, and if $p \geq 5$, then $S \geq p + 1$. However, it is **not known whether these bounds** are **sharp** in **all cases**; for example, all **known methods** of **order 8** have at least **11 stages**, though it is **possible** that there are **methods** with **fewer stages**. (The **bound above suggests** that there could be a **method with 9 stages**; but it could also be that the **bound is simply not sharp**.)

Indeed, it is **an open problem** what the **precise minimum number** of **stages** $\boldsymbol{S}$is for an **explicit Runge–Kutta method** to have order $\boldsymbol{p}$ in those cases where no methods have yet been discovered that satisfy the bounds above with equality. Some values which are known are:

$\boldsymbol{p}$	1	2	3	4	5	6	7	8
$\mathbf{min}\ \boldsymbol{S}$	1	2	3	4	6	7	9	11

The **provable bounds above** then imply that cannot find methods of orders $\boldsymbol{p = 1, 2, \dots, 6}$ that **require fewer stages** than the **methods** for these orders. However, it is **conceivable** that **might find** a **method** of **order** $\boldsymbol{p = 7}$ that **has only 8 stages**, whereas the only **ones known** today have at **least 9 stages** as shown in the table.

<u>**Example**</u>**:** The **RK4 method** falls in this **framework** its **tableau** is:

$$\begin{array}{c|cccc}
\mathbf{0} & & & & \\
\mathbf{\frac{1}{2}} & \mathbf{\frac{1}{2}} & & & \\
\mathbf{\frac{1}{2}} & \mathbf{0} & \mathbf{\frac{1}{2}} & & \\
\mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \\
\hline
 & \mathbf{\frac{1}{6}} & \mathbf{\frac{1}{3}} & \mathbf{\frac{1}{3}} & \mathbf{\frac{1}{6}}
\end{array}$$

A **slight variation** of "the" **Runge – Kutta method** is also due to **Kutta** in **1901** and is **called** the **3/8 – rule**. The **primary advantage** this method has is that almost all of the **error coefficients** are **smaller than** in the **popular method**, but it **requires slightly more FLOPs** (**floating-point operations**) per **time step**. Its **Butcher tableau** is:

$$\begin{array}{c|cccc} 0 & & & & \\ \frac{1}{3} & \frac{1}{3} & & & \\ \frac{2}{3} & -\frac{1}{3} & 1 & & \\ 1 & 1 & -1 & 1 & \\ \hline & \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8} \end{array}$$

However, the simplest **Runge – Kutta method** is the (**forward**) **Euler method**, **given by** the formula

$$y_{n+1} = y_n + h\, f(t_n, y_n)$$

This is the **only consistent explicit Runge – Kutta method** with **one stage**. The **corresponding tableau** is:

$$\begin{array}{c|c} 0 & \\ \hline & 1 \end{array}$$

<u>Example</u> 1: Find **y (0.2)** for

$$y' = \frac{x-y}{2} \;, \; x_0 = 0, \;\; y(0) = 1$$

with **step length 0.1** using **Runge – Kutta 2 method (1st order derivative).**

<u>Solution</u>:

Given: $y' = f(x, y) = \dfrac{x-y}{2}$, $x_0 = 0, y(0) = 1, h = 0.1,$ $\qquad y(0.2) = ?$

$$y_1 = y_0 + \frac{k_1 + k_2}{2}$$

<u># Method – 1: Using formula</u>

$$k_2 = hf(x_0 + h, y_0 + k_1)$$

<u>Second order R – K method</u>

$k_1 = hf(x_0, y_0) = 0.1f(0,1) = 0.1 \cdot (-0.5) = -0.05$

$k_2 = hf(x_0 + h, y_0 + k_1) = 0.1f[0 + 0.1, 1 + (-0.05)]$

$\quad = 0.1f(0.1, 0.95) = 0.1 \cdot (-0.425) = -0.0425$

$$y_1 = y_0 + \frac{\mathbf{k_1} + \mathbf{k_2}}{2} = 1 + \frac{-0.05 + (-0.0425)}{2} = 1 - 0.04625 = 0.9538$$

$$y_1 = 0.9538$$

$$y_2 = y_1 + \frac{\mathbf{k_1} + \mathbf{k_2}}{2}$$

$$\mathbf{k_1} = hf(x_1, y_1) = 0.1f(0.1, 0.9538) = 0.1 \cdot (-0.4269) = -0.0427$$

$$\mathbf{k_2} = hf(x_1 + h, y_1 + k_1) = 0.1f[0.1 + 0.1, 0.9538 + (-0.0427)]$$

$$= 0.1f(0.2, 0.9111) = 0.1 \cdot (-0.3555) = -0.0356$$

$$y_2 = y_1 + \frac{\mathbf{k_1} + \mathbf{k_2}}{2} = 0.9538 + \frac{-0.0427 + (-0.0356)}{2}$$

$$y_2 = 0.9538 - 0.0391 = 0.9146$$

$$y_3 = y_2 + \frac{\mathbf{k_1} + \mathbf{k_2}}{2}$$

$$\mathbf{k_1} = hf(x_2, y_2) = 0.1f(0.2, 0.9146) = 0.1 \cdot (-0.3573) = -0.0357$$

$$\mathbf{k_2} = hf(x_2 + h, y_2 + k_1) = 0.1f[0.2 + 0.1, 0.9146 + (-0.0357)]$$

$$= 0.1f(0.3, 0.8789) = 0.1 \cdot (-0.2895) = -0.02895$$

$$y_3 = y_2 + \frac{\mathbf{k_1} + \mathbf{k_2}}{2} = 0.9146 + \frac{-0.0357 + (-0.0289)}{2}$$

$$y_3 = 0.9146 - 0.0323 = 0.8823$$

$$\therefore \qquad y(0.2) = 0.83823$$

 # Method – 2: Using formula

$$y_1 = y_0 + \mathbf{k_2}$$

$$\mathbf{k_2} = hf\left(x_0 + \frac{h}{2}, y_0 + \frac{k_1}{2}\right)$$

$$\mathbf{k_1} = hf(x_0, y_0)$$

Second order R – K method

$$y_1 = y_0 + \mathbf{k_2}$$

$$\mathbf{k_1} = hf(x_0, y_0) = 0.1f(0, 1) = 0.1 \cdot (-0.5) = -0.05$$

$$\mathbf{k}_2 = hf\left(x_0 + \frac{h}{2}, y_0 + \frac{k_1}{2}\right) = 0.1f\left[0 + \frac{0.1}{2}, 1 + \left(-\frac{005}{2}\right)\right]$$

$$= 0.1f(0.05, 0.975) = 0.1 \cdot (-0.4625) = -0.04625$$

$$y_1 = y_0 + \mathbf{k}_2 = 1 + (-0.0462) = 0.9538$$

$$y_2 = y_1 + \mathbf{k}_2$$

$$\mathbf{k}_1 = hf(x_1, y_1) = 0.1f(0.1, 0.9538) = 0.1 \cdot (-0.4269) = -0.0427$$

$$\mathbf{k}_2 = hf\left(x_0 + \frac{h}{2}, y_0 + \frac{k_1}{2}\right) = 0.1f\left[0.1 + \frac{0.1}{2}, 0.9538 + \left(-\frac{0.0427}{2}\right)\right]$$

$$= 0.1f(0.15, 0.9324) = 0.1 \cdot (-0.43912) = -0.0391$$

$$y_2 = y_1 + \mathbf{k}_2 = 0.9538 + (-0.0391) = 0.9146$$

$$y_3 = y_2 + \mathbf{k}_2$$

$$\mathbf{k}_1 = hf(x_2, y_2) = 0.1f(0.2, 0.9146) = 0.1 \cdot (-0.3573) = -0.0357$$

$$\mathbf{k}_2 = hf\left(x_2 + \frac{h}{2}, y_2 + \frac{k_1}{2}\right) = 0.1f\left[0.2 + \frac{0.1}{2}, 0.9146 + \left(-\frac{0.0357}{2}\right)\right]$$

$$= 0.1f(0.25, 0.8968) = 0.1 \cdot (-0.3234) = -0.0323$$

$$y_3 = y_2 + \mathbf{k}_2 = 0.9146 + (-0.0323) = 0.8823$$

#Third order R – K method

$$y_1 = y_0 + \frac{1}{6}(\mathbf{k}_1 + 4\mathbf{k}_2 + \mathbf{k}_3)$$

$$\mathbf{k}_1 = hf(x_0, y_0)$$

$$\mathbf{k}_2 = hf\left(x_0 + \frac{h}{2}, y_0 + \frac{k_1}{2}\right)$$

$$\mathbf{k}_3 = hf(x_0 + h, y_0 + 2\mathbf{k}_2 - \mathbf{k}_1)$$

Example 2: Find y **(0.2)** for

$$y' = \frac{x - y}{2}\ ,\ x_0 = 0,\ \ y(0) = 1$$

with **step length 0.1** using **Runge – Kutta 3 method (1st order derivative).**

Solution:

Given: $y' = f(x, y) = \dfrac{x - y}{2}$, $x_0 = 0, y(0) = 1, h = 0.1,$ $y(0.2)$ =?

$$y_1 = y_0 + \frac{1}{6}(\mathrm{k}_1 + 4\mathrm{k}_2 + \mathrm{k}_3)$$

$$\mathrm{k}_1 = hf(x_0, y_0) = 0.1f(0, 1) = 0.1 \cdot (-0.5) = -0.05$$

$$\mathrm{k}_2 = hf\left(x_0 + \frac{h}{2}, y_0 + \frac{k_1}{2}\right) = 0.1f\left[0 + \frac{0.1}{2}, 1 + \left(-\frac{0.05}{2}\right)\right]$$

$$= 0.1f(0.05, 0.975) = 0.1 \cdot (-0.4625) = -0.04625$$

$$\mathrm{k}_3 = hf(x_0 + h, y_0 + 2\mathrm{k}_2 - \mathrm{k}_1)$$

$$= 0.1f[0 + 0.1, 1 + 2(-0.04625) - (-0.05)]$$

$$= 0.1f[0.1, 1 + 2(-0.04625) - (-0.05)] = 0.1f(0.1, 0.9574)$$

$$= 0.1 \cdot (-0.42875) = -0.04288$$

$$y_1 = y_0 + \frac{1}{6}(\mathrm{k}_1 + 4\mathrm{k}_2 + \mathrm{k}_3)$$

$$= 1 + \frac{1}{6}[-0.05 + 4(-0.04625) + (-0.04288)] = 0.95369$$

$$y_2 = y_1 + \frac{1}{6}(\mathrm{k}_1 + 4\mathrm{k}_2 + \mathrm{k}_3)$$

$$\mathrm{k}_1 = hf(x_1, y_1) = 0.1f(0.1, 0.9537) = 0.1 \cdot (-0.4268) = -0.0427$$

$$\mathrm{k}_2 = hf\left(x_1 + \frac{h}{2}, y_1 + \frac{k_1}{2}\right) = 0.1f\left[0.1 + \frac{0.1}{2}, 0.9537 + \left(-\frac{0.0427}{2}\right)\right]$$

$$= 0.1f(0.15, 0.9324) = 0.1 \cdot (-0.3912) = -0.03912$$

$$\mathrm{k}_3 = hf(x_1 + h, y_1 + 2\mathrm{k}_2 - \mathrm{k}_1)$$

$$= 0.1f[0.1 + 0.1, 0.9537 + 2(-0.03912) - (-0.0427)]$$

$$= 0.1f(0.2, 0.91814) = 0.1 \cdot (-0.0359) = -0.0359$$

$$y_2 = y_1 + \frac{1}{6}(\mathrm{k}_1 + 4\mathrm{k}_2 + \mathrm{k}_3)$$

$$= 0.9537 + \frac{1}{6}[-0.0427 + 4(-0.0391) + (-0.0359)] = 0.91451$$

$$y_3 = y_2 + \frac{1}{6}(\mathrm{k}_1 + 4\mathrm{k}_2 + \mathrm{k}_3)$$

$$\mathrm{k}_1 = hf(x_2, y_2) = 0.1f(0.2, 0.9145) = 0.1 \cdot (-0.35725) = -0.0357$$

$$\mathrm{k}_2 = hf\left(x_2 + \frac{h}{2}, y_2 + \frac{k_1}{2}\right) = 0.1f\left[0.2 + \frac{0.1}{2}, 0.9145 + \left(-\frac{0.0357}{2}\right)\right]$$

$$= 0.1f(0.25, 0.8967) = 0.1 \cdot (-0.3233) = -0.0323$$

$$\mathbf{k_3} = \boldsymbol{hf}(\boldsymbol{x_2} + \boldsymbol{h}, \boldsymbol{y_2} + \mathbf{2k_2} - \mathbf{k_1})$$

$$= \mathbf{0.1}\boldsymbol{f}[\mathbf{0.2} + \mathbf{0.1}, \mathbf{0.9145} + \mathbf{2}(-\mathbf{0.0323}) - (-\mathbf{0.0357})]$$

$$= \mathbf{0.1}\boldsymbol{f}(\mathbf{0.3}, \mathbf{0.8856}) = \mathbf{0.1} \cdot (-\mathbf{0.2928}) = -\mathbf{0.0293}$$

$$\boldsymbol{y_3} = \boldsymbol{y_2} + \frac{\mathbf{1}}{\mathbf{6}}(\mathbf{k_1} + \mathbf{4k_2} + \mathbf{k_3})$$

$$= \mathbf{0.9145} + \frac{\mathbf{1}}{\mathbf{6}}[-\mathbf{0.0357} + \mathbf{4}(-\mathbf{0.0323}) + (-\mathbf{0.0293})] = \mathbf{0.8823}$$

$$\therefore \quad \boldsymbol{y}(\mathbf{0.2}) = \mathbf{0.8823}$$

<u>#Forth order R – K method</u>

$$\boldsymbol{y_1} = \boldsymbol{y_0} + \frac{\mathbf{1}}{\mathbf{6}}(\mathbf{k_1} + \mathbf{2k_2} + \mathbf{2k_3} + \mathbf{k_4})$$

$$\mathbf{k_1} = \boldsymbol{hf}(\boldsymbol{x_0}, \boldsymbol{y_0})$$

$$\mathbf{k_2} = \boldsymbol{hf}\left(\boldsymbol{x_0} + \frac{\boldsymbol{h}}{\mathbf{2}}, \boldsymbol{y_0} + \frac{\boldsymbol{k_1}}{\mathbf{2}}\right)$$

$$\mathbf{k_3} = \boldsymbol{hf}\left(\boldsymbol{x_0} + \frac{\boldsymbol{h}}{\mathbf{2}}, \boldsymbol{y_0} + \frac{\boldsymbol{k_2}}{\mathbf{2}}\right)$$

$$\mathbf{k_4} = \boldsymbol{hf}(\boldsymbol{x_0} + \boldsymbol{h}, \boldsymbol{y_0} + \mathbf{k_3})$$

<u>Example</u> 2: Find **y (0.2)** for

$$\boldsymbol{y'} = \frac{\boldsymbol{x} - \boldsymbol{y}}{\mathbf{2}} \;, \; \boldsymbol{x_0} = \mathbf{0}, \;\; \boldsymbol{y}(\mathbf{0}) = \mathbf{1}$$

with **step length 0.1** using **Runge – Kutta 4 method (1st order derivative).**

<u>Solution</u>:

Given: $\boldsymbol{y'} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y}) = \frac{\boldsymbol{x} - \boldsymbol{y}}{\mathbf{2}}$, $\boldsymbol{x_0} = \mathbf{0}, \boldsymbol{y}(\mathbf{0}) = \mathbf{1}, \boldsymbol{h} = \mathbf{0.1},$ $\quad \boldsymbol{y}(\mathbf{0.2}) =?$

$$\boldsymbol{y_1} = \boldsymbol{y_0} + \frac{\mathbf{1}}{\mathbf{6}}(\mathbf{k_1} + \mathbf{2k_2} + \mathbf{2k_3} + \mathbf{k_4})$$

$$\mathbf{k_1} = \boldsymbol{hf}(\boldsymbol{x_0}, \boldsymbol{y_0}) = \mathbf{0.1}\boldsymbol{f}(\mathbf{0}, \mathbf{1}) = \mathbf{0.1} \cdot (-\mathbf{0.5}) = -\mathbf{0.05}$$

$$\mathbf{k_2} = \boldsymbol{hf}\left(\boldsymbol{x_0} + \frac{\boldsymbol{h}}{\mathbf{2}}, \boldsymbol{y_0} + \frac{\boldsymbol{k_1}}{\mathbf{2}}\right) = \mathbf{0.1}\boldsymbol{f}\left[\mathbf{0} + \frac{\mathbf{0.1}}{\mathbf{2}}, \mathbf{1} + \left(-\frac{\mathbf{0.05}}{\mathbf{2}}\right)\right]$$

$$= \mathbf{0.1}\boldsymbol{f}(\mathbf{0.05}, \mathbf{0.975}) = \mathbf{0.1} \cdot (-\mathbf{0.4625}) = -\mathbf{0.0463}$$

$$\mathbf{k}_3 = hf\left(x_0 + \frac{h}{2}, y_0 + \frac{k_2}{2}\right) = 0.1f\left[0 + \frac{0.1}{2}, 1 + \left(-\frac{0.0463}{2}\right)\right]$$

$$= 0.1f(0.05, 0.9769) = 0.1(-0.46344) = -0.0463$$

$$\mathbf{k}_4 = hf(x_0 + h, y_0 + \mathbf{k}_3) = 0.1[0 + 0.1, 1 + (-0.0463)]$$

$$= 0.1f(0.1, 0.9537) = 0.1(-0.4268) = -0.0427$$

$$y_1 = y_0 + \frac{1}{6}(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4)$$

$$= 1 + \frac{1}{6}[-0.05 + 2(-0.0463) + 2(-0.0463) + (-0.427)]$$

$$y_1 = 0.9537$$

$$y_2 = y_1 + \frac{1}{6}(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4)$$

$$\mathbf{k}_1 = hf(x_1, y_1) = 0.1f(0.1, 0.9537) = 0.1 \cdot (-0.4268) = -0.0427$$

$$\mathbf{k}_2 = hf\left(x_1 + \frac{h}{2}, y_1 + \frac{k_1}{2}\right) = 0.1f\left[0.1 + \frac{0.1}{2}, 0.9537 + \left(-\frac{0.0427}{2}\right)\right]$$

$$= 0.1f(0.15, 0.9323) = 0.1 \cdot (-0.3912) = -0.0391$$

$$\mathbf{k}_3 = hf\left(x_1 + \frac{h}{2}, y_1 + \frac{k_2}{2}\right) = 0.1f\left[0.1 + \frac{0.1}{2}, 0.9537 + \left(-\frac{0.0391}{2}\right)\right]$$

$$= 0.1f(0.15, 0.9341) = 0.1(-0.3921) = -0.0392$$

$$\mathbf{k}_4 = hf(x_1 + h, y_1 + \mathbf{k}_3) = 0.1[0.1 + 0.1, 0.9537 + (-0.0392)]$$

$$= 0.1f(0.2, 0.9145) = 0.1(-0.3572) = -0.0357$$

$$y_2 = y_1 + \frac{1}{6}(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4)$$

$$= 0.9537 + \frac{1}{6}[-0.0427 + 2(-0.0391) + 2(-0.0392) + (-0.0357)]$$

$$y_2 = 0.91451$$

$$y_3 = y_2 + \frac{1}{6}(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4)$$

$$\mathbf{k}_1 = hf(x_2, y_2) = 0.1f(0.2, 0.91451) = 0.1 \cdot (-0.35725) = -0.0357$$

$$\mathbf{k}_2 = hf\left(x_2 + \frac{h}{2}, y_2 + \frac{k_1}{2}\right) = 0.1f\left[0.2 + \frac{0.1}{2}, 0.91451 + \left(-\frac{0.0357}{2}\right)\right]$$

$$= 0.1f(0.25, 0.73589) = 0.1 \cdot (-48589) = -0.0486$$

$$\mathbf{k_3} = \boldsymbol{hf}\left(\boldsymbol{x_2} + \frac{\boldsymbol{h}}{\mathbf{2}}, \boldsymbol{y_2} + \frac{\boldsymbol{k_2}}{\mathbf{2}}\right) = \mathbf{0.1}\boldsymbol{f}\left[\mathbf{0.2} + \frac{\mathbf{0.1}}{\mathbf{2}}, \mathbf{0.91451} + \left(-\frac{\mathbf{0.0486}}{\mathbf{2}}\right)\right]$$

$$= \mathbf{0.1}\boldsymbol{f}(\mathbf{0.25}, \mathbf{0.6716}) = \mathbf{0.1}(-\mathbf{0.4216}) = -\mathbf{0.0422}$$

$$\mathbf{k_4} = \boldsymbol{hf}(\boldsymbol{x_2} + \boldsymbol{h}, \boldsymbol{y_2} + \mathbf{k_3}) = \mathbf{0.1}[\mathbf{0.2} + \mathbf{0.1}, \mathbf{0.9145} + (-\mathbf{0.0392})]$$

$$= \mathbf{0.1}\boldsymbol{f}(\mathbf{0.3}, \mathbf{0.8753}) = \mathbf{0.1}(-\mathbf{0.5753}) = -\mathbf{0.05753}$$

$$\boldsymbol{y_3} = \boldsymbol{y_2} + \frac{\mathbf{1}}{\mathbf{6}}(\mathbf{k_1} + \mathbf{2k_2} + \mathbf{2k_3} + \mathbf{k_4})$$

$$= \mathbf{0.9145} + \frac{\mathbf{1}}{\mathbf{6}}[-\mathbf{0.0357} + \mathbf{2}(-\mathbf{0.0486}) + \mathbf{2}(-\mathbf{0.0422}) + (-\mathbf{0.05753})]$$

$$\boldsymbol{y_3} = \mathbf{0.7824}$$

<u>Second – order methods with two stages</u>

An **example** of a **second – order method** with **two stages** is **provided** by the **midpoint method**

$$\boldsymbol{y_{n+1}} = \boldsymbol{y_n} + \boldsymbol{h\,f}\ \left[\boldsymbol{t_n} + \frac{\mathbf{1}}{\mathbf{2}}\boldsymbol{h}\,, \boldsymbol{y_n} + \frac{\mathbf{1}}{\mathbf{2}}\boldsymbol{h\,f}(\boldsymbol{t_n}, \boldsymbol{y_n})\right]$$

