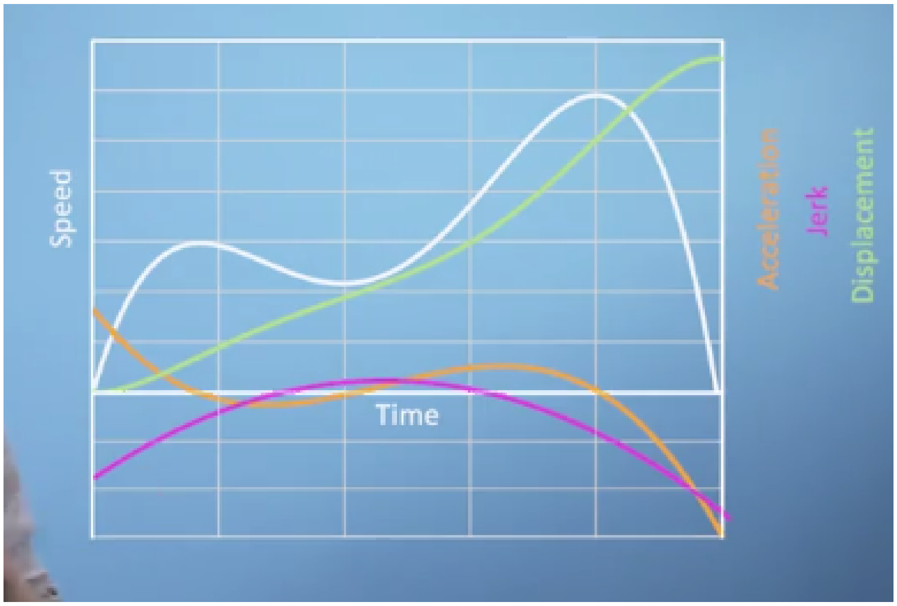
Week - 01:

Derivatives:

The video is introducing **calculus** as a way to understand how things change—specifically how one quantity (like speed) changes with another (like time). It's using the example of a car's motion to make this idea more understandable.



---

1. Speed-Time Graph:

- Imagine a graph where **speed** is on the vertical axis and **time** is on the horizontal axis.
- If the car was moving at the **same speed the whole time**, the graph would be a **flat horizontal line**.
- But in this case, the graph is **not flat**, which means the speed is changing — the car is **accelerating and decelerating**.

---

2. Acceleration from the Graph:

- The **slope** (or steepness) of the speed-time graph at any point tells us the **acceleration** at that point.
    - If the line is going **up**, the car is **accelerating**.
    - If the line is going **down**, the car is **decelerating**.

- The slope at a single point is called the **local gradient**, and it can be visualized with a **tangent line** (a straight line that just touches the curve).

---

3. Acceleration-Time Graph:

- By calculating the slope at every point on the speed-time graph, we can create a **new graph** that shows **acceleration versus time**.
- If a car moves at **constant speed**, the speed-time graph is flat (slope = 0), so the acceleration-time graph would also be a flat line **at zero**.
- In more complex cases:
  - Acceleration is **positive** when speed increases.
  - **Zero** when speed stops increasing (i.e., reaches its peak).
  - **Negative** when the car slows down.

---

4. Understanding Derivatives:

- What we've done by finding the slope of the speed graph is essentially taking its **derivative**.
- The **derivative** is a new function that tells us the **rate of change** (how fast something changes).
- In this case:
  - Speed is the derivative of **distance**.
  - Acceleration is the derivative of **speed**.
  - **Jerk** is the derivative of **acceleration** (how quickly acceleration is changing).

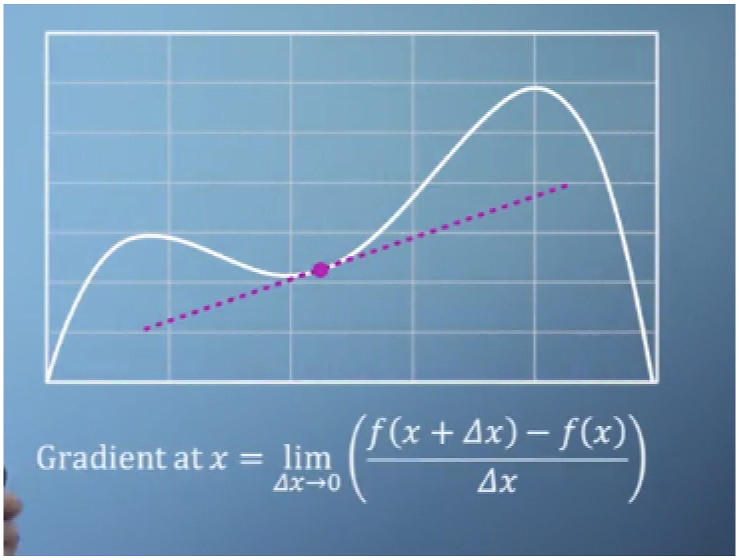
---

5. Going Backwards – Anti-derivative:

- You can also go **in reverse**:
  - If you have a speed-time graph, and you ask: "What function would give this speed when I take its slope?" — you're looking for the **anti-derivative** or **integral**.

- In this example, the anti-derivative of speed is **distance**.
    - Why? Because the rate of change of distance over time (its slope) is speed.

