1 Introduction to Differentiation

Why do we care about derivatives?

Derivatives in natural science are widely used as they help understanding some phenomena. Some of the most common case uses of derivatives are:

- Rate of change: Chemistry deals with many dynamic processes that evolve over time.

- Optimization: Derivatives are widely use to find minima or maxima of functions.

The derivative of a function $f(x)$ at a point x is the slope of the straight line tangent to $f(x)$ at x. In the limiting process, the derivative of $f(x)$ at a point x is defined as,

$$f'(x) = \frac{df(x)}{dx} = \frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \tag{1}$$

Fig. 1 illustrates the limiting process for when $\Delta x \to 0$. Eq. 1 is also know as the definition of the **ordinary derivate**.
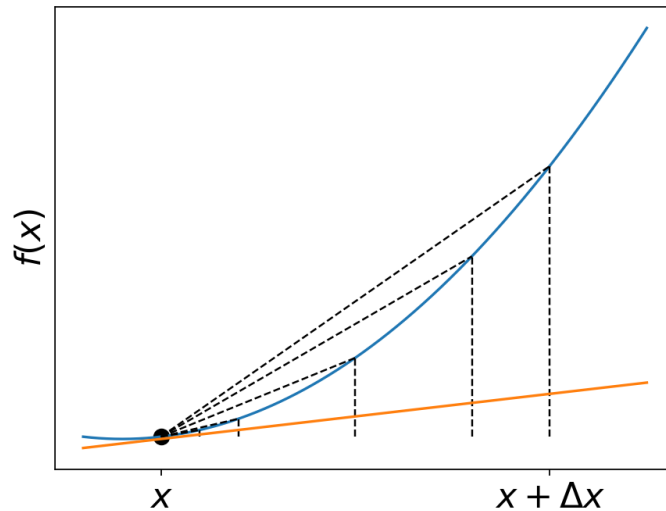

Figure 1: An illustration of the limiting process in the definition of the derivative for $f(x)$.

Exercise: Let's compute the derivative of $f(x) = x^2$ using Eq. 1.

$$\begin{aligned}
f'(x) &= \lim_{\Delta x \to 0} \frac{(x + \Delta x)^2 - x^2}{\Delta x} = \lim_{\Delta x \to 0} \frac{\cancel{x^2} + 2x\Delta x + \Delta x^2 \cancel{- x^2}}{\Delta x} & (2) \\
&= \lim_{\Delta x \to 0} \frac{\Delta x(2x + \Delta x)}{\Delta x} = \lim_{\Delta x \to 0} 2x + \Delta x = 2x + \cancel{\lim_{\Delta x \to 0} \Delta x} & (3) \\
&= 2x & (4)
\end{aligned}$$

Exercise: Let's compute the derivative of the so called function `ReLU` (rectified linear unit), $f(x) = max(0, x)$ using Eq. 1. For this exercise, we will consider three scenarios, 1) x is positive $x \in \mathbb{R}^+$, 2) x is negative $x \in \mathbb{R}^-$, and 3) $x = 0$.

1. x is positive:

$$f'(x) \quad = \quad \lim_{\Delta x \to 0} \frac{\max(0, x + \Delta x) - \max(0, x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{\cancel{x} + \Delta x - \cancel{x}}{\Delta x} = 1. \tag{5}$$

2. x is negative:

$$f'(x) \quad = \quad \lim_{\Delta x \to 0} \frac{\max(0, x + \Delta x) - \max(0, x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{0 - 0}{\Delta x} = 0. \tag{6}$$

3.1 x is equal to zero and $\Delta x \in \mathbb{R}^+$:

$$f'(x) \quad = \quad \lim_{\Delta x \to 0} \frac{\max(0, x + \Delta x) - \max(0, x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{\Delta x}{\Delta x} = 1. \tag{7}$$

3.2 x is equal to zero and $\Delta x \in \mathbb{R}^-$:

$$f'(x) \quad = \quad \lim_{\Delta x \to 0} \frac{\max(0, x + \Delta x) - \max(0, x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{0}{\Delta x} = 0. \tag{8}$$

As we can observe, the derivative $f'(x)$ when $x = 0$ has two different possible values,

$$f'(x = 0) = \begin{cases} 1 \text{ for } \Delta x \in \mathbb{R}^+ \\ 0 \text{ for } \Delta x \in \mathbb{R}^- \end{cases} \tag{9}$$

.