The **corresponding tableau** is:

$$\begin{array}{c|cc} \mathbf{0} & & \\ \frac{\mathbf{1}}{\mathbf{2}} & \frac{\mathbf{1}}{\mathbf{2}} & \\ \hline & \mathbf{0} & \mathbf{1} \end{array}$$

The **midpoint method** is **not** the **only second-order Runge – Kutta method** with **two stages**; there is a family of such methods, parameterized by **α** and given by the formula

$$\boldsymbol{y_{n+1}} = \boldsymbol{y_n} + \boldsymbol{h}\left[\left(\mathbf{1} - \frac{\mathbf{1}}{\mathbf{2}\boldsymbol{\alpha}}\right)\boldsymbol{f}(\boldsymbol{t_n}\,, \boldsymbol{y_n}) + \frac{\mathbf{1}}{\mathbf{2}\boldsymbol{\alpha}}\ \boldsymbol{f}\{(\boldsymbol{t_n} + \boldsymbol{\alpha h}\,, \boldsymbol{y_n} + \boldsymbol{\alpha h f}(\boldsymbol{t_n}, \boldsymbol{y_n})\}\right]$$

Its **Bucher tableau** is:

$$\begin{array}{c|cc} \mathbf{0} & & \\ \boldsymbol{\alpha} & \boldsymbol{\alpha} & \\ \hline & \left(\mathbf{1} - \frac{\mathbf{1}}{\mathbf{2}\boldsymbol{\alpha}}\right) & \frac{\mathbf{1}}{\mathbf{2}\boldsymbol{\alpha}} \end{array}$$

In this family,

- $\boldsymbol{\alpha = \frac{1}{2}}$ gives the **midpoint method**
- $\boldsymbol{\alpha = 1}$ is **Heun`s method**
- $\boldsymbol{\alpha = \frac{2}{3}}$ is **Ralston`s method**

Use

As an **example**, consider the **two-stage second – order Runge – Kutta method** with $\boldsymbol{\alpha} = \frac{2}{3}$, also **known** as **Ralston method**. It is **given** by the **tableau**

$$\begin{array}{c|cc} \mathbf{0} & & \\ \mathbf{\frac{2}{3}} & \mathbf{\frac{2}{3}} & \\ \hline & \mathbf{\frac{1}{4}} & \mathbf{\frac{3}{4}} \end{array}$$

with the corresponding equations

$$\boldsymbol{k_1 = f(t_n, y_n)},$$

$$\boldsymbol{k_2 = f\left[t_n + \frac{2}{3}h\,, y_n + \frac{2}{3}hk_1\right]},$$

$$\boldsymbol{y_{n+1} = y_n + h\left(\frac{1}{4}k_1 + \frac{3}{4}k_2\right)}$$

The method is **used** to **solve** the **initial – value problem**

$$\boldsymbol{\frac{dy}{dt} = tan(y) + 1} \quad , \qquad \boldsymbol{y_0 = 1\,, t \in (1, 1.1)}$$

with **step size** $\boldsymbol{h = 0.025}$, so the **method needs** to take **four steps**.

The **method proceeds as follows**:

$\boldsymbol{t_0 = 1}$:

$$\boldsymbol{y_0 = 1}$$

$\boldsymbol{t_1 = 1.025}$:

$$\boldsymbol{y_0 = 1}$$

$$\boldsymbol{k_1 = 2.557407725}$$

$$k_2 = f\left(t_0 + \frac{2}{3}h\,, y_0 + \frac{2}{3}hk_1\right) = 2.7138981400$$

$$y_1 = y_0 + h\left(\frac{1}{4}k_1 + \frac{3}{4}k_2\right) = 1.066869388$$

$t_2 = 1.05$:

$$y_1 = 1.066869388$$

$$k_1 = 2.813524695$$

$$k_2 = f\left(t_1 + \frac{2}{3}h\,, y_1 + \frac{2}{3}hk_1\right)$$

$$y_2 = y_1 + h\left(\frac{1}{4}k_1 + \frac{3}{4}k_2\right) = 1.141332181$$

$t_3 = 1.075$:

$$y_2 = 1.141332181$$

$$k_1 = 3.183536647$$

$$k_2 = f\left(t_2 + \frac{2}{3}h\,, y_2 + \frac{2}{3}hk_1\right) = 2.7138981400$$

$$y_3 = y_2 + h\left(\frac{1}{4}k_1 + \frac{3}{4}k_2\right) = 1.227417567$$

$t_4 = 1.1$:

$$y_3 = 1.227417567$$

$$k_1 = 3.796866512$$

$$k_2 = f\left(t_3 + \frac{2}{3}h\,, y_3 + \frac{2}{3}hk_1\right) = 2.7138981400$$

$$y_4 = y_3 + h\left(\frac{1}{4}k_1 + \frac{3}{4}k_2\right) = 1.335079087$$

Second Order R – K Method (Second Order Differential Equation)

Method – 1:

$$y_1 = y_0 + \frac{k_1 + k_2}{2}$$

$k_1 = hf(x_0, y_0, z_0)$

$k_2 = hf(x_0 + h, y_0 + k_1, z_0 + l_1)$

$l_1 = hg(x_0, y_0, z_0)$

$l_2 = hg(x_0 + h, y_0 + k_1, z_0 + l_1)$

Method – 2:

$$y_1 = y_0 + k_2$$

$k_2 = hf\left(x_0 + \frac{h}{2}, y_0 + \frac{k_1}{2}, z_0 + \frac{l_1}{2}\right)$

$k_1 = hf(x_0, y_0, z_0)$

$l_1 = hg(x_0, y_0, z_0)$

$l_2 = hg\left(x_0 + \frac{h}{2}, y_0 + \frac{k_1}{2}, z_0 + \frac{l_1}{2}\right)$

Example: Find $y(0.1)$ for

$$y'' = 1 + 2xy - x^2 z$$

using the initial conditions, $x_0 = 0$, $y_0 = 1$, and $z_0 = 0$ with step length 0.1 using **Runge – Kutta 2 method (2nd order derivative).**

Solution:

The Given equation is: $y'' = \frac{d^2y}{dx^2} = 1 + 2xy - x^2 z$

Put $\frac{dy}{dx} = z$ and differentiate with respect to x, yields $\frac{d^2y}{dx^2} = \frac{dz}{dx}$

Thus, the given **equation** becomes:

$$\frac{d^2y}{dx^2} = \frac{dz}{dx} = 1 + 2xy - x^2 z$$

with $\boldsymbol{y(0) = 1, y'(0) = 0, x_0 = 0, h = 0.1,} \quad \boldsymbol{y(0.1)} =?$

Thus, the system of equations is:

$$\frac{dy}{dx} = z = f(x, y, z)$$

$$\frac{dz}{dx} = 1 + 2xy - x^2 z = g(x, y, z)$$

<u>Method – 1</u>: Using formula

$$y_1 = y_0 + \frac{k_1 + k_2}{2}$$

$$k_1 = hf(x_0, y_0, z_0) = (0.1)f(0, 1, 0) = (0.1) \cdot (0) = 0$$

$$k_2 = hf(x_0 + h, y_0 + k_1, z_0 + l_1)$$

$$l_1 = hg(x_0, y_0, z_0) = (0.1)g(0, 1, 0)$$

$$= (0.1)[1 + 2(0)(1) - (0)^2(0)] = 0.1$$

$$\Rightarrow k_2 = (0.1)f(0 + 0.1, 1 + 0, 0 + 0.1) = (0.1)f(0.1, 1, 0.1)$$

$$= (0.1)(0.1) = 0.01$$

$$l_2 = hg(x_0 + h, y_0 + k_1, z_0 + l_1) = (0.1)g(0 + 0.1, 1 + 0, 0 + 0.1)$$

$$= (0.1)g(0.1, 1, 0.1) = (0.1)[1 + 2(0.1)(1) - (0.1)^2(0.1)]$$

$$= (0.1)(1 + 0.2 - 0.001) = (0.1)(1.199) = 0.1199$$

$$\Rightarrow \quad y_1 = y_0 + \frac{k_1 + k_2}{2} = 1 + \frac{0 + 0.01}{2} = 1.005$$

$$\therefore \quad y(0.1) = 1.005$$

<u>Method – 2</u>: Using formula

$$y_1 = y_0 + k_2$$

$$k_2 = hf\left(x_0 + \frac{h}{2}, y_0 + \frac{k_1}{2}, z_0 + \frac{l_1}{2}\right)$$

$$k_1 = hf(x_0, y_0, z_0) = (0.1)f(0, 1, 0) = (0.1) \cdot (0) = 0$$

$$l_1 = hg(x_0, y_0, z_0) = (0.1)g(0, 1, 0)$$

$$= (0.1)[1 + 2(0)(1) - (0)^2(0)] = 0.1$$

$$\Rightarrow \boldsymbol{k_2} = (\mathbf{0.1})\boldsymbol{f}\left(\mathbf{0}+\frac{\mathbf{0.1}}{\mathbf{2}},\mathbf{1}+\frac{\mathbf{0}}{\mathbf{2}},\mathbf{0}+\frac{\mathbf{0.1}}{\mathbf{2}}\right) = (\mathbf{0.1})\boldsymbol{f}(\mathbf{0.05},\mathbf{1},\mathbf{0.05})$$
$$= (\mathbf{0.1})(\mathbf{0.05}) = \mathbf{0.005}$$
$$\boldsymbol{l_2} = \boldsymbol{h}\mathbf{g}\left(\boldsymbol{x_0}+\frac{\boldsymbol{h}}{\mathbf{2}},\boldsymbol{y_0}+\frac{\boldsymbol{k_1}}{\mathbf{2}},\boldsymbol{z_0}+\frac{\boldsymbol{l_1}}{\mathbf{2}}\right) = (\mathbf{0.1})\boldsymbol{f}\left(\mathbf{0}+\frac{\mathbf{0.1}}{\mathbf{2}},\mathbf{1}+\frac{\mathbf{0}}{\mathbf{2}},\mathbf{0}+\frac{\mathbf{0.1}}{\mathbf{2}}\right)$$
$$= (\mathbf{0.1})\mathbf{g}(\mathbf{0.05},\mathbf{1},\mathbf{0.05}) = (\mathbf{0.1})[\mathbf{1}+\mathbf{2}(\mathbf{0.05})(\mathbf{1}) - (\mathbf{0.05})^{\mathbf{2}}(\mathbf{0.05})]$$
$$= (\mathbf{0.1})(\mathbf{1.099875}) = \mathbf{0.10998751}$$
$$\Rightarrow \boldsymbol{y_1} = \boldsymbol{y_0} + \mathbf{k_2} = \mathbf{1} + \mathbf{0.005} = \mathbf{1.005}$$
$$\therefore \quad \boldsymbol{y}(\mathbf{0.1}) = \mathbf{1.005}$$

<u>Third Order R – K Method (Second Order Differential Equation)</u>

$$\boldsymbol{y_1} = \boldsymbol{y_0} + \frac{\mathbf{1}}{\mathbf{6}}(\boldsymbol{k_1} + \mathbf{4}\boldsymbol{k_2} + \boldsymbol{k_3})$$

$$\boldsymbol{k_1} = \boldsymbol{hf}(\boldsymbol{x_0},\boldsymbol{y_0},\boldsymbol{z_0})$$
$$\boldsymbol{l_1} = \boldsymbol{h}\mathbf{g}(\boldsymbol{x_0},\boldsymbol{y_0},\boldsymbol{z_0})$$
$$\boldsymbol{k_2} = \boldsymbol{hf}\left(\boldsymbol{x_0}+\frac{\boldsymbol{h}}{\mathbf{2}},\boldsymbol{y_0}+\frac{\boldsymbol{k_1}}{\mathbf{2}},\boldsymbol{z_0}+\frac{\boldsymbol{l_1}}{\mathbf{2}}\right)$$
$$\boldsymbol{l_2} = \boldsymbol{h}\mathbf{g}\left(\boldsymbol{x_0}+\frac{\boldsymbol{h}}{\mathbf{2}},\boldsymbol{y_0}+\frac{\boldsymbol{k_1}}{\mathbf{2}},\boldsymbol{z_0}+\frac{\boldsymbol{l_1}}{\mathbf{2}}\right)$$
$$\boldsymbol{k_3} = \boldsymbol{hf}(\boldsymbol{x_0}+\boldsymbol{h},\boldsymbol{y_0}+\mathbf{2}\boldsymbol{k_2}-\boldsymbol{k_1},\boldsymbol{z_0}+\mathbf{2}\boldsymbol{l_2}-\boldsymbol{l_1})$$
$$\boldsymbol{l_3} = \boldsymbol{h}\mathbf{g}(\boldsymbol{x_0}+\boldsymbol{h},\boldsymbol{y_0}+\mathbf{2}\boldsymbol{k_2}-\boldsymbol{k_1},\boldsymbol{z_0}+\mathbf{2}\boldsymbol{l_2}-\boldsymbol{l_1})$$

<u>**Example**</u>**: Find $\boldsymbol{y}(\mathbf{0.1})$** for

$$\boldsymbol{y''} = \mathbf{1} + \mathbf{2}\boldsymbol{xy} - \boldsymbol{x}^{\mathbf{2}}\boldsymbol{z}$$

using the initial conditions, $\boldsymbol{x_0} = \mathbf{0}$, $\boldsymbol{y_0} = \mathbf{1}$, and $\boldsymbol{z_0} = \mathbf{0}$ with step length $\mathbf{0.1}$ using **Runge – Kutta 3 method (2nd order derivative).**

<u>**Solution**</u>**:**

The Given equation is: $\boldsymbol{y''} = \dfrac{\boldsymbol{d^2y}}{\boldsymbol{dx^2}} = \mathbf{1} + \mathbf{2}\boldsymbol{xy} - \boldsymbol{x}^{\mathbf{2}}\boldsymbol{z}$

Put $\dfrac{\boldsymbol{dy}}{\boldsymbol{dx}} = \boldsymbol{z}$ and differentiate with respect to $\boldsymbol{x}$, yields $\dfrac{\boldsymbol{d^2y}}{\boldsymbol{dx^2}} = \dfrac{\boldsymbol{dz}}{\boldsymbol{dx}}$

Thus, the given **equation** becomes:

$$\frac{\boldsymbol{d^2y}}{\boldsymbol{dx^2}} = \frac{\boldsymbol{dz}}{\boldsymbol{dx}} = \mathbf{1} + \mathbf{2}\boldsymbol{xy} - \boldsymbol{x}^{\mathbf{2}}\boldsymbol{z}$$

with $\boldsymbol{y}(\mathbf{0}) = \mathbf{1}, \boldsymbol{y'}(\mathbf{0}) = \mathbf{0}, \boldsymbol{x_0} = \mathbf{0}, \boldsymbol{h} = \mathbf{0.1},\qquad \boldsymbol{y}(\mathbf{0.1}) = ?$

Thus, the system of equations is:

$$\frac{dy}{dx} = z = f(x, y, z)$$

$$\frac{dz}{dx} = 1 + 2xy - x^2 z = g(x, y, z)$$

Third Order R – K Method

$$y_1 = y_0 + \frac{1}{6}(k_1 + 4k_2 + k_3)$$

$$k_1 = hf(x_0, y_0, z_0) = (0.1)f(0, 1, 0) = (0.1) \cdot (0) = 0$$

$$l_1 = hg(x_0, y_0, z_0) = (0.1)g(0, 1, 0)$$

$$= (0.1)[1 + 2(0)(1) - (0)^2(0)] = 0.1$$

$$k_2 = hf\left(x_0 + \frac{h}{2}, y_0 + \frac{k_1}{2}, z_0 + \frac{l_1}{2}\right) = (0.1)f\left(0 + \frac{0.1}{2}, 1 + \frac{0}{2}, 0 + \frac{0.1}{2}\right)$$

$$= (0.1)(0.05) = 0.005$$

$$l_2 = hg\left(x_0 + \frac{h}{2}, y_0 + \frac{k_1}{2}, z_0 + \frac{l_1}{2}\right) = (0.1)g\left(0 + \frac{0.1}{2}, 1 + \frac{0}{2}, 0 + \frac{0.1}{2}\right)$$

$$= (0.1)g(0.05, 1, 0.05) = (0.1)[1 + 2(0.05)(1) - (0.05)^2(0.05)]$$

$$= (0.1)(1.099875) = 0.10998751$$

$$k_3 = hf(x_0 + h, y_0 + 2k_2 - k_1, z_0 + 2l_2 - l_1)$$

$$= (0.1)[0 + 0.1, 1 + 2(0.005) - 0, 0 + 2(0.10999) - 0.1]$$

$$= (0.1)(0.1, 1.01, 0.11998) = (0.1)(0.12) = 0.012$$

$$l_3 = hg(x_0 + h, y_0 + 2k_2 - k_1, z_0 + 2l_2 - l_1)$$

$$= (0.1)g[0 + 0.1, 1 + 2(0.005) - 0, 0 + 2(0.10999) - 0.1]$$

$$= (0.1)g(0.1, 1.01, 0.11998)$$

$$= (0.1)[1 + 2(0.1)(1.01) - (0.1)^2(0.12)] = (0.1)(1.20202)$$

$$= 0.1202$$

$$\Rightarrow \; y_1 = y_0 + \frac{1}{6}(k_1 + 4k_2 + k_3)$$

$$= 1 + \frac{1}{6}[0 + 4(0.005) + 0.012] = 1.0053$$

$$\therefore \quad y(0.1) = 1.005$$

Fourth Order R – K Method for 2nd Order Differential Equation)

$$y_1 = y_0 + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

$k_1 = hf(x_0, y_0, z_0)$

$l_1 = hg(x_0, y_0, z_0)$

$k_2 = hf\left(x_0 + \frac{h}{2}, y_0 + \frac{k_1}{2}, z_0 + \frac{l_1}{2}\right)$

$l_2 = hg\left(x_0 + \frac{h}{2}, y_0 + \frac{k_1}{2}, z_0 + \frac{l_1}{2}\right)$

$k_3 = hf\left(x_0 + \frac{h}{2}, y_0 + \frac{k_2}{2}, z_0 + \frac{l_2}{2}\right)$

$l_3 = hg\left(x_0 + \frac{h}{2}, y_0 + \frac{k_2}{2}, z_0 + \frac{l_2}{2}\right)$

$k_4 = hf(x_0 + h, y_0 + k_3, z_0 + l_3)$

$l_4 = hg(x_0 + h, y_0 + k_3, z_0 + l_3)$

Example: Find $y(0.1)$ for

$$y'' = 1 + 2xy - x^2 z$$

using the initial conditions, $x_0 = 0$, $y_0 = 1$, and $z_0 = 0$ with step length 0.1 using **Runge – Kutta 4 method (2nd order derivative).**

Solution:

The Given equation is: $y'' = \frac{d^2y}{dx^2} = 1 + 2xy - x^2 z$

Put $\frac{dy}{dx} = z$ and differentiate with respect to x, yields $\frac{d^2y}{dx^2} = \frac{dz}{dx}$

Thus, the given **equation** becomes:

$$\frac{d^2y}{dx^2} = \frac{dz}{dx} = 1 + 2xy - x^2 z$$

with $y(0) = 1, y'(0) = 0, x_0 = 0, h = 0.1,$ $y(0.1) = ?$

Thus, the system of equations is:

$$\frac{dy}{dx} = z = f(x, y, z)$$

$$\frac{dz}{dx} = 1 + 2xy - x^2 z = g(x, y, z)$$

<u>Fourth Order R – K Method for 2nd order differential equation)</u>

$$y_1 = y_0 + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

$k_1 = hf(x_0, y_0, z_0) = (0.1)f(0,1,0) = (0.1)\cdot(0) = 0$

$l_1 = hg(x_0, y_0, z_0) = (0.1)g(0,1,0)$

$= (0.1)[1 + 2(0)(1) - (0)^2(0)] = 0.1$

$k_2 = hf\left(x_0 + \frac{h}{2}, y_0 + \frac{k_1}{2}, z_0 + \frac{l_1}{2}\right) = (0.1)f\left(0 + \frac{0.1}{2}, 1 + \frac{0}{2}, 0 + \frac{0.1}{2}\right)$

$= (0.1)(0.05) = 0.005$

$l_2 = hg\left(x_0 + \frac{h}{2}, y_0 + \frac{k_1}{2}, z_0 + \frac{l_1}{2}\right) = (0.1)g\left(0 + \frac{0.1}{2}, 1 + \frac{0}{2}, 0 + \frac{0.1}{2}\right)$

$= (0.1)g(0.05, 1, 0.05) = (0.1)[1 + 2(0.05)(1) - (0.05)^2(0.05)]$

$= (0.1)(1.099875) = 0.10998751 = 0.11$

$k_3 = hf\left(x_0 + \frac{h}{2}, y_0 + \frac{k_2}{2}, z_0 + \frac{l_2}{2}\right)$

$= (0.1)f\left(0 + \frac{0.1}{2}, 1 + \frac{0.005}{2}, 0 + \frac{0.11}{2}\right)$

$= (0.1)f(0.05, 1.0025, 0.055) = (0.1)(0.055) = 0.0055$

$l_3 = hg\left(x_0 + \frac{h}{2}, y_0 + \frac{k_2}{2}, z_0 + \frac{l_2}{2}\right)$

$= (0.1)f\left(0 + \frac{0.1}{2}, 1 + \frac{0.005}{2}, 0 + \frac{0.11}{2}\right)$

$= (0.1)f(0.05, 1.0025, 0.055)$

$= (0.1)[1 + 2(0.05)(1.0025) - (0.05)^2(0.055)]$

$= (0.1)(1.1001) = 0.11$

$k_4 = hf(x_0 + h, y_0 + k_3, z_0 + l_3)$

$= (0.1)f(0 + 0.1, 1 + 0.0055, 0 + 0.11)$

$= (0.1)f(0.1, 1.0055, 0.11) = (0.1)(0.11) = 0.011$

$l_4 = hg(x_0 + h, y_0 + k_3, z_0 + l_3) =$

$= (0.1)g(0 + 0.1, 1 + 0.0055, 0 + 0.11)$

$= (0.1)g(0.1, 1.0055, 0.11) = (0.1)$

$$= (\mathbf{0.1})[\mathbf{1} + \mathbf{2}(\mathbf{0.1})(\mathbf{1.0025}) - (\mathbf{0.1})^2(\mathbf{0.055})]$$

$$= (\mathbf{0.1})(\mathbf{1.9995}) = \mathbf{0.12}$$

$$\Longrightarrow \boldsymbol{y_1} = \boldsymbol{y_0} + \frac{\mathbf{1}}{\mathbf{6}}(\boldsymbol{k_1} + \mathbf{2}\boldsymbol{k_2} + \mathbf{2}\boldsymbol{k_3} + \boldsymbol{k_4})$$

$$= \mathbf{1} + \frac{\mathbf{1}}{\mathbf{6}}[\mathbf{0} + \mathbf{2}(\mathbf{0.005}) + \mathbf{2}(\mathbf{0.0055}) + \mathbf{0.011}] = \mathbf{1.0053}$$

$$\therefore \quad \boldsymbol{y}(\mathbf{0.1}) = \mathbf{1.005}$$

<u>Numerical Integration</u>

The **area under curve** is **one** of the **common most application** for "**Numerical Integration**" because the "**Analytical Methods**" are **difficult solution** or **there's no solution** in this way.