When the derivative of a function has two different possible values, the function is **discontinuous**.

In the infinite limit when $\Delta x \to 0$, we can defined as,

$$dy = f'(x)dx, \tag{10}$$

where dy and dx are known as the **differentials**.

2 Common derivatives

In real-life, one usually consult a table or memorize the most common derivatives. Later in the course we will see that the derivative of more complicated functions can be simplified using the **chain rule**. For the following derivatives we assume a is a constant, and $f(x)$ and $g(x)$ both are functions of x.

1. **Derivative of a constant**

$$\frac{da}{dx} = 0 \tag{11}$$

2. **Derivative of a sum**

$$\frac{d \ f(x) + g(x)}{dx} = \frac{d \ f(x)}{dx} + \frac{d \ g(x)}{dx} \tag{12}$$

3. **Derivative of a product**

$$\frac{d\ f(x)g(x)}{dx} = \frac{d\ f(x)}{dx}g(x) + f(x)\frac{d\ g(x)}{dx} \qquad (13)$$

4. **Derivative of a polynomial**

$$\frac{d\ x^n}{dx} = nx^{n-1} \qquad (14)$$

5. **Derivative of trigonometric functions**

$$\frac{d\ \cos(x)}{dx} = -\sin(x) \qquad (15)$$

$$\frac{d\ \sin(x)}{dx} = \cos(x) \qquad (16)$$

6. **Derivative of logarithm**

$$\frac{d\ \ln(x)}{dx} = \frac{1}{x} \qquad (17)$$

7. **Derivative of exponential**

$$\frac{d\ e^x}{dx} = e^x \qquad (18)$$

With the previous list of derivatives one can derive the derivate of,

1. **Derivative of** $af(x)$

$$\frac{d\ af(x)}{dx} = a\frac{d\ f(x)}{dx} + f(x)\frac{d\ \cancel{a}}{\cancel{dx}} = a\frac{d\ f(x)}{dx} \qquad (19)$$

2. **Derivative of the ratio between two functions**

$$\frac{d\ \frac{f(x)}{g(x)}}{dx} = \frac{d\ f(x)g(x)^{-1}}{dx} = g(x)^{-1}\frac{d\ f(x)}{dx} + f(x)\underbrace{\frac{d\ g(x)^{-1}}{dx}}_{\text{chain rule}} \qquad (20)$$

$$= g(x)^{-1}\frac{d\ f(x)}{dx} + f(x)(-1)g(x)^{-2}\frac{d\ g(x)}{dx} \qquad (21)$$

$$= \left(\frac{g(x)}{g(x)}\right)\frac{\frac{d\ f(x)}{dx}}{g(x)} - \frac{f(x)\frac{d\ g(x)}{dx}}{g(x)^2} = \frac{g(x)\frac{d\ f(x)}{dx} - f(x)\frac{d\ g(x)}{dx}}{g(x)^2} \qquad (22)$$

$$= \frac{1}{g(x)}\frac{d\ f(x)}{dx} - \frac{f(x)}{g(x)^2}\frac{d\ g(x)}{dx} \qquad (23)$$

2.1 Chain rule: derivative of a function of functions

As you will see through this course, it is convenient sometimes to see more complicated functions as a composition of simple functions, $(f \circ g)(x) = f(g(x))$, for example,

$$f(x) = (x+1)^2 \qquad (24)$$

$$f(u) = u^2, \text{ where we defined } u(x) = x+1. \qquad (25)$$

The derivative of $f(u)$ w.r.t. to u and $u(x)$ w.r.t. to x are simpler than the derivative of $f(x)$ w.r.t. to x. We can compute $\frac{df(x)}{dx}$ using the chain rule[1],

$$\frac{df(x)}{dx} = \frac{df(g(x))}{x} = \frac{df(u)}{du} \times \frac{du}{dx}, \tag{26}$$

where for connivance we define the variable $u = g(x)$.

Exercise: Compute the derivative of $f(x) = (x+1)^2$ using the chain rule.

$$\frac{df(x)}{dx} = \frac{du}{dx} \times \frac{df(u)}{du} = (1) \times (2u) = 2(x+1) = 2x + 2. \tag{27}$$

Exercise: Let's compute the derivative of $f(x) = e^{-(x^2+a^2)^{\frac{1}{2}}}$ using the chain rule multiple times.

First, we can define $u_1 = -(x^2 + a^2)^{\frac{1}{2}}$,

$$f'(x) = f'(u_1) = \frac{df(u_1)}{du_1} = \frac{de^{u_1}}{du_1} \underbrace{\frac{du_1}{dx}}_{\text{chain rule}} = e^{u_1} \frac{du_1}{dx}. \tag{28}$$

Similarly, we can defined a secondary auxiliary function $u_2 = x^2 + a^2$, meaning $u_1 = -u_2^{\frac{1}{2}}$. Because of this relation, $f(x)$ can be decomposed as,

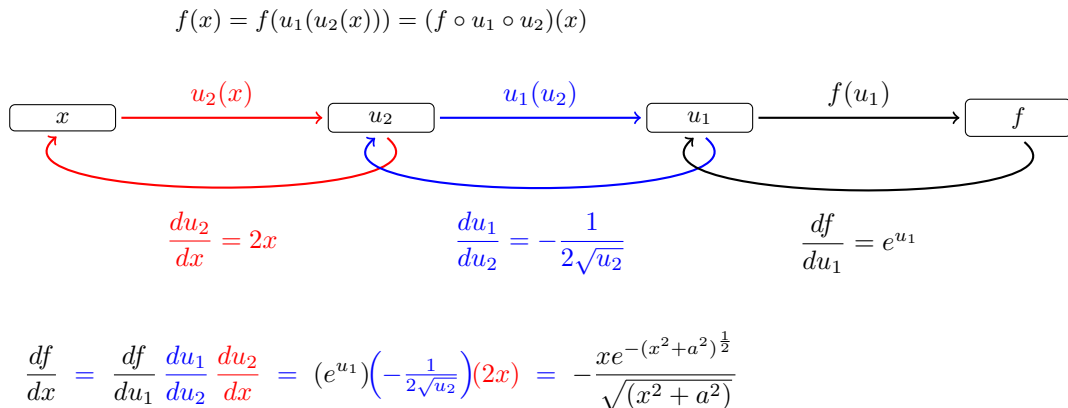$$f(x) = f(u_1) = f(u_1(u_2(x))) = (f \circ u_1 \circ u_2)(x) \tag{29}$$

and its derivative

$$f'(x) = \left(\frac{df(u_1)}{du_1}\right)\left(\frac{du_1}{du_2}\right)\left(\frac{du_2}{dx}\right) = (e^{u_1})\left(-\frac{1}{2\sqrt{u_2}}\right)(2x), \tag{30}$$

if we clean this expression and plug back in the definition of u_1 and u_2, we get,

$$f'(x) = -\frac{xe^{-(x^2+a^2)^{\frac{1}{2}}}}{\sqrt{(x^2+a^2)}}. \tag{31}$$

Here is the chain rule diagram for $f'(x)$:

$$f(x) = f(u_1(u_2(x))) = (f \circ u_1 \circ u_2)(x)$$

$$\frac{df}{dx} = \frac{df}{du_1}\frac{du_1}{du_2}\frac{du_2}{dx} = (e^{u_1})\left(-\frac{1}{2\sqrt{u_2}}\right)(2x) = -\frac{xe^{-(x^2+a^2)^{\frac{1}{2}}}}{\sqrt{(x^2+a^2)}}$$

---

[1]The chain rule equation is one of the main components in modern Machine learning libraries, we will talk more about this throughout the course.