It is known that a **definite integral** of the form

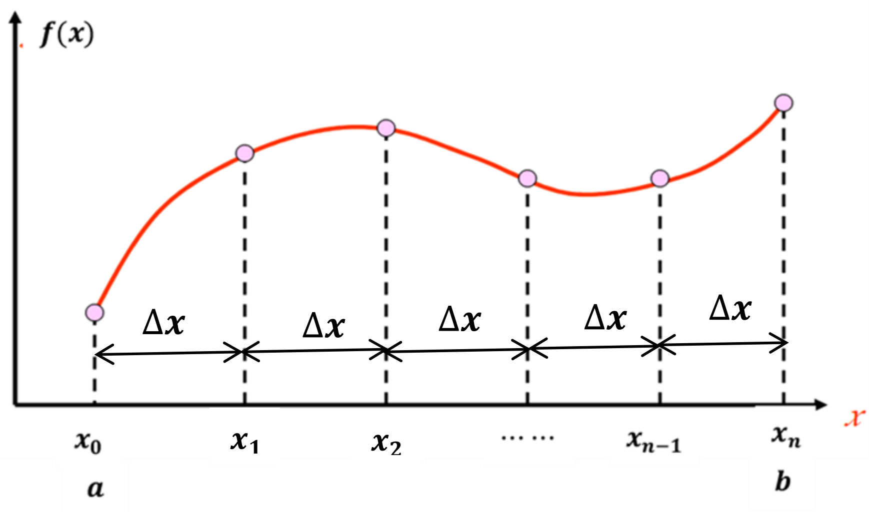
$$\int_{a}^{b} f(x)dx$$

Represents the **area unde**r the **curve**

$$y = f(x)$$

enclosed between the **limits** $x = a$ and $x = b$.

Thus, the **goal** is to **approximate** the **definite integral** of a **function** $f(x)$ **over the interval** $[a . b]$ by **evaluating** the **function** $f(x)$ at a **finite number** of **sample points. Thus,** the **process** of **finding** the **approximation** to the **definite Integral** is **known** as <u>**Numerical Integration**</u>.



Let $x_0 . x_1 . x_2 . \cdots . x_n$ be **given set** of **observations**, and let $y_0 . y_1 . y_2 . \cdots . y_n$ be the **corresponding values** for the **curve** $y = f(x)$.

Suppose that $a = x_0 < x_1 < x_2 < \cdots < x_m < b$.

A **formula** of the **form**:

$$\boldsymbol{Q}[\boldsymbol{f}] = \sum_{k=0}^{m} \boldsymbol{w}_k\, \boldsymbol{f}(\boldsymbol{x}_k) = \boldsymbol{w}_0\, \boldsymbol{f}(\boldsymbol{x}_0) + \boldsymbol{w}_1\, \boldsymbol{f}(\boldsymbol{x}_1) + \cdots + \boldsymbol{w}_m\, \boldsymbol{f}(\boldsymbol{x}_m)$$

with the **property** that:

$$\int_a^b \boldsymbol{f}(\boldsymbol{x})\boldsymbol{dx} = \boldsymbol{Q}[\boldsymbol{f}] + \boldsymbol{E}[\boldsymbol{f}]$$

is **called** a **numerical integration** or **quadrature formula**.

- The **term** $\boldsymbol{E}[\boldsymbol{f}]$ is **called** the **truncation error** for **integration**.
- The **values** $\{\boldsymbol{x}_k\}_{k=0}^{m}$ are **called** the **quadrature nodes** and
- $\{\boldsymbol{w}_k\}_{k=0}^{m}$ are **called** **weights**.

Closed Newton-Cotes Quadrature Formulas

The **Newton – Cotes Integration formulas** are the **most numerical integration schemes.** They are **based on** the **strategy** of **replacing** a **complicated function** or **tabulated data** with an **approximating function** that is **easy to integrate**:

$$\int_a^b \boldsymbol{f}(\boldsymbol{x})\boldsymbol{dx} = \int_a^b \boldsymbol{f}_n(\boldsymbol{x})\boldsymbol{dx}$$

$$\boldsymbol{f}_n(\boldsymbol{x}) = \sum_{i=0}^{n} \boldsymbol{a}_i \boldsymbol{f}(\boldsymbol{x}_i) = \boldsymbol{a}_0 \boldsymbol{f}(\boldsymbol{x}_0) + \boldsymbol{a}_1 \boldsymbol{f}(\boldsymbol{x}_1) + \cdots + \boldsymbol{a}_n \boldsymbol{f}(\boldsymbol{x}_n)$$

where

$$\boldsymbol{x}_n = \boldsymbol{x}_0 + \boldsymbol{n}\,\Delta\boldsymbol{x} \quad . \quad \boldsymbol{n} = \textbf{number of subintervals} = \mathbf{1}, \mathbf{2}, \mathbf{3}, \cdots$$

$$\Longrightarrow \quad \Delta\boldsymbol{x} = \frac{\boldsymbol{x}_n - \boldsymbol{x}_0}{\boldsymbol{n}}$$

are **equally spaced nodes** and

$$\boldsymbol{f}_n = \boldsymbol{f}(\boldsymbol{x}_n).$$

The **first four closed Newton-Cotes quadrature formulas** are:

$$\int_{x_0}^{x_1} f(x)dx \approx \frac{\Delta x}{2}(f_0 + f_1) \text{ -------- the Trapezoidal Rule}$$

$$\int_{x_0}^{x_2} f(x)dx \approx \frac{\Delta x}{3}(f_0 + 4f_1 + f_2) \text{ ---- the Simpson`s 1/3 Rule}$$

$$\int_{x_0}^{x_3} f(x)dx \approx \frac{3\Delta x}{8}(f_0 + 3f_1 + 3f_2 + f_3) \text{ the Simpson`s 3/8 Rule}$$

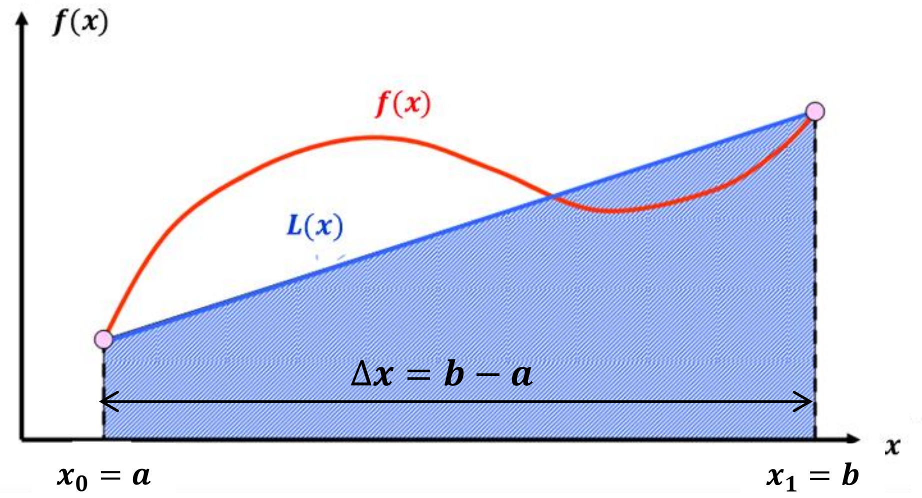
$$\int_{x_0}^{x_4} f(x)dx \approx \frac{2\Delta x}{45}(7f_0 + 32f_1 + 12f_2 + 32f_3 + 7f_4) \text{ the Boole`s Rule}$$

✪ Trapezoidal Rule

$$I = \frac{\Delta x}{2}\left[(f(x_0) + 2\{f(x_1) + f(x_2) + f(x_3) + \cdots + f(x_{n-1}\} + f(x_n)\right]$$

Here n is **a number** of **intervals,** and

$$\Delta x = \frac{b-a}{n} \text{ . when } n = 1 \implies \Delta x = b - a$$



Lagrange Interpolation:

$$L(x) = \frac{x - x_1}{x_0 - x_1}\, f(x_0) + \frac{x - x_1}{x_1 - x_0}\, f(x_1)$$

Let $a = x_0$, $b = x_1$, $\Delta x = b - a$

$$\xi = \frac{x - a}{b - a} \quad , \quad d\xi = \frac{dx}{\Delta x}$$

$$\left\{\begin{matrix} x = a & \Rightarrow & \xi = \dfrac{x - a}{b - a} = \dfrac{a - a}{b - a} = 0 \\ x = b & \Rightarrow & \xi = \dfrac{x - a}{b - a} = \dfrac{b - a}{b - a} = 1 \end{matrix}\right\}$$

$$\Rightarrow \quad L(\xi) = (1 - \xi) f(a) + (\xi)\, f(b)$$

Integrating to **obtain** the **rule:**

$$I = \int_a^b f(x)\, dx \approx \int_a^b L(x)\, dx = \Delta x \int_0^1 L(\xi)\, d\xi$$

$$= \Delta x\, f(a) \int_0^1 (1 - \xi)\, d\xi + \Delta x\, f(b) \int_0^1 L(\xi)\, d\xi$$

$$= \Delta x\, f(a) \left(\xi - \frac{\xi^2}{2}\right)\Bigg|_0^1 + \Delta x\, f(b) \frac{\xi^2}{2}\Bigg|_0^1$$

$$\Rightarrow \quad I \approx \frac{\Delta x}{2} [f(a) + f(b)]$$

<u>Example</u> 1: Evaluate the **integral** by using **<u>Trapezoidal method</u>**

$$\int_0^4 x\, e^{2x}\, dx$$

<u>Solution</u>:

- The **<u>exact solution</u>** is:

$$\int_0^4 x\, e^{2x}\, dx = \int u dv = uv - \int v\, du = \left| \begin{matrix} u = x \Rightarrow & du = dx \\ dv = e^{2x} dx \Rightarrow & v = \dfrac{e^{2x}}{2} \end{matrix} \right|$$

$$= x\, \frac{e^{2x}}{2} - \int_0^4 \frac{e^{2x}}{2} dx = \left[\frac{1}{2} x\, e^{2x} - \frac{e^{2x}}{4} \right]_0^4$$

$$= \left[\frac{x}{2} e^{2x} - \frac{1}{4}\, e^{2x} \right]_0^4 = \frac{1}{4}\, e^{2x}(2x - 1) \Big|_0^4$$

$$= 5216.93$$

- **<u>Trapezoidal Rule</u>:**

Let $a = x_0 = 0 \quad .b = x_1 = 4 \quad \Rightarrow \quad \Delta x = x_1 - x_0 = 4 - 0 = 4$

$$I = \int_0^4 x\, e^{2x}\, dx \approx \frac{\Delta x}{2} [f(x_0) + f(x_1)] \approx \frac{4}{2} [f(0) + f(4)]$$

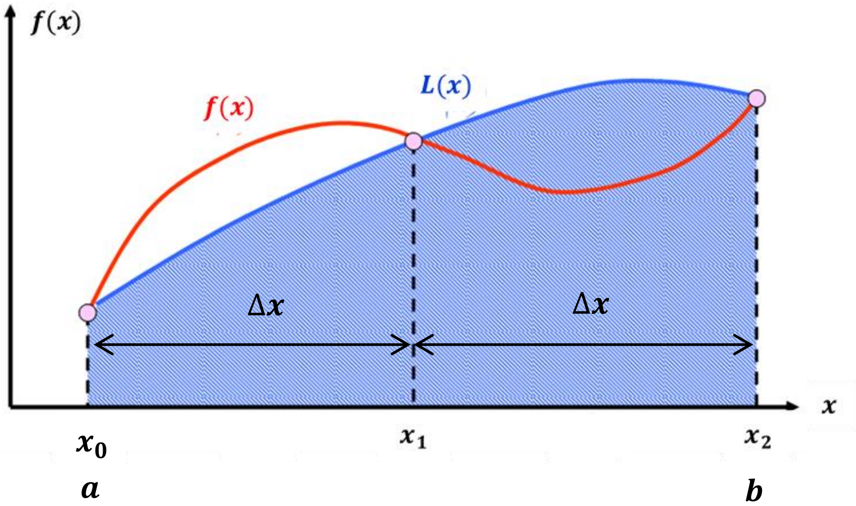
$$= 2(0 + 4e^8) = 23847.66$$

$$\varepsilon_a = \left| \frac{5216.926 - 23847.66}{5216.926} \right| \times 100\% = 357.12\%$$

✪ Simpson`s $1/3\,rd$ Rule

Approximate the **function** by a **parabola**

$$I = \frac{\Delta x}{3}[(f(x_0) + 4f(x_1) + f(x_2)]$$



Lagrange Interpolation:

$$L(x) = \frac{(x-x_1)}{(x_0-x_1)}\frac{(x-x_2)}{(x_0-x_2)}f(x_0) + \frac{(x-x_1)}{(x_1-x_0)}\frac{(x-x_2)}{(x_1-x_2)}f(x_1) + \frac{(x-x_0)}{(x_2-x_0)}\frac{(x-x_1)}{(x_2-x_1)}f(x_2)$$

Let $\quad a = x_0 \quad , \quad b = x_2 \quad , \Delta x = \dfrac{b-a}{2} = x_1 - x_0 = x_2 - x_1$

$$x_1 = \frac{a+b}{2} = x_0 + \Delta x \quad , \quad x_2 = \frac{a+b}{2} = x_0 + 2\Delta x = x_1 + \Delta x$$

$$\xi = \frac{x-x_1}{\Delta x} \quad , \quad d\xi = \frac{dx}{\Delta x}$$

$$\begin{cases} x = x_0 & \Rightarrow \quad \xi = -1 \\ x = x_1 & \Rightarrow \quad \xi = 0 \\ x = x_2 & \Rightarrow \quad \xi = 1 \end{cases}$$

$$\Rightarrow \quad L(\xi) = \frac{\xi(1-\xi)}{2}f(x_0) + (1-\xi^2)f(x_1) + \frac{\xi(\xi+1)}{2}f(x_2)$$

Integrating to **obtain** the **rule:**

$$I = \int_a^b f(x)\,dx \approx \Delta x \int_{-1}^{1} L(\xi)\,d\xi$$

$$= f(x_0)\frac{\Delta x}{2}\int_{-1}^{1} \xi(\xi - 1)\,d\xi + f(x_1)\,\Delta x \int_{-1}^{1} (1 - \xi^2)\,d\xi$$

$$+ f(x_2)\frac{\Delta x}{2}\int_{-1}^{1} \xi(\xi + 1)\,d\xi$$

$$= f(x_0)\,\frac{\Delta x}{2}\left(\frac{\xi^3}{3} - \frac{\xi^2}{2}\right)\Bigg|_{-1}^{1} + f(x_1)\,\Delta x\left(\xi - \frac{\xi^3}{3}\right)\Bigg|_{-1}^{1}$$

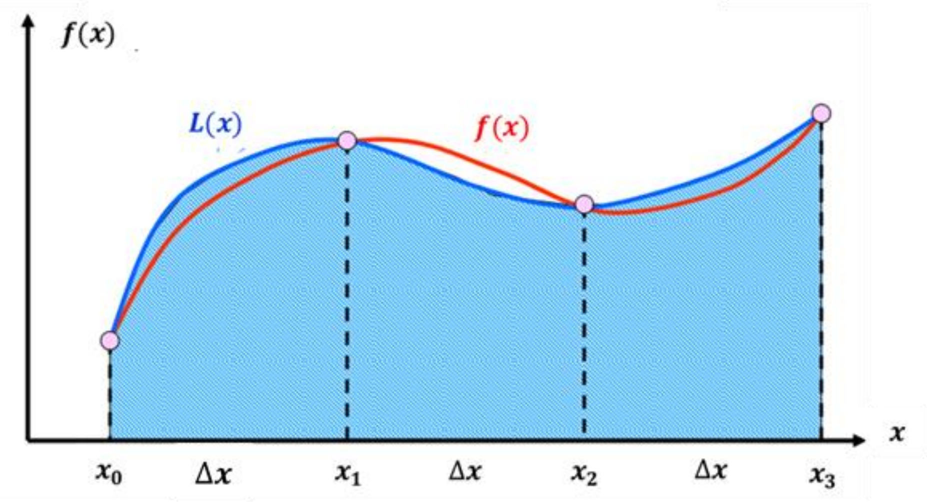
$$+ f(x_2)\frac{\Delta x}{2}\left(\frac{\xi^3}{3} + \frac{\xi^2}{2}\right)\Bigg|_{-1}^{1}$$

$$\Rightarrow \quad I = \int_a^b f(x)\,dx \approx \frac{\Delta x}{3}[f(x_0) + 4f(x_1) + f(x_2)]$$

✪ <u>Simpson\`s $3/8\,th$ Rule</u>

Approximate by a **<u>cubic polynomial</u>**

$$I \approx \frac{3}{8}\Delta x[f(x_0) + 3f(x_1) + 3f(x_2) + f(x_3)]$$



<u>Lagrange Interpolation</u>:

$$\boldsymbol{L(x)=\frac{(x-x_1)}{(x_0-x_1)}\frac{(x-x_2)}{(x_0-x_2)}\frac{(x-x_3)}{(x_0-x_3)}f(x_0)}$$
$$\boldsymbol{+\frac{(x-x_0)}{(x_1-x_0)}\frac{(x-x_2)}{(x_1-x_2)}\frac{(x-x_3)}{(x_1-x_3)}f(x_1)}$$
$$\boldsymbol{+\frac{(x-x_0)}{(x_2-x_0)}\frac{(x-x_1)}{(x_2-x_1)}\frac{(x-x_3)}{(x_2-x_3)}f(x_2)}$$
$$\boldsymbol{+\frac{(x-x_0)}{(x_3-x_0)}\frac{(x-x_1)}{(x_3-x_1)}\frac{(x-x_2)}{(x_3-x_2)}f(x_3)}$$

Integrating to **obtain** the **rule:**

$$\boldsymbol{I=\int_a^b f(x)\,dx\approx\int_a^b L(x)\,dx}\quad ,\qquad \boldsymbol{\Delta x=\frac{b-a}{3}}$$

$$\boldsymbol{\Rightarrow I=\int_a^b f(x)\,dx\approx\frac{3\Delta x}{8}[f(x_0)+3f(x_1)+3f(x_2)+f(x_3)]}$$

Note the $\boldsymbol{3/8}$ in the **formula**, and hence the **name** of **method** as the **Simpson's $\boldsymbol{3/8}$ rule**.