3 Differentiation for multi-dimensional functions

3.1 Partial Derivatives

In natural sciences, we commonly need functions that depend on multiple variables to describe a phenomenon. For example, in thermodynamics the equation of state depends on temperature T, pressure P and volume V, $f(P, V, T)$. Similarly, we can define the derivate of a multi variable function with respect to (w.r.t.) one of its variables, this is known as **partial derivative** and it is defined as,

$$\left(\frac{\partial f(x,y)}{\partial x}\right)_y = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}, \tag{32}$$

where $(\ )_y$ means that the variable y is considered as a constant. Commonly, we do not use this notation $(\ )_y$ as for functions with large number of variables it makes the writing cumbersome, so we define $\frac{\partial f(x,y)}{\partial x} = \left(\frac{\partial f(x,y)}{\partial x}\right)_y$.

Eq. 32 is analogous to the definition of the ordinary derivative (Eq. 1). The only difference is the different used symbol, $\frac{d}{dx} \to \frac{\partial}{\partial x}$. We can also consider the partial derivative of $f(x, y)$ but w.r.t. y,

$$\left(\frac{\partial f(x,y)}{\partial y}\right)_x = \frac{\partial f(x,y)}{\partial y} = \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}. \tag{33}$$

The list of ordinary derivatives can be used also in partial derivatives.

Exercise: Let's compute the partial derivative of $f(x, y) = ax^2 y^3 + be^{xy}$ w.r.t. x and y,

1. w.r.t. x

$$\begin{aligned}
\frac{\partial f(x,y)}{\partial x} &= \frac{\partial}{\partial x} ax^2 y^3 + \frac{\partial}{\partial x} be^{xy} = ay^3 \frac{\partial}{\partial x} x^2 + b\frac{\partial}{\partial x} e^{xy} & (34) \\
&= ay^3(2x) + bye^{xy} = 2axy^3 + bye^{xy} & (35) \\
& & (36)
\end{aligned}$$

2. w.r.t. y

$$\begin{aligned}
\frac{\partial f(x,y)}{\partial y} &= \frac{\partial}{\partial y} ax^2 y^3 + \frac{\partial}{\partial y} be^{xy} = ax^2 \frac{\partial}{\partial y} y^3 + b\frac{\partial}{\partial y} e^{xy} & (37) \\
&= ax^2(3y^2) + bxe^{xy} = 3ax^2 y^2 + bxe^{xy} & (38) \\
& & (39)
\end{aligned}$$

- Observe that $\frac{\partial f(x,y)}{\partial x} \neq \frac{\partial f(x,y)}{\partial y}$.

If the function $f(x, y)$ now undergoes an infinitesimal change in x and y, we can define the **total differential** of $f(x, y)$ as,

$$df(x, y) = \left(\frac{\partial f(x,y)}{\partial x}\right)_y dx + \left(\frac{\partial f(x,y)}{\partial y}\right)_x dy. \tag{40}$$

Eq. 40 often appears in thermodynamics and metrology.

3.2 Partial derivative identities

Here we will derive two identities for partial derivatives, i) **reprocity relation** and ii) **cyclic relation**. These two relations are widely used in themordynamics, see Bridgman's thermodynamic equations. The following derivations and examples are based on the following links, triple product rule and exact differential.

The variables x, y and z are bounded by some differentiable function $f(x, y, z)$, the following total differentials (Eq. 40) exist,

$$dx = \left(\frac{\partial x}{\partial y}\right)_z dy + \left(\frac{\partial x}{\partial z}\right)_y dz; \qquad \text{meaning} \qquad x = f(y, z) \tag{41}$$

$$dz = \left(\frac{\partial z}{\partial x}\right)_y dx + \left(\frac{\partial z}{\partial y}\right)_x dy; \qquad \text{meaning} \qquad z = f(x, y) \tag{42}$$

Substituting the equation of dx into the equation for dz and rearranging, we get,

$$dz = \left(\frac{\partial z}{\partial x}\right)_y \underbrace{\left[\left(\frac{\partial x}{\partial y}\right)_z dy + \left(\frac{\partial x}{\partial z}\right)_y dz\right]}_{dx} + \left(\frac{\partial z}{\partial y}\right)_x dy \tag{43}$$

$$dz - \left(\frac{\partial z}{\partial x}\right)_y \left(\frac{\partial x}{\partial z}\right)_y dz = \left(\frac{\partial z}{\partial x}\right)_y \left(\frac{\partial x}{\partial y}\right)_z dy + \left(\frac{\partial z}{\partial y}\right)_x dy \tag{44}$$

$$\left[1 - \left(\frac{\partial z}{\partial x}\right)_y \left(\frac{\partial x}{\partial z}\right)_y\right] dz = \left[\left(\frac{\partial z}{\partial x}\right)_y \left(\frac{\partial x}{\partial y}\right)_z + \left(\frac{\partial z}{\partial y}\right)_x\right] dy \tag{45}$$

The left-hand and the right-hand side depend on dz and dy, selectively, and since y and z are independent variables, dy and dz may be chosen without restriction. For this last equation to be valid in general, the bracketed terms must be equal to zero.

Reprocity relation: Let start with the bracket from the left-hand side in Eq. 45.

$$\left[1 - \left(\frac{\partial z}{\partial x}\right)_y \left(\frac{\partial x}{\partial z}\right)_y\right] = 0 \tag{46}$$

$$\left(\frac{\partial z}{\partial x}\right)_y \left(\frac{\partial x}{\partial z}\right)_y = 1 \tag{47}$$

We can also rearrange this equation to get the reciprocity relation,

$$\left(\frac{\partial z}{\partial x}\right)_y = \frac{1}{\left(\frac{\partial x}{\partial z}\right)_y} \tag{48}$$

• Reciprocity relations show that the inverse of a partial derivative is equal to its reciprocal.

Example: Latter in the course we will revise the ideal gas law that relates pressure (P), volume (V), temperature (T), and number of moles (n); $PV = nRT$. We can define,

$$T = f(P, V) = \frac{PV}{nR}, \qquad \left(\frac{\partial T}{\partial P}\right)_V = \frac{V}{nR} \tag{49}$$

$$P = f(V, T) = \frac{nRT}{V}, \qquad \left(\frac{\partial P}{\partial T}\right)_V = \frac{nR}{V} \tag{50}$$

$$\tag{51}$$

from these two equations we can see that

$$\left(\frac{\partial T}{\partial P}\right)_V \left(\frac{\partial P}{\partial T}\right)_V = \left(\frac{V}{nR}\right)\left(\frac{nR}{V}\right) = 1, \tag{52}$$

meaning the reciprocity relation holds for $PV = nRT$.

4 Higher-order Derivatives

Functions can also have higher-order derivatives, for example, the kinetic energy is the second derivative of the position. $f(x, y)$ has four different second partial derivatives. Let's first consider the partial derivatives w.r.t. x and y of $\frac{\partial f(x,y)}{\partial x}$; this partial derivative is also a function of x and y.

$$\frac{\partial}{\partial x}\frac{\partial f(x,y)}{\partial x} = \frac{\partial^2 f(x,y)}{\partial x^2} \quad \text{and} \quad \frac{\partial}{\partial y}\frac{\partial f(x,y)}{\partial x} = \frac{\partial f(x,y)}{\partial y \partial x} \tag{53}$$

The partial derivatives of $\frac{\partial f(x,y)}{\partial y}$ w.r.t. x and y are, $f(x, y)$ has four different second partial derivatives. Let's first consider the partial derivatives w.r.t. x and y of $\frac{\partial f(x,y)}{\partial x}$; this partial derivative is also a function of x and y.

$$\frac{\partial}{\partial x}\frac{\partial f(x,y)}{\partial y} = \frac{\partial f(x,y)}{\partial x \partial y} \quad \text{and} \quad \frac{\partial}{\partial y}\frac{\partial f(x,y)}{\partial y} = \frac{\partial^2 f(x,y)}{\partial y^2} \tag{54}$$

Later in the course we will see that for multivariate functions, the second oder derivative is known as the **Hessian**,

$$\mathbf{H}_{f(x,y)} = \begin{pmatrix} \frac{\partial^2 f(x,y)}{\partial x^2} & \frac{\partial f(x,y)}{\partial y \partial x} \\ \frac{\partial f(x,y)}{\partial x \partial y} & \frac{\partial^2 f(x,y)}{\partial y^2} \end{pmatrix} \tag{55}$$

.