<u>Example</u> 2: Evaluate the following **integral** using **Simpson 1/3** and **Simpson 3/8 rule** and then **compar**e with **exact value**

$$\boldsymbol{\int_0^4 x\,e^{2x}\,dx}$$

<u>Solution</u>:

The **<u>exact solution</u>** is:

$$\boldsymbol{\int_0^4 x\,e^{2x}\,dx=\left[\frac{x}{2}e^{2x}-\frac{1}{4}\,e^{2x}\right]_0^4=\frac{1}{4}\,e^{2x}(2x-1)\Big|_0^4=5216.93}$$

- **Simpson 1/3 rule**

$$I = \int_0^4 x\, e^{2x}\, dx \approx \frac{\Delta x}{3}[f(0) + 4f(2) + f(4)]$$

$$\Delta x = \frac{4-0}{2} = 2$$

$$\Rightarrow \int_0^4 x\, e^{2x}\, dx = \frac{2}{3}[0 + 4(2e^4) + 4e^4] = 8240.41$$

$$\varepsilon = \frac{5216.93 - 8240.41}{5216.93} = -57.96\%$$

- **Simpson 3/8 rule**

$$I = \int_0^4 x\, e^{2x}\, dx \approx \frac{3\Delta x}{8}\left[f(0) + 3f\left(\frac{4}{3}\right) + 3f\left(\frac{8}{3}\right) + f(4)\right]$$

$$\Delta x = \frac{4-0}{3} = \frac{4}{3}$$

$$\Rightarrow \mathrm{I} \approx \frac{3\left(\frac{4}{3}\right)}{8}[0 + 3(19.18922) + 3(552.33933) + 11923.832]$$
$$= 6819.21$$

$$\varepsilon = \frac{5216.93 - 6819.21}{5216.93} = -30.71\%$$

Example 3: A **certain fluid** has **volume** (**0.08** m^3) is **expansion reversibly** in a **cylinder behind a piston according** to a **law** $PV = 0.25$ to **become** (**0.09** m^3), where (P) pressure in (kPa) and (V) **volume** in (m^3). **Find** by **Trapezoidal Rule work done** from

$$W = \int_{V_1}^{V_2} P dV \text{ and } n = 5.$$

Solution:

By **numerical integration:**

$$\Delta x = \frac{b-a}{n} = \frac{0.09-0.08}{5} = 0.002$$

$V_0 = 0.08$	$P_0 = f(V_0) = \frac{0.25}{V_0} = \frac{0.25}{0.08} = 3.125$
$V_1 = V_0 + 0.002 = 0.082$	$P_1 = f(V_1) = \frac{0.25}{V_1} = \frac{0.25}{0.082} = 3.0487$
$V_2 = V_1 + 0.002 = 0.084$	$P_2 = f(V_2) = \frac{0.25}{V_2} = \frac{0.25}{0.084} = 2.9762$
$V_3 = V_2 + 0.002 = 0.086$	$P_3 = f(V_3) = \frac{0.25}{V_3} = \frac{0.25}{0.086} = 2.907$
$V_4 = V_3 + 0.002 = 0.088$	$P_4 = f(V_4) = \frac{0.25}{V_4} = \frac{0.25}{0.088} = 2.841$
$V_5 = V_4 + 0.002 = 0.09$	$P_5 = f(V_5) = \frac{0.25}{V_5} = \frac{0.25}{0.09} = 2.7777$

$$I = W = \int_{V_1}^{V_2} P dV \approx \frac{\Delta x}{2}[P_0 + 2(P_1 + P_2 + P_3 + P_4) + P_5]$$

$$\approx \frac{0.002}{2}[3.125 + 2(3.0487 + 2.9762 + 2.907 + 2.841) + 2.7777]$$

$$\Rightarrow W = 0.02944\ kJ$$

The **exact solution** is:

$$W = \int_{V_1}^{V_2} P dV = \int_{0.08}^{0.09} \frac{0.25}{V} dV = 0.25[lnV]_{0.08}^{0.09} = 0.02944\ \ kJ$$

$$\varepsilon = \frac{0.02944 - 0.02944}{0.02944} = 0\%$$

Example 4: The **vertical distance covered** by a **rocket** from $x = 8$ to $x = 30$ **seconds** is given by

$$s = \int_{8}^{30} \left[2000\ ln\left(\frac{140000}{140000 - 2100\ x}\right) - 9.8\ x\right] dx$$

Use **Simpson 3/8 rule** to **find** the **approximate value** of the **integral**.

Solution:

$$\Delta x = \frac{b-a}{n} = \frac{30-8}{3} = 7.3333$$

$$I \approx \frac{3}{8}\Delta x[f(x_0) + 3f(x_1) + 3f(x_2) + f(x_3)]$$

$$x_0 = a = 8$$

$$f(x_0) = 2000\ ln\left(\frac{140000}{140000 - 2100 \times 8}\right) - 9.8\ \times 8 = 177.2667$$

$$x_1 = x_0 + \Delta x = 8 + 7.3333 = 15.3333$$

$$f(x_1) = 2000\ ln\left(\frac{140000}{140000 - 2100 \times 15.3333}\right) - 9.8\ \times 15.3333$$
$$= 372.4629$$

$$x_2 = x_0 + 2\Delta x = 8 + 2(7.3333) = 22.6666$$

$$f(x_2) = 2000\ ln\left(\frac{140000}{140000 - 2100 \times 22.6666}\right) - 9.8\ \times 22.6666$$
$$= 608.8976$$

$$x_3 = x_0 + 3\Delta x = 8 + 3(7.3333) = 30$$

$$f(x_3) = 2000\ ln\left(\frac{140000}{140000 - 2100 \times 30}\right) - 9.8\ \times 30 = 901.6740$$

Substituting these values into the f**ormula** for **Simpson 3/8 rule**, yields

$$I \approx \frac{3}{8} \times 7.3333 \times (177.2667 + 3 \times 372.4629 + 3 \times 608.8976$$
$$+901.6740) \approx 11063.31$$

The **exact answer** can be computed as

$$I_{exact} = 11061.34$$
$$\varepsilon = \frac{11061.34 - 11063.31}{11061.34} = -178.098\%$$

✪ The Boole`s Rule

Boole's rule is a **method** of **numerical integration** to **approximates an integral**

$$\int_{x_0}^{x_4} f(x)\, dx$$

by **using** the **values** of the **function** $f(x)$ at **five equally spaced points**

x_0 .

$x_1 = x_0 + \Delta x$.

$x_2 = x_0 + 2\Delta x$.

$x_3 = x_0 + 3\Delta x$.

$x_4 = x_0 + 4\Delta x$

It expressed as

$$\int_{x_0}^{x_4} f(x)\, dx \approx \frac{2\Delta x}{45}[7\, f(x_0) + 32f(x_1) + 12f(x_2) + 32f(x_3) + 7f(x_4)]$$

it is **applicable only** when the **number** of **subintervals** is a **multiple** of 4 $(n = 4.8.12.\cdots)$. Thus **dividing** the **range** into **four equal parts.**

Example 5: Apply the **various quadrature formulas** to **evaluate** the **integration** of the **function**

$$f(x) = 1 + e^{-x}sin(4x)$$

over the **interval** $[0.2]$ with **equally spaced quadrature nodes**

$$x_0 = 0\,, \quad x_1 = 0.5\,, \quad x_2 = 1\,, \quad x_3 = 1.5\,, \quad x_4 = 2$$

Solution:

The **step size** is:

$$\Delta x = x_1 - x_0 = 0.5 - 0 = 0.5$$

and the **corresponding function values** are:

$$f_0 = f(0) = 1 + e^{-0}sin(4 \times 0) = 1$$

$$f_1 = f(0.5) = 1 + e^{-0.5}sin(4 \times 0.5) = 1.55152$$

$$f_2 = f(1) = 1 + e^{-1}sin(4 \times 1) = 0.72159$$

$$f_3 = f(1.5) = 1 + e^{-1.5}sin(4 \times 1.5) = 0.93765$$

$$f_4 = f(2) = 1 + e^{-2}sin(4 \times 2) = 1.13390$$

- **<u>The Trapezoidal Rule</u>**

$$\int_{x_0}^{x_1} f(x)dx = \int_0^{0.5} f(x)dx \approx \frac{\Delta x}{2}(f_0 + f_1) = \frac{0.5}{2}(1 + 1.55152)$$
$$\approx 0.63788$$

- **<u>The Simpson`s 1/3 Rule</u>**

$$\int_{x_0}^{x_2} f(x)dx = \int_0^{1} f(x)dx \approx \frac{\Delta x}{3}(f_0 + 4f_1 + f_2)$$
$$= \frac{0.5}{3}[1 + 4(1.55152) + 0.72159] = 1.32128$$

- **<u>The Simpson`s 3/8 Rule</u>**

$$\int_{x_0}^{x_3} f(x)dx = \int_0^{1.5} f(x)dx \approx \frac{3\Delta x}{8}(f_0 + 3f_1 + 3f_2 + f_3)$$
$$= \frac{3(0.5)}{8}(1 + 3(1.55152) + 3(0.72159) + 0.93765)$$
$$= 1.64193$$

- **<u>The Boole`s Rule</u>**

$$\int_{x_0}^{x_4} f(x)dx = \int_0^{2} f(x)dx \approx \frac{2\Delta x}{45}(7f_0 + 32f_1 + 12f_2 + 32f_3 + 7f_4)$$
$$= \frac{2(0.5)}{45}[7(1) + 32(1.55152) + 12(0.72159) + 32(0.93765) + 7(1.1339)] = 2.29444$$

In this **example** the **quadrature rules** with $\Delta x = 0.5$ have been **applied**.

If the **endpoints** of the **interval [*a*, *b*]** are **<u>held fixed,</u>** the **step size must be adjusted** for **each rule** as follows:

$$\Delta x = h = b - a \qquad \textbf{For trapezoidal rule}$$

$$\Delta x = h = \frac{b-a}{2} \qquad \textbf{For Simpson`s 1/3 rule}$$

$$\Delta x = h = \frac{b-a}{3} \qquad \textbf{For Simpson`s } \frac{3}{8} \textbf{ rule}, \text{and}$$

$$\Delta x = h = \frac{b-a}{4} \qquad \textbf{For Boole`s rule}$$

The **next example illustrates this point**.

<u>Example</u> 6: Evaluate the **integration** of the **function**

$$f(x) = 1 + e^{-x} sin(4x)$$

over the **interval** $[a, b] = [0, 1]$ by applying the **various quadrature formulas**.

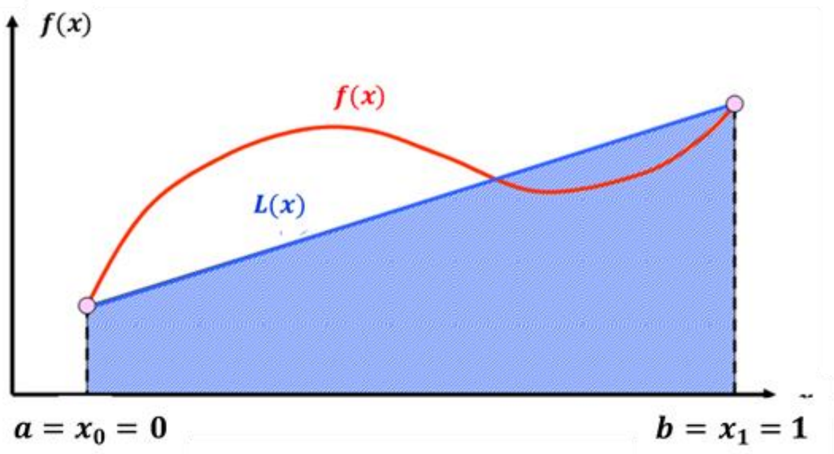
<u>Solution</u>:

➢ For the **<u>trapezoidal rule</u>**, where

$$a = x_0 = 0 \quad , \qquad b = x_1 = 1$$

The **step size** is:

$$\Delta x = b - a = 1 - 0 = 1$$



$$\int_{x_0}^{x_1} f(x)dx = \int_0^1 f(x)dx \approx \frac{\Delta x}{2}(f_0 + f_1)$$

The **values** of the **corresponding function** are:

$$f_0 = f(0) = 1 + e^{-0} sin(4 \times 0) = 1$$

$$f_1 = f(1) = 1 + e^{-1} sin(4 \times 1) = 0.72159$$

Thus,

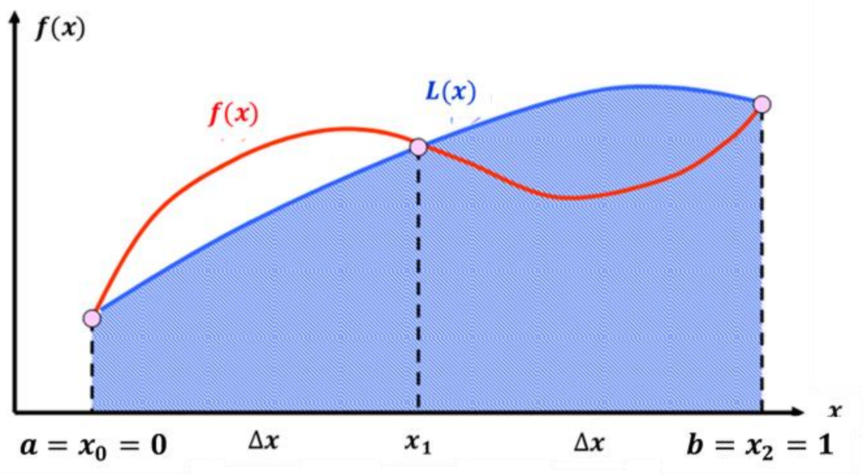
$$\int_0^1 f(x)dx \approx \frac{\Delta x}{2}(f_0 + f_1) = \frac{1}{2}(1 + 1.55152) \approx 0.86079$$

- For the **Simpson`s 1/3 rule**, where

$$a = x_0 \quad , \quad b = x_2 \quad , \quad x_1 = \frac{b-a}{2}$$

The **step size** is:

$$\Delta x = \frac{b-a}{2} = \frac{1-0}{2} = 0.5$$



$$\int_{x_0}^{x_2} f(x)dx = \int_0^1 f(x)dx \approx \frac{\Delta x}{3}(f_0 + 4f_1 + f_2)$$

The **values** of the **corresponding function** are:

$$f_0 = f(0) = 1 + e^{-0} sin(4 \times 0) = 1$$

$$f_1 = f(0.5) = 1 + e^{-0.5} sin(4 \times 0.5) = 1.55152$$

$$f_2 = f(1) = 1 + e^{-1} sin(4 \times 1) = 0.72159$$

Thus,

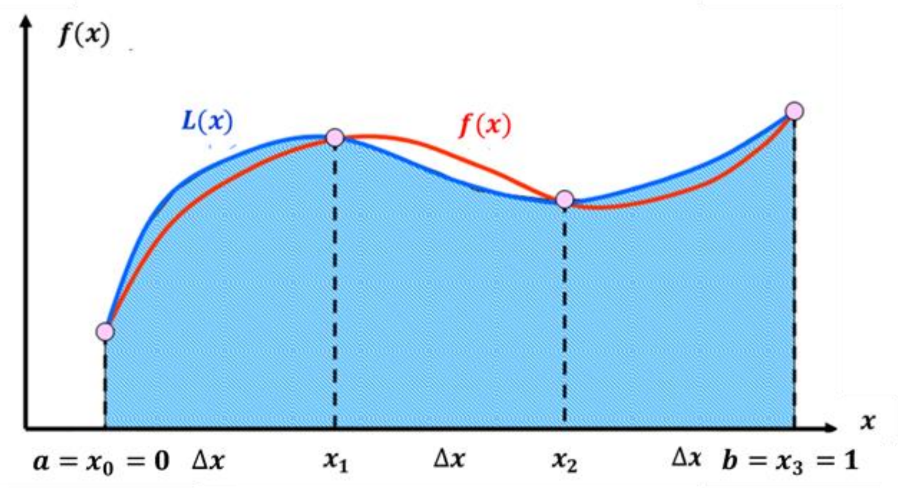
$$\int_{x_0}^{x_2} f(x)dx = \int_0^1 f(x)dx \approx \frac{\Delta x}{3}(f_0 + 4f_1 + f_2)$$
$$= \frac{0.5}{3}[1 + 4(1.55152) + 0.72159] = 1.32128$$

- For the **Simpson`s 3/8 rule**, where

$$a = x_0 = 0 \quad , \qquad b = x_3 = 1$$

The **step size** is:

$$\Delta x = \frac{b-a}{3} = \frac{1-0}{3} = \frac{1}{3}$$



$$\int_{x_0}^{x_3} f(x)dx = \int_0^1 f(x)dx \approx \frac{3\Delta x}{8}(f_0 + 3f_1 + 3f_2 + f_3)$$

The **values** of the **corresponding function** are:

$$f_0 = f(0) = 1 + e^{-0}sin(4 \times 0) = 1$$
$$f_1 = f(1/3) = 1 + e^{-1/3}sin(4 \times 1/3) = 1.69642$$
$$f_2 = f(2/3) = 1 + e^{-2/3}sin(4 \times 2/3) = 1.23447$$
$$f_3 = f(1) = 1 + e^{-1}sin(4 \times 1) = 0.72159$$

Thus,

$$\int_{x_0}^{x_3} f(x)dx = \int_0^1 f(x)dx \approx \frac{3\Delta x}{8}(f_0 + 3f_1 + 3f_2 + f_3)$$

$$= \frac{3(1/3)}{8}(1 + 3(1.69642) + 3(1.23447) + 0.72159)$$

$$= 1.3144$$

➢ For the **<u>Boole`s rule</u>**

Since this rule is **applicable only** when the **number** of **subintervals** is a **multiple** of **4** ($n = 4, 8, 12, \cdots$). Thus **dividing** the **range** $[0, 1]$ into **<u>four equal parts</u> by taking** $\Delta x = 1/4$

$$\int_{x_0}^{x_4} f(x)dx = \int_0^1 f(x)dx \approx \frac{2\Delta x}{45}(7f_0 + 32f_1 + 12f_2 + 32f_3 + 7f_4)$$

The **corresponding function values** are:

$$f_0 = f(0) = 1 + e^{-0} sin(4 \times 0) = 1$$

$$f_1 = f(1/4) = 1 + e^{-1/4} sin(4 \times 1/4) = 1.65534$$

$$f_2 = f(1/2) = 1 + e^{-1/2} sin(4 \times 1/2) = 1.55152$$

$$f_3 = f(3/4) = 1 + e^{-3/4} sin(4 \times 3/4) = 1.06666$$

$$f_4 = f(1) = 1 + e^{-1} sin(4 \times 1) = 0.72159$$

$$\Rightarrow \int_0^1 f(x)dx \approx \frac{2(1/4)}{45}[7(1) + 32(1.65534) + 12(1.55152)$$

$$+32(1.06666) + 7(0.72159)] = 1.30859$$

The **true value** of the **definite integral** is:

$$\int_0^1 f(x)dx = \frac{21\, e - 4\, cos(4) - sin(4)}{17\, e} = 1.308251$$

The **approximation** from **Boole's rule** is the **<u>best</u>**,

➔ Composite Integration

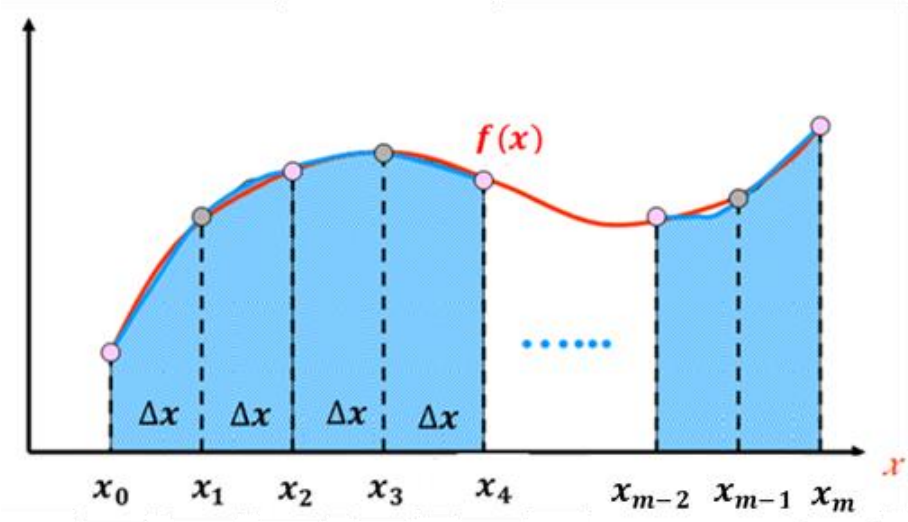
To **make a fair comparison** of **quadrature methods**, the **same number** of **function evaluations** in **each method must be used**. The **final example** is **concerned** with **comparing integration** over a **fixed interval** $[\boldsymbol{a}, \boldsymbol{b}]$ using **exactly five function evaluation** $\boldsymbol{f_k = f(x_k)}$, for $\boldsymbol{k = 0, 1, 2, 3, 4}$ for **each method**.

✪ Composite Trapezoidal Rule

The **trapezoidal rule** when is **applied on** the **many subintervals**

$$[x_0, x_1] \ , \ [x_1, x_2] \ , \quad [x_2, x_3], \cdots, [x_{m-2}, x_{m-1}], \ [x_{m-1}, x_m]$$

it is **called** a **composite trapezoidal rule.**



The **composite trapezoidal rule** for $\boldsymbol{n}$ **subintervals** can be **expressed as**:

$$\int_a^b f(x)dx \approx \frac{\Delta x}{2}[f(x_0) + 2\{f(x_1) + f(x_i) + \cdots + f(x_{m-1})\} + f(x_m)]$$

Proof:

Suppose that the **interval** $[\boldsymbol{a}, \boldsymbol{b}]$ is **subdivided** into **subinterval** $[x_k, x_{k+1}]$ of **width**

$$\Delta x = h = \frac{b-a}{n} \quad \text{where } n = \textbf{number of subintervals}$$

by **using <u>equally spaced nodes</u>**

$$x_k = a + k\Delta x \quad \text{for} \quad k = 0, 1, 2, 3, \cdots, m$$

Applying the **trapezoidal rule over each subinterval** and **use** the **additive property** of the **integral** for **subintervals**:

$$\int_a^b f(x)dx = \int_{x_0}^{x_1} f(x)\, dx + \int_{x_1}^{x_2} f(x)dx + \cdots + \int_{x_{m-1}}^{x_m} f(x)dx$$

$$\approx \frac{\Delta x}{2}[f(x_0) + f(x_1)] + \frac{\Delta x}{2}[f(x_1) + f(x_2)] + \cdots$$

$$+ \frac{\Delta x}{2}[f(x_{m-1}) + f(x_m)]$$

$$\approx \frac{\Delta x}{2}[f(x_0) + 2f(x_1) + 2f(x_i) + \cdots + 2f(x_{m-1}) + f(x_m)]$$

$$\approx \frac{\Delta x}{2}[f(x_0) + 2\{f(x_1) + f(x_i) + \cdots + f(x_{m-1})\} + f(x_m)]$$

<u>**Example**</u> **7: Use** the <u>**composite trapezoid rule**</u> to **evaluate** the **integral**

$$\int_0^4 x\, e^{2x}\, dx$$

<u>**Solution**</u>**:**

- $n = 1$, $\Delta x = \frac{b-a}{n} = \frac{4-0}{1} = 4$

$$\Rightarrow I \approx \frac{\Delta x}{2}[f(0) + f(4)] = 23847.66$$

$$\varepsilon = -357.12\%$$

- $n = 2$, $\Delta x = \frac{b-a}{n} = \frac{4-0}{2} = 2$

$$\Rightarrow I \approx \frac{\Delta x}{2}[f(0) + 2f(2) + f(4)] = 12142.23$$

$$\varepsilon = -132.75\%$$

- $n = 4$, $\Delta x = \frac{b-a}{n} = \frac{4-0}{4} = 1$

$$\Rightarrow I \approx \frac{\Delta x}{2}[f(0) + 2f(1) + 2f(2) + 2f(3) + f(4)] = 7288.79$$

$$\varepsilon = -39.17\%$$

- $n = 8$, $\Delta x = \frac{b-a}{n} = \frac{4-0}{8} = 0.5$

$$\Rightarrow I \approx \frac{\Delta x}{2}[f(0) + 2f(0.5) + 2f(1) + 2f(1.5) + 2f(2) + 2f(2.5) + 2f(3) + 2f(3.5) + f(4)] = 5764.76$$

$$\varepsilon = -10.50\%$$

- $n = 16$, $\Delta x = \frac{b-a}{n} = \frac{4-0}{16} = 0.25$

$$\Rightarrow I \approx \frac{\Delta x}{2}[f(0) + 2f(0.25) + 2f(0.5) + \cdots + 2f(3.5) + f(4)] = 5355.95$$

$$\varepsilon = -2.66\%$$

Example 8: Use the **composite trapezoidal rule** with **11 sample points** to compute an **approximation** to the **integral** of the function

$$f(x) = 2 + sin\left(2\sqrt{x}\right)$$

taken over [1,6].

Solution:

To **generate 11 sample points**, we use $n = 10$ and

$$\Delta x = h = \frac{b-a}{n} = \frac{6-1}{10} = 0.5$$

x	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6
$f(x)$	2.91	2.64	2.31	1.98	1.68	1.44	1.24	1,11	1.03	1.00	1.02

$$\int_1^6 f(x)dx \approx \frac{\Delta x}{2}[f(x_0) + 2\{f(x_1) + f(x_2) + \cdots + f(x_5)\} + f(x_6)]$$

$$\approx \frac{\Delta x}{2}[f(1) + 2\{f(1.5) + f(2) + f(2.5) + f(3) + f(3.5) + f(4) + f(4.5) + f(5) + f(5.5\} + f(x_6)] = 8.19386$$

❖ <u>Composite Trapezoid Rule with Unequal Segments</u>

<u>Example</u> 9: Use the **<u>composite trapezoid rule</u>** to **evaluate** the **integral**

$$\int_0^4 x\, e^{2x}\, dx$$

with $(\Delta x)_1 = 2$, $(\Delta x)_2 = 1$, $(\Delta x)_3 = 0.5$, $(\Delta x)_4 = 0.5$

<u>Solution</u>:

$$I = \int_0^2 f(x)dx + \int_2^3 f(x)\, dx + \int_3^{3.5} f(x)dx + \int_{3.5}^4 f(x)dx$$

$$\approx \frac{(\Delta x)_1}{2}[f(0) + f(2)] + \frac{(\Delta x)_2}{2}[f(2) + f(3)]$$

$$+ \frac{(\Delta x)_3}{2}[f(3) + f(3.5)] + \frac{(\Delta x)_4}{2}[f(3.5) + f(4)]$$

$$\approx \frac{2}{2}[0 + 2\, e^4] + \frac{1}{2}[2\, e^4 + 3\, e^6] + \frac{0.5}{2}[3\, e^6 + 3.5\, e^7]$$

$$+ \frac{0.5}{2}[3.5\, e^7 + 4\, e^8] = 5971.58$$

$$\varepsilon = \frac{5216.93 - 5971.58}{5216.93} = -14.47\%$$

✪ <u>Composite Simpson's Rule</u>

<u>Simpson's rule</u> can also be **used** in **same manner as for trapezoidal rule** when is **applied on** the **many subintervals** it is **called** a **<u>composite Simpson rule</u>.**

Suppose that $[a, b]$ is **subdivided** into n **<u>subintervals</u>** $[x_k\,, x_{k+1}]$ of **equal width**

$$\Delta x = \frac{b - a}{n}$$

by **using**

$$x_k = a + i\Delta x \quad \text{for} \quad i = 0, 1, 2, 3, \cdots, 2$$

The **composite Simpson rule** for $2m$ **subintervals** can **be expressed** as:

$$\int_a^b f(x)dx \approx \frac{\Delta x}{3}[f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + \cdots$$
$$+ 2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)]$$
$$\approx \frac{\Delta x}{3}\left[f(a) + 4\sum_{\substack{i=1\\ i=odd}}^{n-1} f(x_i) + 2\sum_{\substack{i=2\\ i=even}}^{n-2} f(x_i) + f(b)\right]$$

When **Simpson's rule** is **applied on** the **two subintervals** $[x_0, x_2]$ and $[x_2, x_4]$, can be **expressed** by

$$\int_{x_0}^{x_4} f(x)dx = \int_{x_0}^{x_2} f(x)dx + \int_{x_2}^{x_4} f(x)dx$$
$$\approx \frac{\Delta x}{3}(f_0 + 4f_1 + f_2) + \frac{\Delta x}{3}(f_2 + 4f_3 + f_4)$$
$$\approx \frac{\Delta x}{3}(f_0 + 4f_1 + 2f_2 + 4f_3 + f_4)$$

Example 10: Use the **composite Simpson`s rule** with **11 sample points** to **compute** an **approximation** to the **integral** of the **function**

$$f(x) = 2 + sin\left(2\sqrt{x}\right)$$

over the **interval [1,6]**.

Solution:

To **generate 11 sample points**, use $n = 10$ and

$$\Delta x = \frac{b-a}{n} = \frac{6-1}{10} = 0.5$$

x	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6
$f(x)$	2.91	2.64	2.31	1.98	1.68	1.44	1.24	1,11	1.03	1.00	1.02

$$\int_a^b f(x)dx \approx \frac{\Delta x}{3}[f(1) + 4f(1.5) + 2f(2) + 4f(2.5) + 2f(3)$$
$$+ 4f(3.5) + 2f(4) + 4f(4.5) + 2f(5) + 4f(5.5)$$
$$+ f(6)]$$

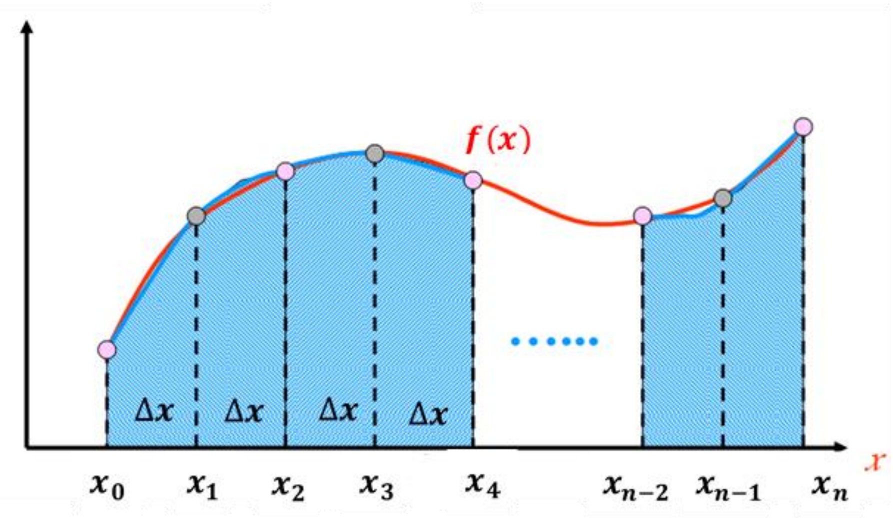
or $$\int_a^b f(x)dx \approx \frac{\Delta x}{3}\left[f(a) + 4\sum_{i=1}^{n-1} f(x_i) + 2\sum_{i=2}^{n-2} f(x_i) + f(b)\right]$$

Thus,

$$\int_1^6 f(x)dx \approx \frac{0.5}{3}[f(1) + 4\{f(1.5) + f(2.5) + f(3.5) + f(4.5) + f(5.5)\} + 2\{f(2) + f(3) + f(4) + f(5)\} + f(6)] = 8.18302$$

➢ Piecewise Quadratic approximations

$$\Delta x = \frac{b-a}{n} \quad , \qquad n = \text{number of subintervals}$$



➢ Multiple applications of Simpson's rule

$$\int_a^b f(x)dx = \int_{x_0}^{x_2} f(x)\,dx + \int_{x_2}^{x_4} f(x)dx + \cdots + \int_{x_{n-2}}^{x_n} f(x)dx$$

$$\approx \frac{\Delta x}{3}[f(x_0) + 4f(x_1) + f(x_2)] + \frac{\Delta x}{3}[f(x_2) + 4f(x_3) + f(x_4)] + \cdots + \frac{\Delta x}{3}[f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)]$$

$$\approx \frac{\Delta x}{3}[f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + 2f(x_4)$$
$$+ \cdots + 4f(x_{2i-1}) + 2f(x_{2i}) + 4f(x_{2i+1})$$
$$+ \cdots + 2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)]$$

<u>Example</u> 11: Use <u>composite Simpson`s rule</u> to **evaluate** the **integral**

$$\int_0^4 x\, e^{2x}\, dx$$

with $n = 2$, $\Delta x = 2$ and $n = 4$, $\Delta x = 1$

<u>Solution</u>:

- $n = 2$, $\Delta x = 2$

$$I \approx \frac{\Delta x}{3}[f(0) + 4f(2) + f(4)] = \frac{2}{3}[0 + 4(2\, e^4) + 4e^8)]$$
$$= 8240.41$$
$$\varepsilon = \frac{5216.93 - 8240.41}{5216.93} = -57.96\%$$

- $n = 4$, $\Delta x = 1$

$$I \approx \frac{\Delta x}{3}[f(0) + 4f(1) + 2f(2) + 4f(3) + f(4)]$$
$$= \frac{2}{3}[0 + 4(\, e^2) + 2\,(2\, e^4) + 4\,(3\, e^6) + 4\, e^8)]$$
$$= 5670.98$$
$$\varepsilon = \frac{5216.93 - 5670.98}{5216.93} = -8.70\%$$

❖ Composite Simpson's Rule with Unequal Segments

Example 12: Use the **composite trapezoid rule** to **evaluate** the **integral**

$$\int_{0}^{4} x\, e^{2x}\, dx$$

with $(\Delta x)_1 = 1.5$, $(\Delta x)_2 = 0.5$

Solution:

$$I = \int_0^3 f(x)dx + \int_3^4 f(x)\, dx$$

$$\approx \frac{(\Delta x)_1}{3}[f(0) + 4\, f(1.5) + 2\, f(3)]$$

$$+ \frac{(\Delta x)_2}{3}[f(3) + 4f(3.5) + 2\, f(4)]$$

$$= \frac{1.5}{3}[0 + 4\, f(1.5\, e^3) + 3\, e^6] + \frac{0.5}{3}[3\, e^6 + 4\,(3.5\, e^7) + 4\, e^8]$$

$$= 5413.23$$

$$\varepsilon = \frac{5216.93 - 5413.23}{5216.93} = -3.76\%$$

Example 13: Evaluate the **integration** of the **function**

$$f(x) = 1 + e^{-x} sin(4x)$$

over the **interval** $[a, b] = [0, 1]$.**Use exactly five function evaluations** and **compare** the **results from** the **composite trapezoidal rule** and **composite Simpson's rule**.

Solution:

Since it is **required five function evaluations,** the **uniform step size** has to be:

$$\Delta x = 1/4.$$

The **composite trapezoidal rule** produces:

$$\int_0^1 f(x)dx = \int_0^{1/4} f(x)dx$$

$$+ \int_{1/4}^{1/2} f(x)dx + \int_{1/2}^{3/4} f(x)dx + \int_{3/4}^{1} f(x)dx$$

$$\approx \frac{\Delta x}{2}(f_0 + 2f_1 + 2f_2 + 2f_3 + f_4)$$

$$= \frac{1/4}{2}\left[f_0 + 2f\left(\frac{1}{4}\right) + 2f\left(\frac{1}{2}\right) + 2f\left(\frac{3}{4}\right) + f(1)\right]$$

The **corresponding function values** are:

$$f_0 = f(0) = 1 + e^{-0}sin(4 \times 0) = 1$$
$$f_1 = f(1/4) = 1 + e^{-1/4}sin(4 \times 1/4) = 1.65534$$
$$f_2 = f(1/2) = 1 + e^{-1/2}sin(4 \times 1/2) = 1.55152$$
$$f_3 = f(3/4) = 1 + e^{-3/4}sin(4 \times 3/4) = 1.06666$$
$$f_4 = f(1) = 1 + e^{-1}sin(4 \times 1) = 0.72159$$

$$\int_0^1 f(x)dx = \frac{1}{8}[1 + 2(1.65534) + 2(1.55152) + 2(1.06666) + 0.72159] = 1.28358$$

- **Using** the **<u>composite Simpson's rule</u>** yields:

$$\int_0^1 f(x)dx \approx \frac{\Delta x}{3}(f_0 + 4f_1 + 2f_2 + 4f_3 + f_4)$$

$$= \frac{1/4}{3}\left[f_0 + 4f\left(\frac{1}{4}\right) + 2f\left(\frac{1}{2}\right) + 4f\left(\frac{3}{4}\right) + f(1)\right]$$

$$= \frac{1}{12}[1 + 4(1.65534) + 2(1.55152) + 4(1.06666) + 0.72159] = 1.30938$$

❖ Combined Simpson's–1/3 & Simpson's–3/8

Example 14: Integrate the **data** listed in the **table** by **using Simpson`s 1/3 rule** and **Simpson`s 3/8 rules**.

i	0	1	2	3	4	5	6	7
x	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7
$f(x)$	1.543	1.669	1.811	1.971	2.151	2.352	2.577	2.828

Solution:

8 abscissas (values of x) ⇒ **n = 7 subintervals**. So

- **Simpson's–1/3 rule cannot be used alone** (***n*** is **not divisible by 2**) or
- **Simpson's–3/8 rule cannot be used alone** (**n** is **not divisible by 3**).

However, in this problem **can combine** the **methods** by **appropriately dividing** the **interval**:

1. **Using Simpson's–1/3 rule** on **interval [1.0, 1.4]** (**4 subintervals** is **divisible** by **2**), and
2. **Using Simpson's–3/8 rule** on **interval [1.4, 1.7]** (**3 subintervals** is **divisible** by **3**).

This way **consistent accuracy** can be **obtained** on the **entire interval [1.0, 1.7].**

➢ **Using Simpson's–1/3 rule** on **interval** $[\mathbf{1.0}, \mathbf{1.4}]$,

$(\mathbf{1.0} \to \mathbf{1.2}$ and $\mathbf{1.2} \to \mathbf{1.4})$ we have 4 subintervals i.e. $\boldsymbol{n = 4}$

$$\Rightarrow \Delta x = \frac{b-a}{n} = \frac{1.4 - 1.0}{4} = \frac{0.4}{4} = 0.1$$

$$\int_{1.0}^{1.4} f(x)dx \approx \frac{\Delta x}{3}[f(0) + 4f(1) + 2f(2) + 4f(3) + f(4)]$$

$$= \frac{0.1}{3}[(1.543) + 4(1.669) + 2(1.811) + 4(1.971) + 2.151]$$

$$= 0.7292$$

➢ **Using Simpson's–3/8 rule** on **interval** $[\mathbf{1.4}, \mathbf{1.7}]$, $\Delta x = \mathbf{0.1}$

$$\Rightarrow \Delta x = \frac{b-a}{n} = \frac{1.7 - 1.4}{3} = \frac{0.3}{3} = 0.1$$

$$\int_{1.4}^{1.7} f(x)dx \approx \frac{3\Delta x}{8}[f(4) + 3f(5) + 3f(6) + f(7)]$$

$$= \frac{3(0.1)}{8}[(2.151) + 3(2.352) + 3(2.577) + 2.828]$$

$$= 0.74123$$

Now **add** the **results**

$$\int_{1.0}^{1.7} f(x)dx = \int_{1.0}^{1.4} f(x)dx + \int_{1.4}^{1.7} f(x)dx$$

$$= 0.7292 + 0.74123 = 1.47043$$

NOTE: Alternatively, **Simpson's–3/8 rule could be used** on interval **[1.0, 1.3]** and **Simpson's–1/3 rule** on **interval [1.3, 1.7]** to **obtain** a **different approximation**: **1.47044**.

BE CAREFUL:

- Could use
  - **Simpson's–1/3 rule** on **interval [1.0, 1.6]** (**6 subintervals** is **divisible by 2**) and
  - the **Trapezoidal rule** on **interval [1.6, 1.7],** but why wouldn't we?
- Could also
  - use **Simpson's–3/8 rule** on **interval [1.0, 1.6]** (**6 subintervals** is **divisible by 3**) and
  - the **Trapezoidal rule on interval [1.6, 1.7]**, but why wouldn't?

➲ If **Even number** of **Intervals** are there, it is **preferred** to use **Simpson's $\frac{1}{3}rd$ Rule** (or) **Trapezoidal Rule**.

➲ If **Number** of **Intervals** is **multiple** of **3**, then **use Simpson's $\frac{3}{8}th$ Rule** (or) **Trapezoidal Rule**.

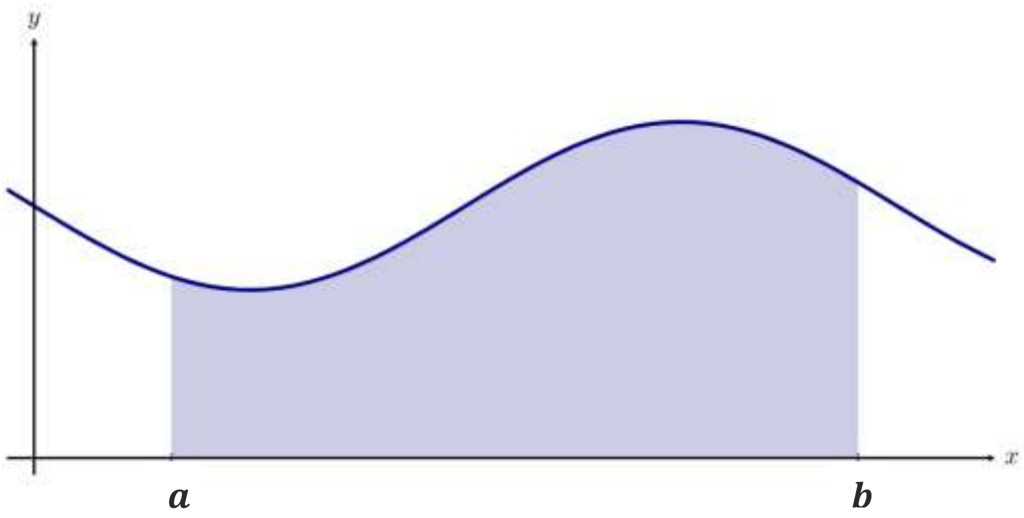
➲ If **Odd number** of **Intervals** are there, and which **is not multiple** of **3**, then **use Trapezoidal Rule**.