Exercise: Let's compute the Hessian of $f(x, y) = ax^2 y^3 + b e^{xy}$,

$$\frac{\partial^2 f(x,y)}{\partial x^2} = \frac{\partial}{\partial x} 2axy^3 + bye^{xy} = 2ay^3 + by^2 e^{xy} \tag{56}$$

$$\frac{\partial f(x,y)}{\partial y \partial x} = \frac{\partial}{\partial y} 3ax^2 y^2 + bxe^{xy} = 6axy^2 + be^{xy}(1 + xy) \tag{57}$$

$$\frac{\partial f(x,y)}{\partial x \partial y} = 6axy^2 + be^{xy}(1 + xy) \tag{58}$$

$$\frac{\partial^2 f(x,y)}{\partial y^2} = 6ax^2 y + bx^2 e^{xy} \tag{59}$$

Form the prev. equations, we can notice that $\frac{\partial f(x,y)}{\partial y \partial x} = \frac{\partial f(x,y)}{\partial x \partial y}$. This property of the partial derivatives is known as the *Clairaut's theorem*.

"Proof": that $\frac{\partial}{\partial x}\frac{\partial}{\partial y} = \frac{\partial}{\partial y}\frac{\partial}{\partial x}$,

$$\frac{\partial}{\partial x}\frac{\partial}{\partial y} = \frac{\partial}{\partial x}\underbrace{\left(\lim_{\Delta y \to 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}\right)}_{\text{function}} \tag{60}$$

$$= \lim_{\Delta x \to 0}\left(\lim_{\Delta y \to 0} \frac{f(x + \Delta x, y + \Delta y) - f(x + \Delta x, y)}{\Delta x \Delta y} - \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta x \Delta y}\right)$$

$$= \lim_{\Delta x \to 0}\lim_{\Delta y \to 0} \frac{f(x + \Delta x, y + \Delta y) - f(x + \Delta x, y) - f(x, y + \Delta y) + f(x, y)}{\Delta x \Delta y}$$

$$\frac{\partial}{\partial y}\frac{\partial}{\partial x} = \frac{\partial}{\partial y}\underbrace{\left(\lim_{\Delta x \to 0}\frac{f(x+\Delta x, y) - f(x,y)}{\Delta x}\right)}_{\text{function}} \tag{61}$$

$$= \lim_{\Delta y \to 0}\left(\lim_{\Delta x \to 0}\frac{f(x+\Delta x, y+\Delta y) - f(x, y+\Delta y)}{\Delta y \Delta x} - \lim_{\Delta y \to 0}\frac{f(x+\Delta x, y) - f(x,y)}{\Delta y \Delta x}\right)$$

$$= \lim_{\Delta x \to 0}\lim_{\Delta y \to 0}\frac{f(x+\Delta x, y+\Delta y) - f(x, y+\Delta y) - f(x+\Delta x, y) + f(x,y)}{\Delta y \Delta x}$$

Example: What is the Hessian of $V(x) = \frac{1}{2}k(x - R_0)^2$?

$$\mathbf{H} = \frac{\partial^2 V(x)}{\partial x^2} = k \tag{62}$$

Example: What is the Hessian of $V(x_1, x_2) = \frac{1}{2}k_1(x_1 - R_0)^2 + \frac{1}{2}k_2(x_2 - R_0)^2$?

$$\mathbf{H} = \begin{pmatrix} \frac{\partial^2 V(x_1, x_2)}{\partial x_1^2} & 0 \\ 0 & \frac{\partial^2 V(x_1, x_2)}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix} \tag{63}$$

Force Constants: The constants k_1 and k_2 represent the "stiffness" of the bonds, i.e., how strongly the potential energy increases as the bond lengths deviate from R_0. Different values of k_1 and k_2 indicate that the bonds to the two atoms may have different stiffness (e.g., if one bond is with a heavier or lighter atom).

5 Numerical differentiation

There are cases where derivatives do not have analytical solutions and one must estimate them numerically.

1. A simple two-point estimation using the definition in Eq. 1

$$\frac{df(x)}{dx} \approx \frac{f(x+\Delta x) - f(x)}{\Delta x} \tag{64}$$

2. Symmetric difference quotient

$$\frac{df(x)}{dx} \approx \frac{f(x+\Delta x) - f(x-\Delta x)}{2\Delta x} \tag{65}$$