➲ For **any number** of **Intervals**, the **default Rule** we can **use** is **Trapezoidal**

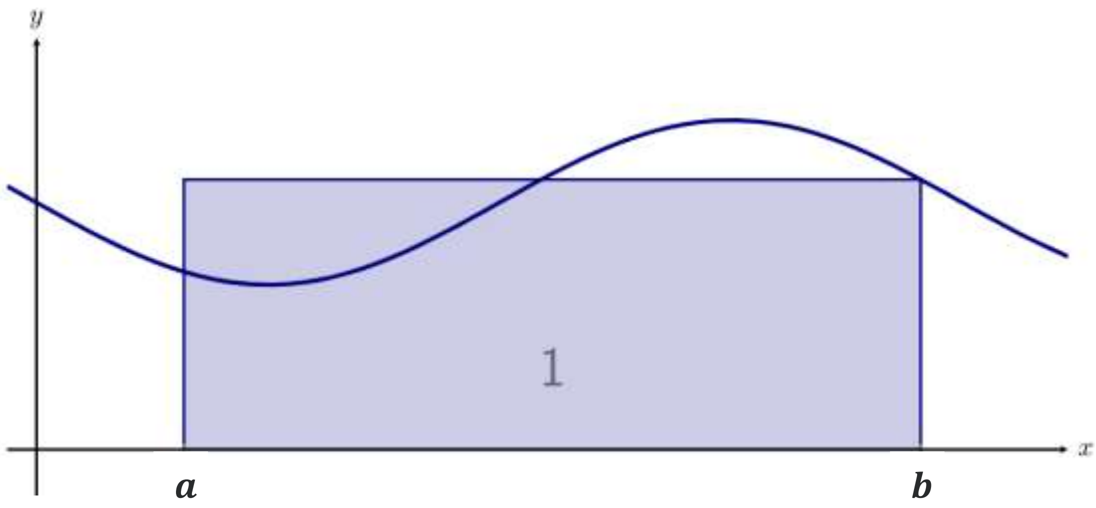
Approximating Definite Integrals Using Rectangles

➢ **Rectangles and areas**

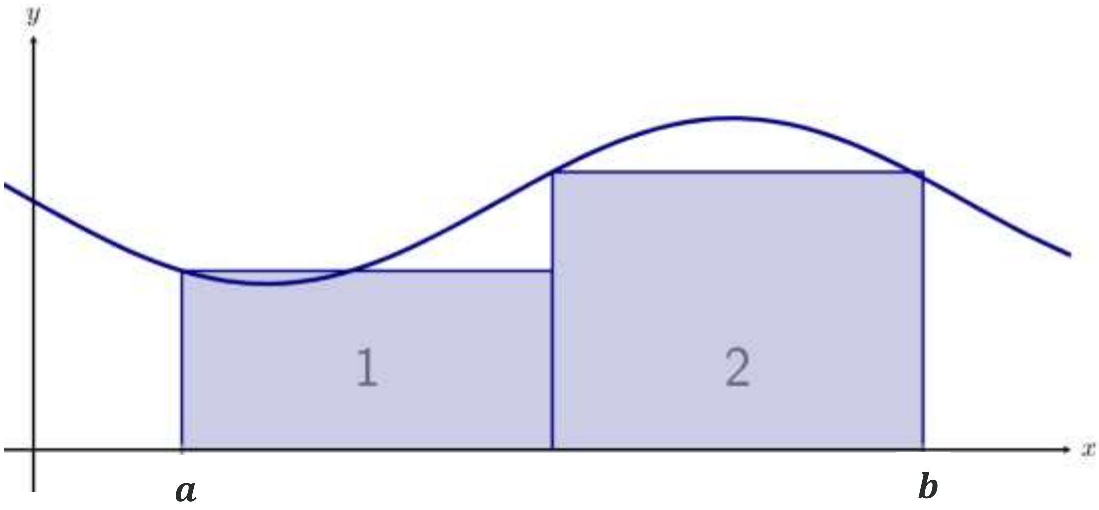
To compute the area between the curve $y = f(x)$ and the horizontal axis on the interval $[a, b]$:



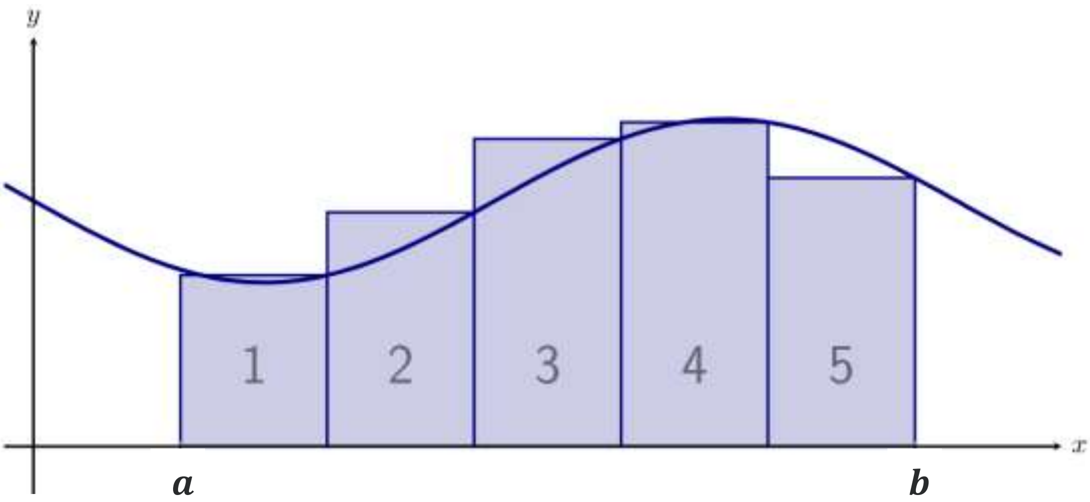
One way to do this would be to approximate the area with rectangles. With one rectangle a rough approximation will gotten:



Two rectangles might make a better approximation:

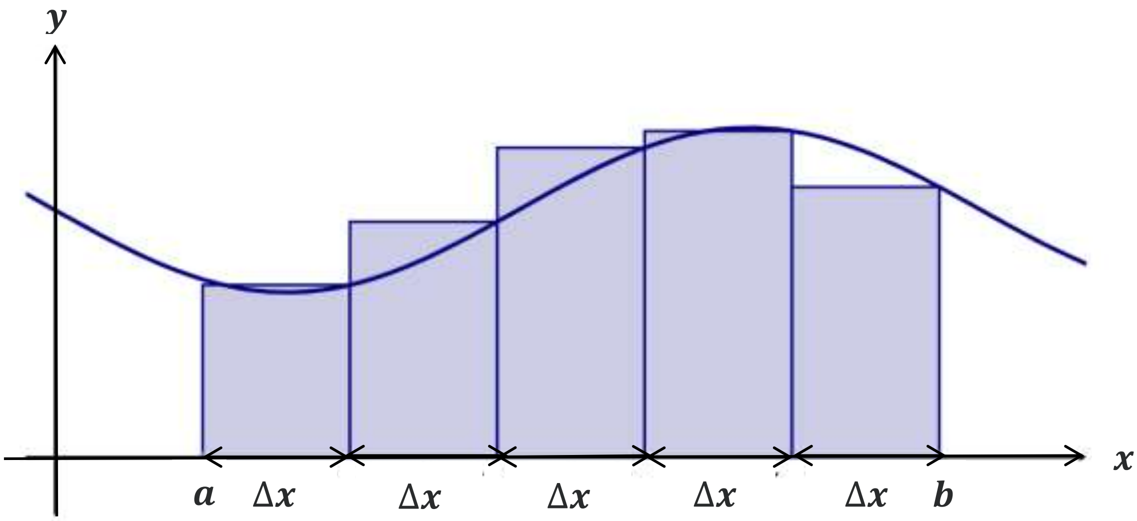


With even more, a closer, and closer, approximation will be gotten:



<u>Definition</u>: Approximating the **area between** a **curve** and the $\boldsymbol{x}$ **– axis** on the **interval** $[\boldsymbol{a}, \boldsymbol{b}]$:with $\boldsymbol{n}$ **rectangles** of width $\Delta\boldsymbol{x}$, then

$$\Delta\boldsymbol{x} = \frac{\boldsymbol{b}-\boldsymbol{a}}{\boldsymbol{n}}$$



Q/ Suppose is wanted to approximate the area between the curve

$$y = x^2 + 1$$

and the x – **axis** on the interval $[-1, 1]$:, with **8 rectangles**. What is Δx?

Solution:

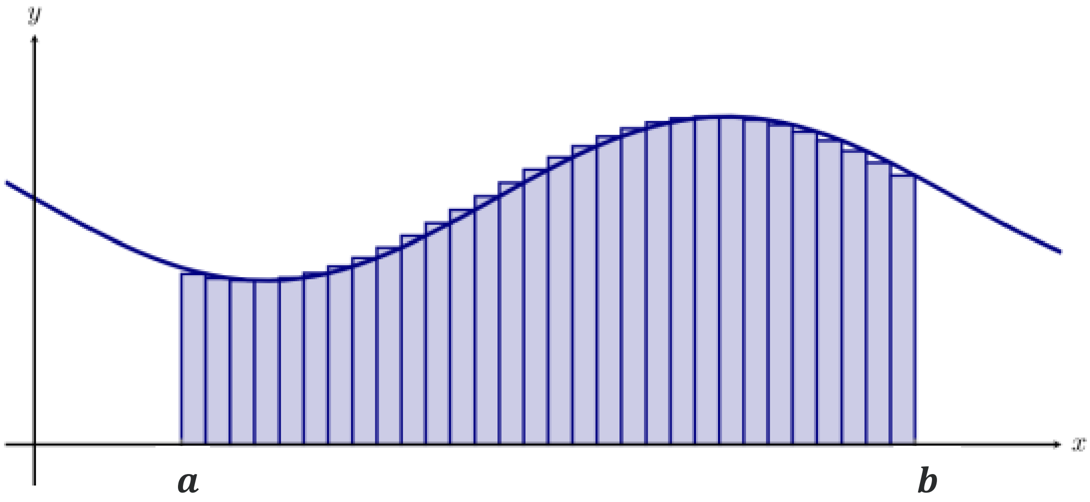
From the problem statement, it is given:

$$a = -1\ , \quad b = 1\ , \quad \text{and} \quad n = 8$$

Then,

$$\Delta x = \frac{b-a}{n} = \frac{1-(-1)}{8} = \frac{2}{8} = \frac{1}{4} = 0.25$$

As adding more rectangles, more closely approximating the area are obtained:

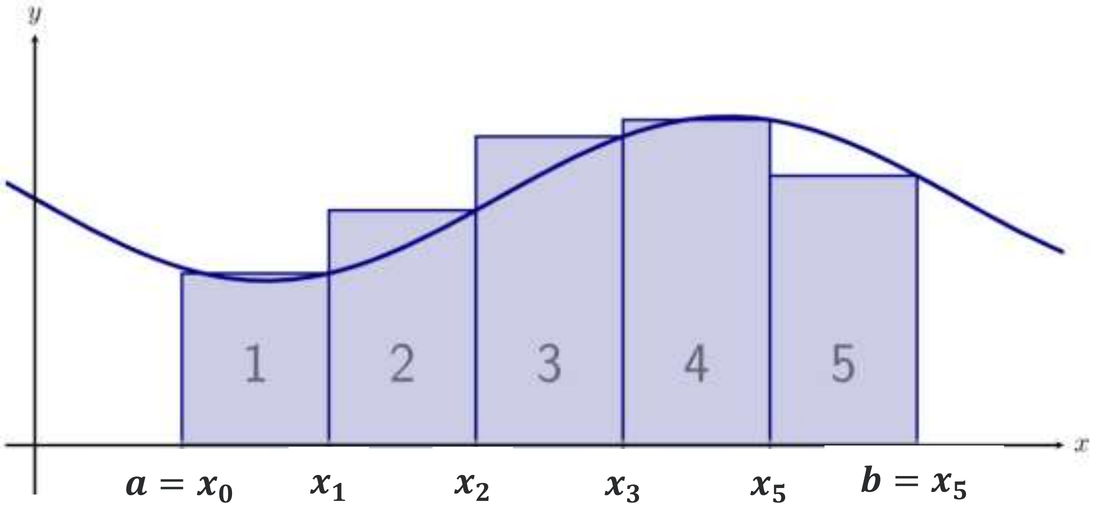


It could find the area exactly if could compute the limit as the width of the rectangles goes to zero and the number of rectangles goes to infinity.

Definition: For approximating an area with nn rectangles, the **grid points**

$$x_0, x_1, x_2, x_3, \cdots, x_n$$

are the x – **coordinates** that determine the edges of the rectangles. In the graph below, the rectangles have been numbered to show the relation between the indices of the grid points and the i^{th} rectangle.



Note, if **approximating** the **area between a curve** and the **horizontal axis** on **interval** $[\boldsymbol{a}, \boldsymbol{b}]$ with $\boldsymbol{n}$ **rectangles**, then it is always the case that

$$\boldsymbol{x} = \boldsymbol{a} \quad \text{and} \quad \boldsymbol{x_n} = \boldsymbol{b}$$

Q/ If approximating the area between a curve and the horizontal axis with **11 rectangles**, how many grid points will be?

<u>Solution</u>:

They will be **12 grid points**, because for 2 – points will be one rectangle, for 3 – points, will be two rectangles and so on, as it illustrated in table below:

Number of points	Number of Rectangles
2	1
3	2
4	3
5	4
6	5
7	6
8	7
9	8
10	9
11	10
12	11

When $\boldsymbol{n}$ **rectangles** are used for **computing** the **area** under a curve, the **width** of **each rectangle** is:

$$\Delta \boldsymbol{x} = \frac{\boldsymbol{b} - \boldsymbol{a}}{\boldsymbol{n}}$$

It is clear that

$$\Delta \boldsymbol{x} = \boldsymbol{x_i} - \boldsymbol{x_{i-1}} \qquad \text{for} \qquad \boldsymbol{i = 1, 2, 3, \cdots, n}.$$

But how to **determine** the **height** of the **rectangle**?

<u>Answer</u>:

Choosing a **sample point** $\boldsymbol{x_i^*}$ and evaluate the function at that point. The value $\boldsymbol{f} = (\boldsymbol{x_i})$ **determines** the **height** of a **rectangle**.

<u>Definition</u>: Approximating an **area with rectangles**, a **sample point** is the x – **coordinate** that **determines** the **height** of $\boldsymbol{i^{th}}$ **rectangle**.

For $\boldsymbol{i = 1, 2, 3, \cdots, n}$ a **sample point** is **denoted** as:

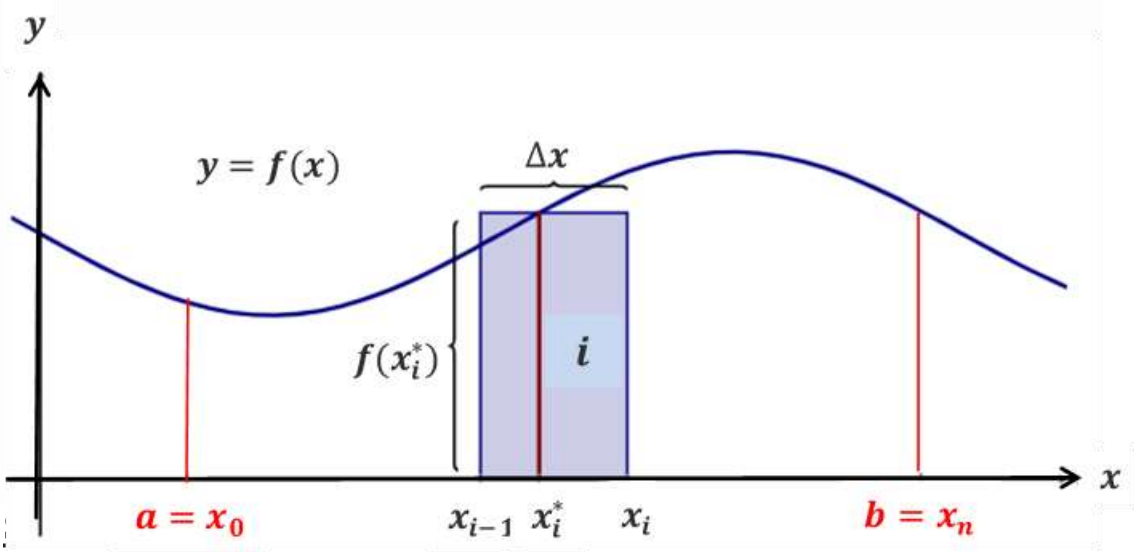
$$\boldsymbol{x_i}$$

and the **value**

$$\boldsymbol{f} = (\boldsymbol{x_i})$$

is the **height** of the $\boldsymbol{i^{th}}$ **rectangle**.

Q/ What is the **area** of the $\boldsymbol{i^{th}}$ **rectangle** shown in the figure below?



$A = \Delta x$
$A = \Delta x \, f(x)$
$A = (x_i^*) f(x_i^*)$
$A = \Delta x f(x_i^*)$
$A = x \, f(x)$
$A = i \, \Delta i$
$A = f(x_i^*)(x_i - x_{i-1})$

There are **three options** for the **sample points** that can be considered.

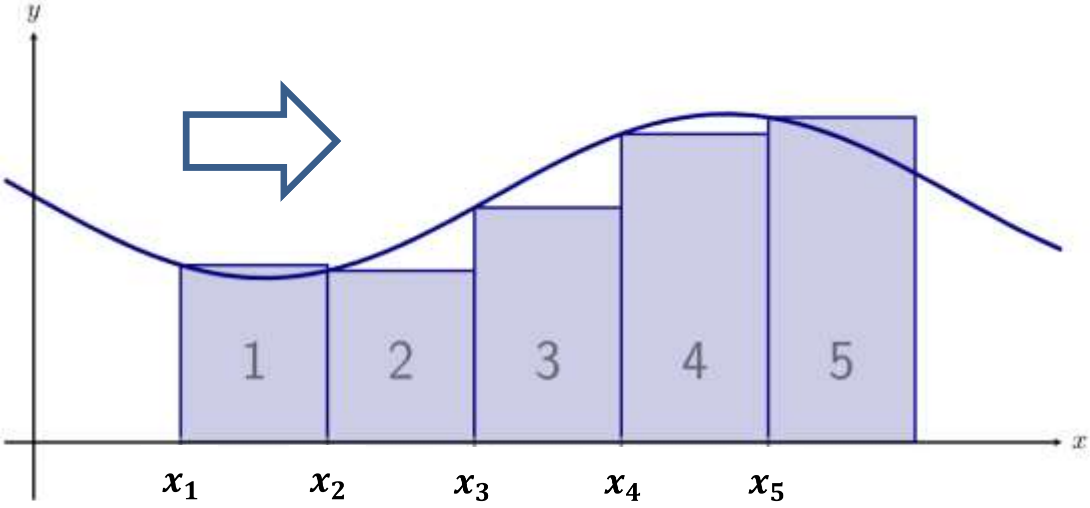
There are **three types** of **rectangles** when **approximating definite integrals**:

- **Left – endpoints,**
- **Right – endpoints**, and
- **Midpoint Rule**.

and **different results** they **do give**.

1. **<u>Rectangles defined by left – endpoints</u>**

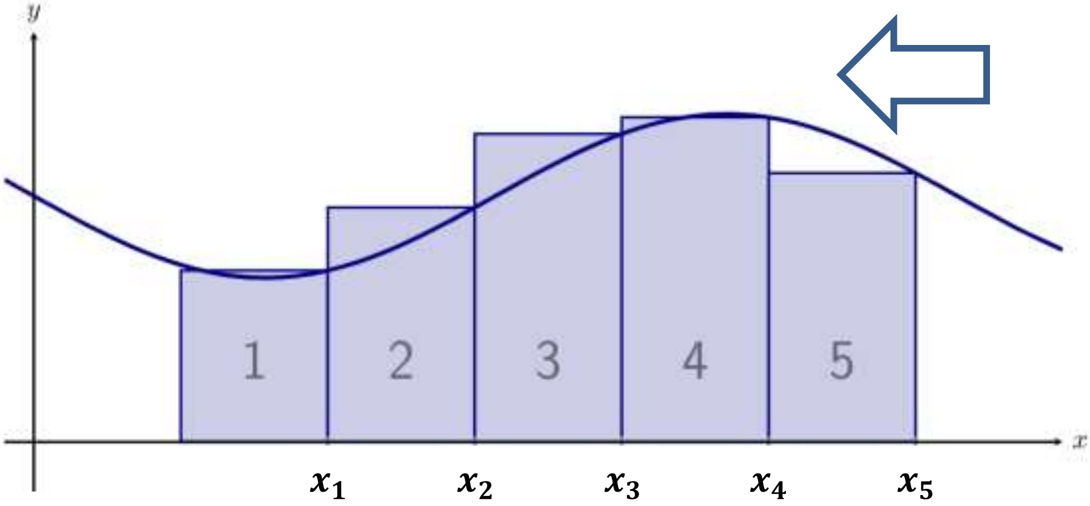
The **rectangles** can be **set up** so that the **left – endpoint** determines the **height** of the **rectangle**.



The $\boldsymbol{i}$ **rectangle's** **left – endpoint** of the **base** determines the **height** of the **rectangle**.

2. **Rectangles defined by right – endpoints**

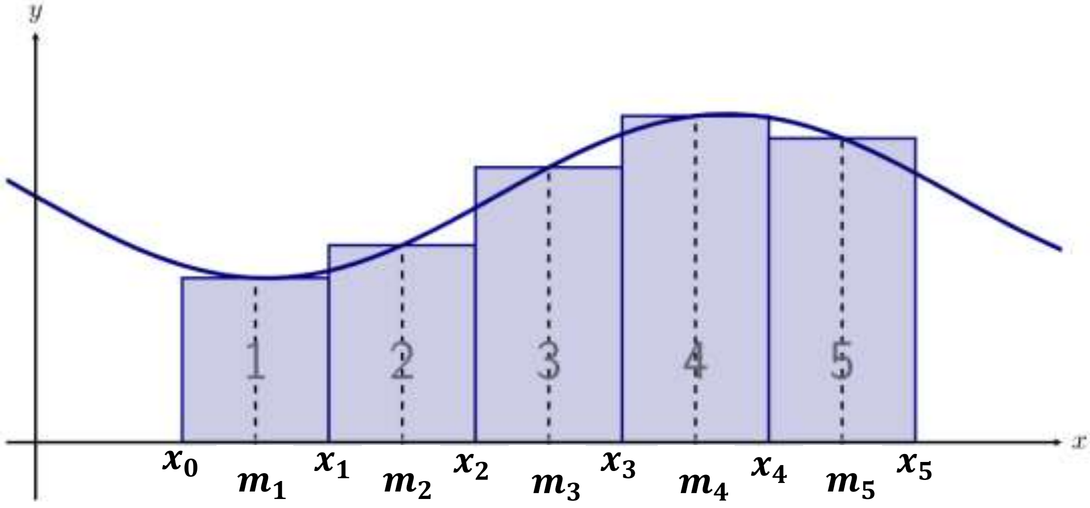
The **rectangles** can be **set up** so that the **right – endpoint** determines the **height** of the **rectangle**.



The $\boldsymbol{i}$ **rectangle's right-endpoint** of the **base** determines the **height** of the **rectangle**.

3. **Rectangles defined by midpoints**

The **rectangles** can be **set up** so that the **midpoint** of the **base** determines the **height** of the **rectangle**.



where,

$$m = \frac{x_i + x_{i+1}}{2}$$

The **midpoint** of the **base** of the $\boldsymbol{i}$ **rectangle** determines the **height** of the **rectangle**.

➢ Riemann sums and approximating area

Once identifying the rectangles, can compute approximations of some areas. If **approximating area** with $\boldsymbol{n}$ **rectangles**, then

$$\boldsymbol{Area} = (\textbf{width of } \boldsymbol{i^{th}} \textbf{ rectangle}) \times \sum_{i=1}^{n} (\textbf{hight of } \boldsymbol{i^{th}} \textbf{ rectangle})$$

$$\Rightarrow \boldsymbol{Area} = \Delta \boldsymbol{x} \sum_{i=1}^{n} \boldsymbol{f}(\boldsymbol{x_i}) = \Delta \boldsymbol{x} \, [\boldsymbol{f}(\boldsymbol{x_1}) + \boldsymbol{f}(\boldsymbol{x_2}) + \boldsymbol{f}(\boldsymbol{x_3}) + \cdots + \boldsymbol{f}(\boldsymbol{x_n})]$$

Definition: A **sum** of the **form**:

$$\Delta \boldsymbol{x} \sum_{i=1}^{n} \boldsymbol{f}(\boldsymbol{x_i}) = \Delta \boldsymbol{x} \, [\boldsymbol{f}(\boldsymbol{x_1}) + \boldsymbol{f}(\boldsymbol{x_2}) + \boldsymbol{f}(\boldsymbol{x_3}) + \cdots + \boldsymbol{f}(\boldsymbol{x_n})]$$

is called a **Riemann sum**, pronounced **"ree-mahn"** sum.

A **Riemann sum** computes an **approximation** of the **area between a curve** and the $\boldsymbol{x}$ **– axis** on the **interval** $[\boldsymbol{a}, \boldsymbol{b}]$. It can be defined in **several different ways** such as:

1. **Left – endpoints,**
2. **Right – endpoints, and**
3. **Midpoints.**

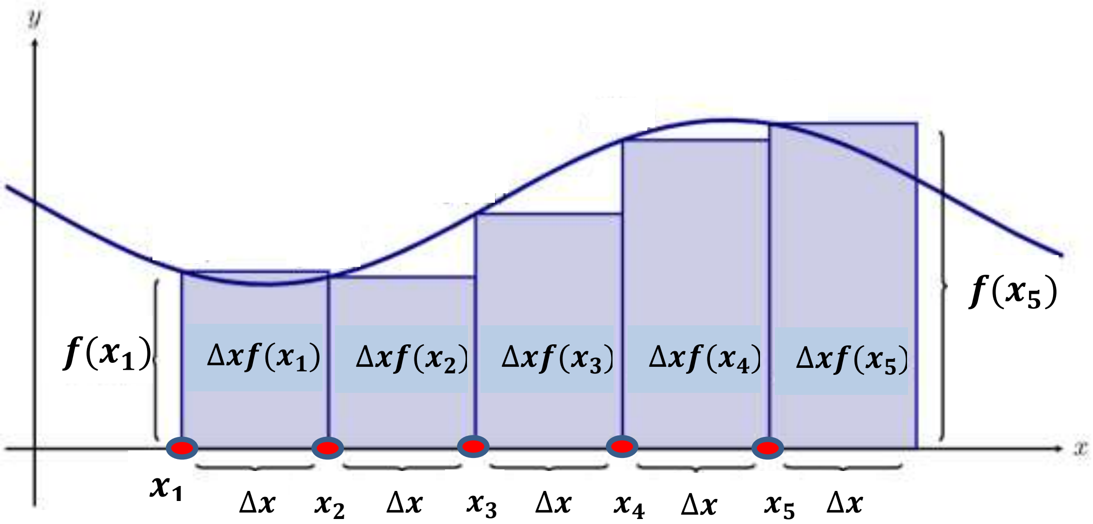
Approximating Definite Integrals Using Left, Right, and Midpoint Riemann Sums with Uniform Partitions

Riemann sums are used to approximate the area under a curve. The areas of a series of n rectangles are summed in which the height of each rectangle is given by the right or left – hand side of the rectangle.

Riemann Sums require both the width and the height of a series of rectangles in order to compute and sum the areas.

1. Left Riemann Sums

It can be seen the **explicit connection** between a **Riemann sum** defined by **left – endpoints** and the **area between a curve** and the ***x* –axis** on the **interval** $[\boldsymbol{a}, \boldsymbol{b}]$:



and here is the associated **Riemann sum**

$$\Delta x \sum_{i=1}^{5} f(x_i) = \Delta x\, [f(x_1) + f(x_2) + f(x_3) + f(x_4) + f(x_n)]$$

Steps for Approximating Using Left Riemann Sums are:

Step 1: Calculate the width, $\Delta \boldsymbol{x}$, of each of the rectangles needed for the Riemann sum. In case of uniform partitions, the width of each rectangle is equal. For $\boldsymbol{n}$ subintervals (or rectangles),

$$\Delta \boldsymbol{x} = \frac{\boldsymbol{b} - \boldsymbol{a}}{\boldsymbol{n}}$$

where $\boldsymbol{a}$ is the lower bound of the definite integral and b is the upper bound of the definite integral.

Step 2: Find the **x – coordinates** of the **left – hand side** of the rectangles, $\boldsymbol{x_i}$ using

$$\boldsymbol{x_i} = \boldsymbol{a} + (\boldsymbol{i} - \boldsymbol{1})\Delta \boldsymbol{x}$$

for each of the $\boldsymbol{n}$ **subintervals**.

Step 3: The height of each rectangle is given by the function value at each point, $\boldsymbol{f}(\boldsymbol{i})$. Therefore, the left Riemann Sum can be computed using the equation:

$$\boldsymbol{A} = \Delta \boldsymbol{x} \sum_{\boldsymbol{i}=\boldsymbol{1}}^{\boldsymbol{n}} \boldsymbol{f}(\boldsymbol{x_i})$$

where $\Delta \boldsymbol{x}$ is the width of each of the $\boldsymbol{n}$ rectangles and $\boldsymbol{f}(\boldsymbol{x_i})$ is the height.

Example 4: Use a left Riemann sum with 3 equal subintervals to approximate the integral

$$\int_{2}^{5} (\boldsymbol{x^2} - \boldsymbol{6x} + \boldsymbol{10})\, \boldsymbol{dx}$$

Solution:

Step 1: First, find the width of each of the rectangles, $\Delta\boldsymbol{x}$. From the problem statement, it is given that $\boldsymbol{n} = \mathbf{3}$, $\boldsymbol{a} = \mathbf{2}$ and $\boldsymbol{b} = \mathbf{5}$. Therefore,

$$\Delta\boldsymbol{x} = \frac{\boldsymbol{b} - \boldsymbol{a}}{\boldsymbol{n}} = \frac{\mathbf{5} - \mathbf{2}}{\mathbf{3}} = \mathbf{1}$$

The width of each rectangle is 1.

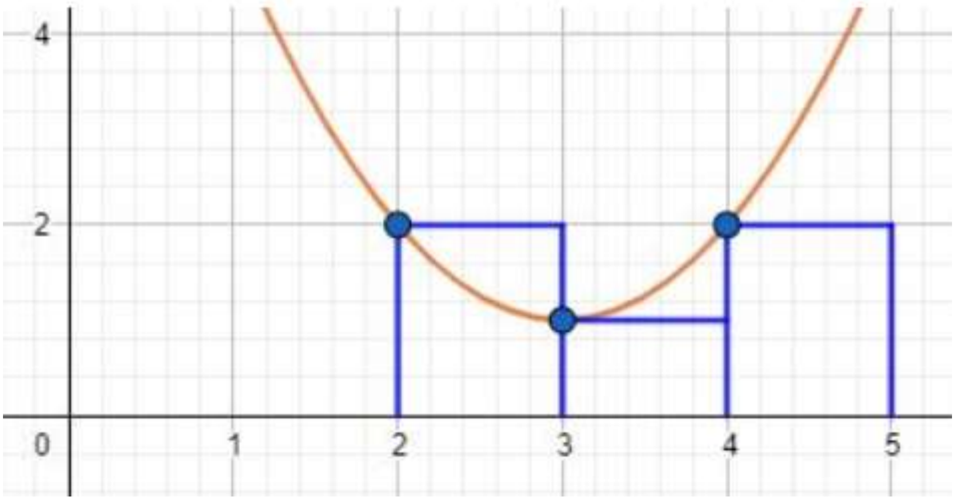
Step 2: Find the x – coordinates of the left – hand side of each rectangle $\boldsymbol{x_i}$.

$$\boldsymbol{x_1} = \boldsymbol{a} + \Delta\boldsymbol{x}(\boldsymbol{i} - \mathbf{1}) = \mathbf{2} + (\mathbf{1}) \times \mathbf{0} = \mathbf{2}$$

$$\boldsymbol{x_2} = \boldsymbol{a} + \Delta\boldsymbol{x}(\boldsymbol{i} - \mathbf{1}) = \mathbf{2} + (\mathbf{1}) \times \mathbf{1} = \mathbf{3}$$

$$\boldsymbol{x_3} = \boldsymbol{a} + \Delta\boldsymbol{x}(\boldsymbol{i} - \mathbf{1}) = \mathbf{2} + (\mathbf{1}) \times \mathbf{2} = \mathbf{4}$$

Step 3: Now the x-coordinates of the left side of our rectangles are known, so the height of each rectangle with $\boldsymbol{x_1}$ can be determined. Already the width of each rectangle is determined to be $\Delta\boldsymbol{x} = \mathbf{1}$. Graphically, our rectangles are the following:



$$\boldsymbol{A} = \Delta\boldsymbol{x}\sum_{\boldsymbol{i}=\mathbf{1}}^{\mathbf{4}} \boldsymbol{f}(\boldsymbol{x_i}) = (\mathbf{1})[\boldsymbol{f}(\boldsymbol{x_1}) + \boldsymbol{f}(\boldsymbol{x_2}) + \boldsymbol{f}(\boldsymbol{x_3}))]$$

$$= \boldsymbol{f}(\mathbf{2}) + \boldsymbol{f}(\mathbf{3}) + \boldsymbol{f}(\mathbf{4}) = \mathbf{2} + \mathbf{1} + \mathbf{2} = \mathbf{5}$$

<u>Example</u> 5: Use the given table to compute the left Riemann Sum of the described function using 6 equal subintervals from $\boldsymbol{x = 0}$ to $\boldsymbol{x = 3}$.

x	y
0	1
1/2	5/4
1	2
3/2	13/4
2	5
5/2	29/4
3	10

<u>Solution</u>:

<u>Step</u> 1: Determine the width of each of rectangles, $\Delta \boldsymbol{x}$. From the problem statement, it is given that $\boldsymbol{n = 6}$, $\boldsymbol{a = 0}$ and $\boldsymbol{b = 3}$. Therefore,

$$\Delta \boldsymbol{x} = \frac{\boldsymbol{b - a}}{\boldsymbol{n}} = \frac{\boldsymbol{3 - 0}}{\boldsymbol{6}} = \frac{\boldsymbol{1}}{\boldsymbol{2}}$$

he width of each rectangle is:

$$\frac{\boldsymbol{1}}{\boldsymbol{2}}$$

<u>Step</u> 2: Find the x - coordinates of the left – hand side of the rectangles using the equation: $\boldsymbol{x_i = a + (i - 1)\,\Delta x}$

$$\boldsymbol{x_1 = a + (i - 1)\,\Delta x = 0 + \frac{1}{2} \times 0 = 0}$$

$$\boldsymbol{x_2 = a + (i - 1)\,\Delta x = 0 + \frac{1}{2} \times 1 = \frac{1}{2}}$$

$$\boldsymbol{x_3 = a + (i - 1)\,\Delta x = 0 + \frac{1}{2} \times 2 = 1}$$

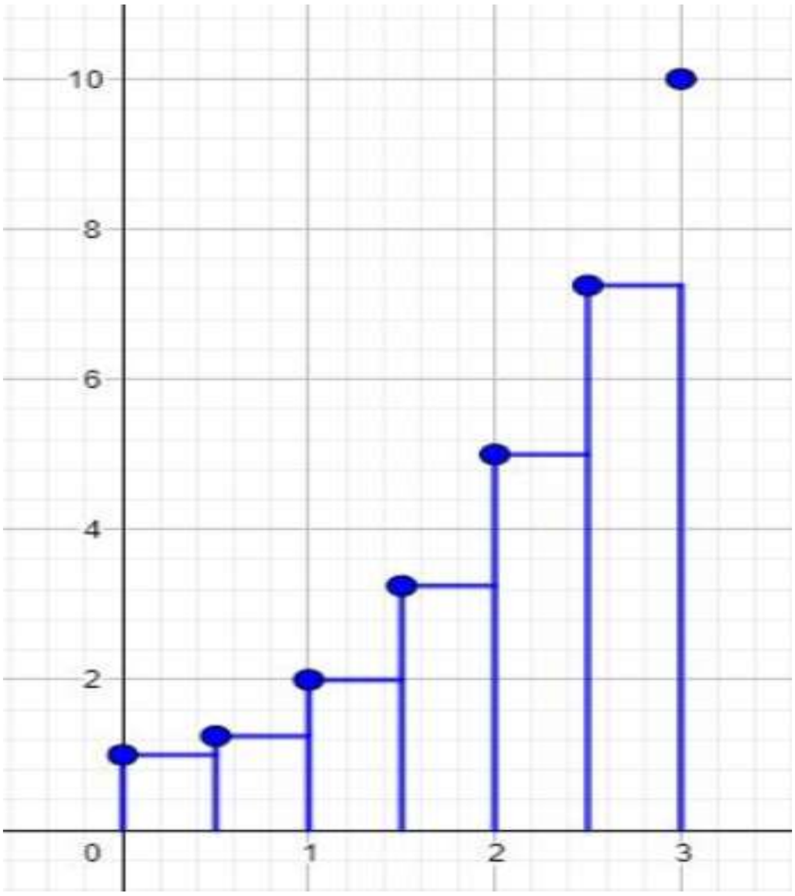
$$\boldsymbol{x_4 = a + (i - 1)\,\Delta x = 0 + \frac{1}{2} \times 3 = \frac{3}{2}}$$

$$\boldsymbol{x_5 = a + (i - 1)\,\Delta x = 0 + \frac{1}{2} \times 4 = 2}$$

$$\boldsymbol{x_6 = a + (i - 1)\,\Delta x = 0 + \frac{1}{2} \times 5 = \frac{5}{2}}$$

<u>Step</u> 3: Now having the locations of the left sides of the rectangles and their width, the rectangles can be drawn graphically, to have a better understanding of what will be calculated.

.



$$\boldsymbol{A} = \Delta \boldsymbol{x} \sum_{\boldsymbol{i}=\boldsymbol{1}}^{\boldsymbol{4}} \boldsymbol{f}(\boldsymbol{x_i})$$

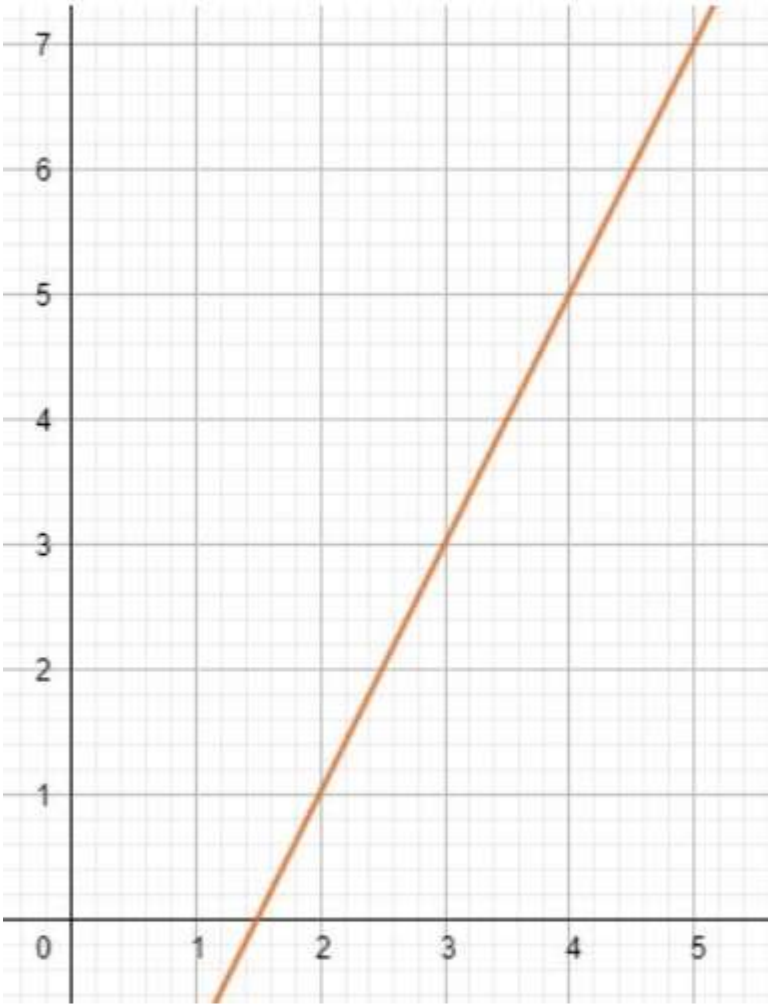
$$= \left(\frac{\mathbf{1}}{\mathbf{2}}\right)[\boldsymbol{f}(\boldsymbol{x_1}) + \boldsymbol{f}(\boldsymbol{x_2}) + \boldsymbol{f}(\boldsymbol{x_3}) + \boldsymbol{f}(\boldsymbol{x_4}) + \boldsymbol{f}(\boldsymbol{x_5}) + \boldsymbol{f}(\boldsymbol{x_6})]$$

$$= \left(\frac{\mathbf{1}}{\mathbf{2}}\right)\left[\boldsymbol{f}(\mathbf{0}) + \boldsymbol{f}\left(\frac{\mathbf{1}}{\mathbf{2}}\right) + \boldsymbol{f}(\mathbf{1}) + \boldsymbol{f}\left(\frac{\mathbf{3}}{\mathbf{2}}\right) + \boldsymbol{f}(\mathbf{2}) + \boldsymbol{f}\left(\frac{\mathbf{5}}{\mathbf{2}}\right)\right]$$

$$= \left(\frac{\mathbf{1}}{\mathbf{2}}\right)\left[\mathbf{1} + \frac{\mathbf{5}}{\mathbf{4}} + \mathbf{2} + \frac{\mathbf{13}}{\mathbf{4}} + \mathbf{5} + \frac{\mathbf{29}}{\mathbf{4}}\right] = \frac{\mathbf{1}}{\mathbf{2}} \times \frac{\mathbf{79}}{\mathbf{4}} = \frac{\mathbf{79}}{\mathbf{8}}$$

<u>Example</u> 6: Approximate the area under the curve using left Riemann Sums and 3 equal subintervals from $\boldsymbol{x = 2}$ to $\boldsymbol{x = 5}$.

.



<u>Solution</u>:

<u>Step</u> 1: Find the width of each of the rectangles, $\Delta \boldsymbol{x}$. From the problem statement it is given that $\boldsymbol{n = 3}$, $\boldsymbol{a = 2}$ and $\boldsymbol{b = 5}$. Therefore,

$$\Delta \boldsymbol{x} = \frac{\boldsymbol{b - a}}{\boldsymbol{n}} = \frac{\boldsymbol{5 - 2}}{\boldsymbol{3}} = \boldsymbol{1}$$

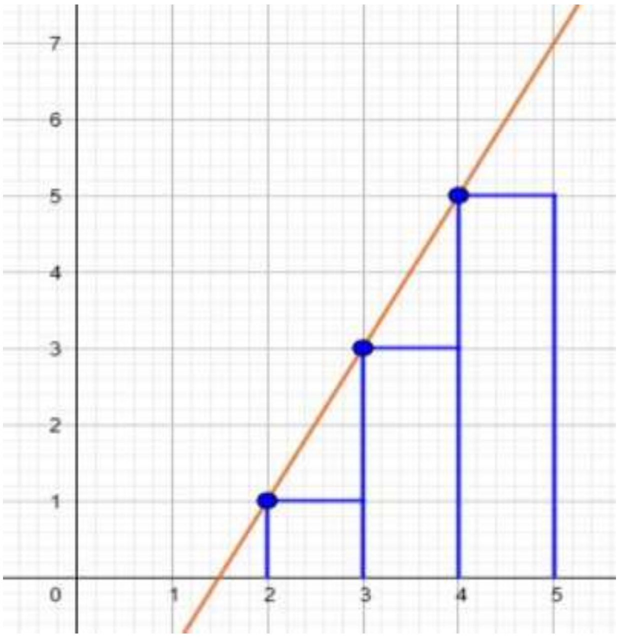
<u>Step</u> 2: Find the x - coordinates of the left – hand side of the rectangles using the equation: $\boldsymbol{x_i = a + (i - 1)\,\Delta x}$

$$\boldsymbol{x_1 = a + (i - 1)\,\Delta x = 2 + (0) \times (1) = 2}$$

$$\boldsymbol{x_2 = a + (i - 1)\,\Delta x = 2 + (1) \times (1) = 3}$$

$$\boldsymbol{x_3 = a + (i - 1)\,\Delta x = 2 + (2) \times (1) = 4}$$

<u>Step</u> 3: Now having the x - coordinates of the left - hand side of the rectangles and their width, the rectangles can be drawn graphically, to aid a better understanding of what will be calculated.



$$\boldsymbol{A} = \Delta \boldsymbol{x} \sum_{\boldsymbol{i}=1}^{2} \boldsymbol{f}(\boldsymbol{x}_{\boldsymbol{i}}) = (\mathbf{1})[\boldsymbol{f}(\boldsymbol{x}_{\mathbf{1}}) + \boldsymbol{f}(\boldsymbol{x}_{\mathbf{2}}) + \boldsymbol{f}(\boldsymbol{x}_{\mathbf{3}})]$$

$$= \boldsymbol{f}(\mathbf{2}) + \boldsymbol{f}(\mathbf{3}) + \boldsymbol{f}(\mathbf{4}) = \mathbf{1} + \mathbf{3} + \mathbf{5} = \mathbf{9}$$

2. <u>Right Riemann Sums</u>

Right Riemann sums are used to approximate the area under a curve. The areas of a series of n rectangles are summed in which the height of each rectangle is given by the right side of the rectangle. Depending on the curve, a right Riemann sum may be an under or over approximation of the actual area. The formula for a right Riemann sum is:

$$\boldsymbol{A} = \Delta \boldsymbol{x} \sum_{\boldsymbol{i}=1}^{\boldsymbol{n}} \boldsymbol{f}(\boldsymbol{x}_{\boldsymbol{i}})$$

where $\Delta \boldsymbol{x}$ is the width of each of the $\boldsymbol{n}$ rectangles and $\boldsymbol{f}(\boldsymbol{x}_{\boldsymbol{i}})$ is the height.

Steps for Approximating Using Right Riemann Sums are:

Step 1: Calculate the width, Δx, of each of the rectangles needed for the Riemann sum. In case of uniform partitions, the width of each rectangle is equal. For nn subintervals (or rectangles),

$$\Delta x = \frac{b-a}{n}$$

where a is the lower bound of the definite integral and b is the upper bound of the definite integral.

Step 2: Locate the right endpoint of each of the rectangles referring to these endpoints as x_i and find the right endpoints using the equation

$$x_i = a + i\,\Delta x$$

Step 3: Compute the right Riemann sum.

Example 1: Approximate the following integral using a right Riemann sum and 4 equal subintervals.

$$\int_0^{12} (x^2 - 4x + 4)\,dx$$

Solution:

Step 1: First, find the width of each of the subintervals, Δx. From the problem statement, it is given that $n = 4$, $a = 0$ and $b = 12$. Therefore,

$$\Delta x = \frac{b-a}{n} = \frac{12-0}{4} = 3$$

Step 2: Find the locations of the right endpoints of the rectangles by finding x_i for each of the 4 rectangles.

$$x_1 = a + (1)\,\Delta x = 0 + (1) \times 3 = 3$$

$$x_2 = a + (2)\,\Delta x = 0 + (2) \times 3 = 6$$

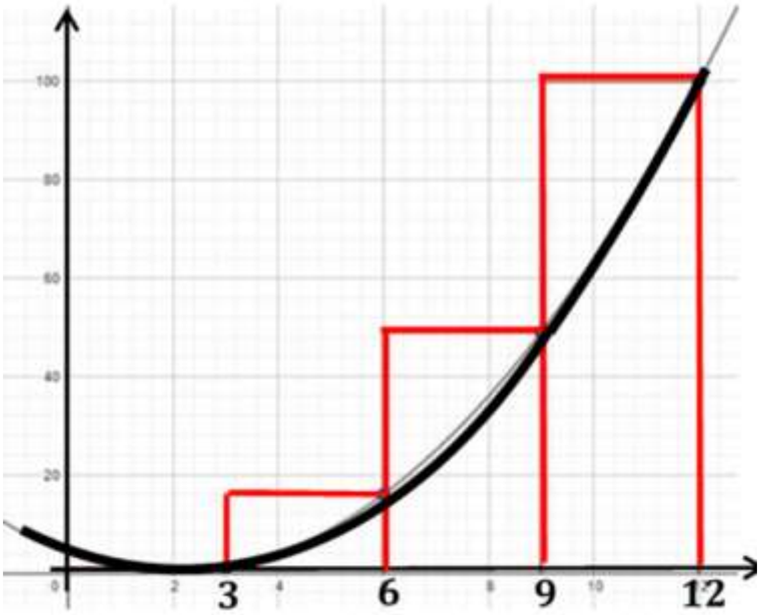
$$x_3 = a + (3)\,\Delta x = 0 + (3) \times 3 = 9$$

$$x_4 = a + (4)\,\Delta x = 0 + (4) \times 3 = 12$$

Step 3: Computing the right Riemann sum where the right endpoints are:

$x_1 = 3$, $x_2 = 6$, $x_3 = 9$, and $x_4 = 12$.

The areas need to be summed are illustrated in the figure.



$$A = \Delta x \sum_{i=1}^{4} f(x_i) = 3[f(x_1) + f(x_2) + f(x_3) + f(x_4)]$$

$$= 3[f(3) + f(6) + f(9) + f(12)]$$

$$= 3(1 + 16 + 49 + 100) = 498$$

The **actual value** of the integral is:

$$\int_{o}^{12} (x^2 - 4x + 4)\,dx = \left[\frac{x^3}{3} - 4\frac{x^2}{2} + 4x\right]_0^{12}$$

$$= 576 - 288 + 48 = 336$$

$$\varepsilon = |336 - 498| = 162$$

<u>Example</u> 2: Approximate the area under the curve from $\boldsymbol{x = 1}$ to $\boldsymbol{x = 6}$ using 5 equal subintervals and the function information provided in the given table.

x	y
1	1
2	8
3	27
4	64
5	125
6	216

<u>Solution</u>:

<u>Step</u> 1: Find the width of each of the subintervals, $\Delta\boldsymbol{x}$. From the problem statement, it is given that $\boldsymbol{n = 5}$, $\boldsymbol{a = 1}$ and $\boldsymbol{b = 6}$. Therefore,

$$\Delta\boldsymbol{x} = \frac{\boldsymbol{b - a}}{\boldsymbol{n}} = \frac{\boldsymbol{6 - 1}}{\boldsymbol{5}} = \boldsymbol{1}$$

<u>Step</u> 2: Find the locations of the right endpoints of the rectangles using the equation: $\boldsymbol{x_i = a + i\,\Delta x}$

$$\boldsymbol{x_1 = a + (1)\,\Delta x = 1 + (1) \times 1 = 2}$$

$$\boldsymbol{x_2 = a + (2)\,\Delta x = 1 + (2) \times 1 = 3}$$

$$\boldsymbol{x_3 = a + (3)\,\Delta x = 1 + (3) \times 1 = 4}$$

$$\boldsymbol{x_4 = a + (4)\,\Delta x = 1 + (4) \times 1 = 5}$$

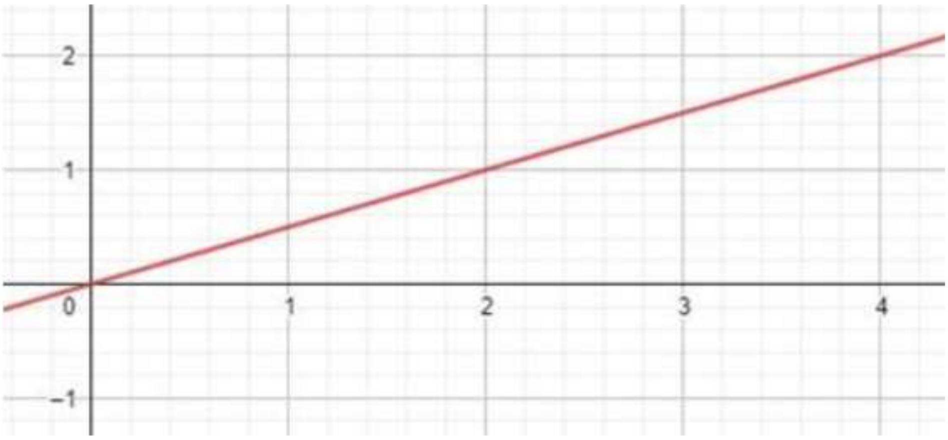
$$\boldsymbol{x_4 = a + (5)\,\Delta x = 1 + (5) \times 1 = 6}$$

<u>Step</u> 3: Compute the right Riemann sum using the table to obtain the function values.

$$\boldsymbol{A = \Delta x \sum_{i=1}^{4} f(x_i) = (1)[f(x_1) + f(x_2) + f(x_3) + f(x_4) + f(x_5)]}$$

$$\boldsymbol{= 8 + 27 + 64 + 125 + 216 = 440}$$

Example 3: Use a right Riemann sum and 2 equal subintervals to approximate the area under the curve from $\boldsymbol{x = 1}$ to $\boldsymbol{x = 6}$ using the given graph.



Solution:

Step 1: Find the width of each of the rectangles, $\Delta\boldsymbol{x}$. From the graph, it is evident that $\boldsymbol{n = 2},\ \boldsymbol{a = 0}$ and $\boldsymbol{b = 4}$. Therefore,

$$\Delta\boldsymbol{x} = \frac{\boldsymbol{b - a}}{\boldsymbol{n}} = \frac{\boldsymbol{4 - 0}}{\boldsymbol{2}} = \boldsymbol{2}$$

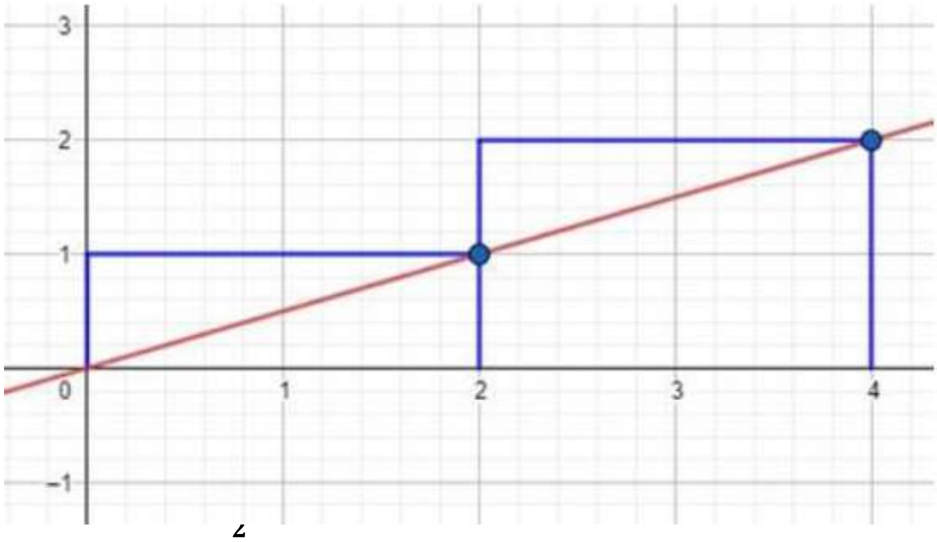
Step 2: Find the locations of the right endpoints of the rectangles using the equation:

$$\boldsymbol{x_i = a + i\,\Delta x}$$

$$\boldsymbol{x_1 = a + (1)\,\Delta x = 0 + (1) \times 2 = 2}$$

$$\boldsymbol{x_2 = a + (2)\,\Delta x = 0 + (2) \times 2 = 4}$$

Step 3: Compute the right Riemann sum using the graph to obtain the function values.



$$A = \Delta x \sum_{i=1}^{2} f(x_i) = (2)[f(x_1) + f(x_2)] = 2[1+2] = 6$$

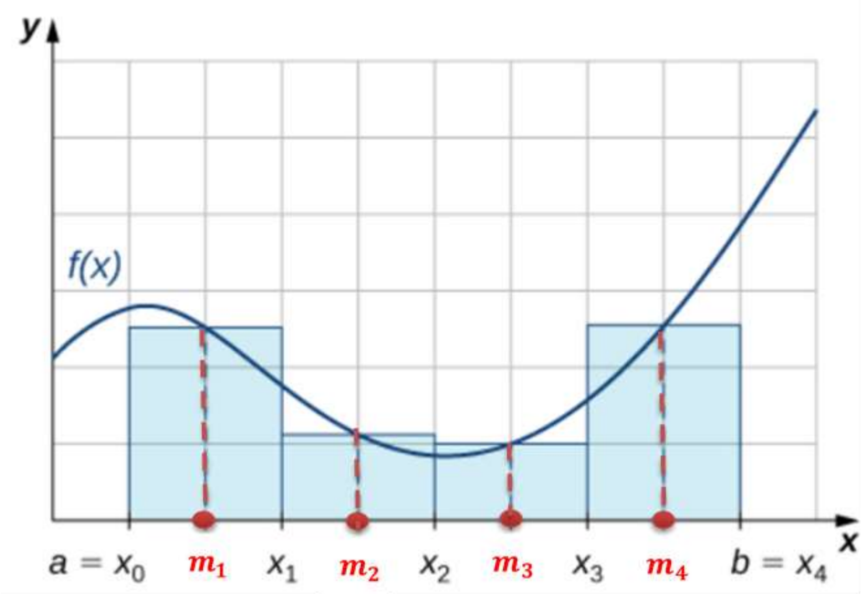
3. **Midpoint Rule**

Assume that $f(x)$ is continuous on $[a, b]$. Let n be a positive integer and

$$\Delta x = \frac{b-a}{n}$$

If $[a, b]$ is divided into n subintervals, each of length Δx, and m_i is the **midpoint** of the i^{th} subinterval, then

$$M_n = \Delta x \sum_{i=1}^{n} f(m_i) \qquad \text{and} \qquad \lim_{n\to\infty} M_n = \int_a^b f(x)dx$$



As it is shown from the Figure;

If $\boldsymbol{f(x) \geq 0}$ over $[\boldsymbol{a}, \boldsymbol{b}]$,

then

$$\Delta \boldsymbol{x} \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{f(m_i)}$$

corresponds to the **sum** of the **areas** of **rectangles** approximating the area between the graph of $\boldsymbol{f(x)}$ and the $\boldsymbol{x}$ **– axis** over $[\boldsymbol{a}, \boldsymbol{b}]$. The graph shows the rectangles corresponding to $\boldsymbol{M_n}$ for a nonnegative function over a closed interval $[\boldsymbol{a}, \boldsymbol{b}]$.

<u>Example</u> 1: Use the **midpoint rule** to estimate the integral

$$\int_{\boldsymbol{0}}^{\boldsymbol{1}} \boldsymbol{x^2\, dx}$$

using **4 subintervals. Compare** the result with **actual value** of this **integral.**

<u>Solution</u>:

Each subinterval has length:

$$\Delta \boldsymbol{x} = \frac{\boldsymbol{b-a}}{\boldsymbol{4}} = \frac{\boldsymbol{1-0}}{\boldsymbol{4}} = \frac{\boldsymbol{1}}{\boldsymbol{4}} = \boldsymbol{0.25}$$

and the subintervals are:

$$\left[\boldsymbol{0}\ , \frac{\boldsymbol{1}}{\boldsymbol{4}}\right], \quad \left[\frac{\boldsymbol{1}}{\boldsymbol{4}}\ , \frac{\boldsymbol{1}}{\boldsymbol{2}}\right], \quad \left[\frac{\boldsymbol{1}}{\boldsymbol{2}}\ , \frac{\boldsymbol{3}}{\boldsymbol{4}}\right], \quad \text{and } \left[\frac{\boldsymbol{3}}{\boldsymbol{4}}\ , \boldsymbol{1}\right]$$

The **midpoints** of these **subintervals** are:

$$\left[\ \frac{\boldsymbol{1}}{\boldsymbol{8}}\ ,\quad \frac{\boldsymbol{3}}{\boldsymbol{8}}\ ,\quad \frac{\boldsymbol{5}}{\boldsymbol{8}}\ ,\quad \frac{\boldsymbol{7}}{\boldsymbol{8}}\ \right]$$

Thus,

$$\boldsymbol{M_4} = \Delta \boldsymbol{x} \sum_{\boldsymbol{i=1}}^{\boldsymbol{4}} \boldsymbol{f(m_i)} = \frac{\boldsymbol{1}}{\boldsymbol{4}}\left[\boldsymbol{f}\left(\frac{\boldsymbol{1}}{\boldsymbol{8}}\right) + \boldsymbol{f}\left(\frac{\boldsymbol{3}}{\boldsymbol{8}}\right) + \boldsymbol{f}\left(\frac{\boldsymbol{5}}{\boldsymbol{8}}\right) + \boldsymbol{f}\left(\frac{\boldsymbol{7}}{\boldsymbol{8}}\right)\right]$$

$$\approx \frac{1}{4}\left(\frac{1}{64}+\frac{9}{64}+\frac{25}{64}+\frac{21}{64}\right)=\frac{21}{64}$$

The **actual value** is:

$$I=\int_0^1 x^2\,dx=\left.\frac{x^3}{3}\right|_0^1=\frac{1}{3}$$

The **error** in this approximation is:

$$\left|\frac{1}{3}-\frac{21}{64}\right|=\frac{1}{192}\approx 0.0052$$

Example 2: Use M_6 to estimate the **length** of the **curve** $y=\frac{1}{2}x^2$ on $[1,4]$.

Solution:

The length of $y=\frac{1}{2}x^2$ on $[1,4]$ is:

$$s=\int_1^4 \sqrt{1+\left(\frac{dy}{dx}\right)^2}\;dx$$

Since $\frac{dy}{dx}=x$, this integral becomes

$$s=\int_1^4 \sqrt{1+x^2}\,dx$$

If $[1,4]$ is divided into **6 subintervals,** then each subinterval has length

$$\Delta x=\frac{b-a}{6}=\frac{4-1}{6}=\frac{1}{2}=0.5$$

and the midpoints of the subintervals are:

$$\frac{5}{4},\quad \frac{7}{4},\quad \frac{9}{4},\quad \frac{11}{4},\quad \frac{13}{4},\quad \frac{15}{4}$$

If set $f(x)=\sqrt{1+x^2}$

$$M_6 = \Delta x \sum_{i=1}^{6} f(m_i)$$

$$= \frac{1}{2}\left[f\left(\frac{5}{4}\right) + f\left(\frac{7}{4}\right) + f\left(\frac{9}{4}\right) + f\left(\frac{11}{4}\right) + f\left(\frac{13}{4}\right) + f\left(\frac{15}{4}\right)\right]$$

$$\approx \frac{1}{2}(1.6008 + 2.0156 + 2.4622 + 2.9262 + 3.4004 + 3.8810)$$

$$= 8.1431$$

HW: Use the **midpoint rule** with **n = 2 to estimate**

$$\int_1^2 \frac{1}{x}\, dx$$

Hint: $\Delta x = \frac{1}{2}$, $m_1 = \frac{3}{4}$, and $m_2 = \frac{7}{4}$, Asw: $\frac{24}{35}$