Definition of a derivative:



📜 Key Concepts

1. **Gradient (Slope)**
    - For straight lines:

$$\text{Gradient} = \frac{\text{Rise}}{\text{Run}} = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

2. **Derivative Definition**
    - For curved functions, gradient at a point is:

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

    - This is the **formal definition of a derivative.**

1. Power Rule

Used to differentiate functions like $f(x) = ax^n$

📌 **Formula:**

$$\frac{d}{dx}[ax^n] = a \cdot n \cdot x^{n-1}$$

🧠 **Example:**

$$\frac{d}{dx}[5x^3] = 15x^2$$

➕ 2. Sum Rule

Used when differentiating a **sum or difference** of functions.

📌 **Formula**:

$$\frac{d}{dx}[f(x) + g(x)] = f'(x) + g'(x)$$

🧠 **Example:**

$$\frac{d}{dx}[3x^2 + 2x] = 6x + 2$$

✖️ 3. Product Rule

Used when differentiating the **product** of two functions.

📌 **Formula**:

$$\frac{d}{dx}[f(x) \cdot g(x)] = f'(x)g(x) + f(x)g'(x)$$

🧠 **Example:**

If $f(x) = x^2$, $g(x) = \sin(x)$, then:

$$\frac{d}{dx}[x^2 \sin(x)] = 2x \sin(x) + x^2 \cos(x)$$

🔗 4. Chain Rule

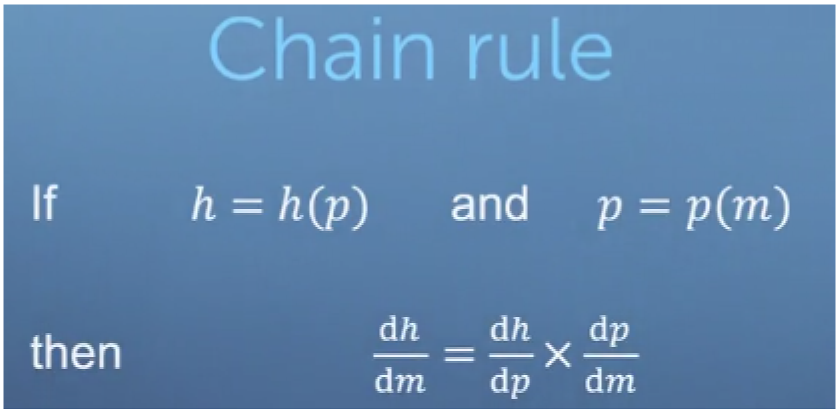
Used to differentiate **composite functions** (function inside a function).

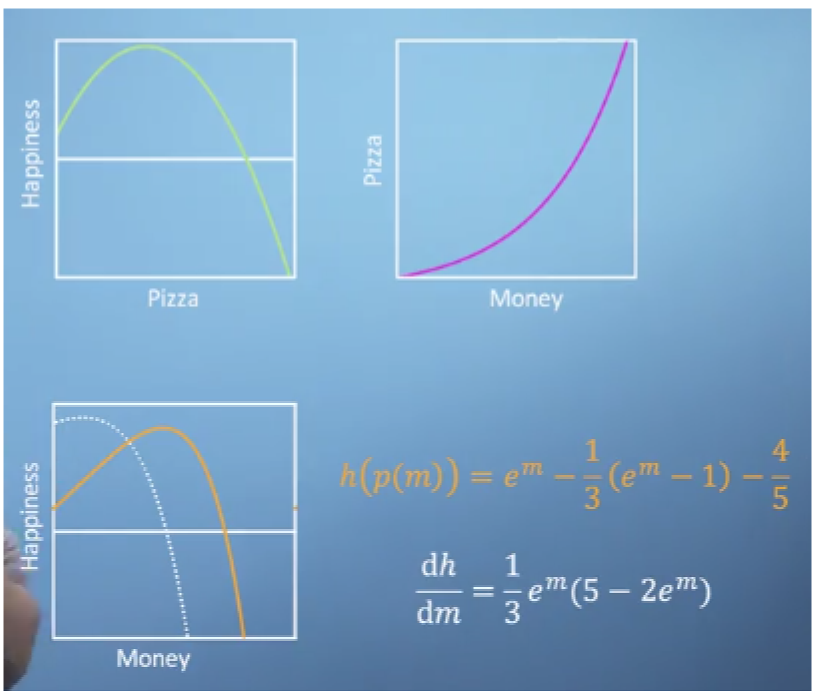
📌 **Formula**:
If $f(x) = g(h(x))$, then:

$$f'(x) = g'(h(x)) \cdot h'(x)$$

🧠 **Example**:

$$\frac{d}{dx}[\sin(x^2)] = \cos(x^2) \cdot 2x$$





Week - 02:

Multivariable systems:

Now, we're moving into **multivariable systems**, where more than one variable affects the outcome — and we use **partial derivatives**.

🧩 What's a Variable, Constant, and Parameter?

- **Dependent variable**: Changes *because* of something else (e.g., speed depends on time).
- **Independent variable**: You choose or control it (e.g., time just passes).
- **Constants**: Don't change in your current context (e.g., car's mass).
- **Parameters**: Variables that can change *across different scenarios* (like drag coefficient for different car models) — not always changing *in the moment*, but adjustable in design.

So, **what's a constant or a variable depends on the context.** Something that is constant in one problem might be a variable in another.

🧪 Partial Derivatives: A Way to Handle Multivariable Systems

In multivariate calculus, functions depend on **many variables**.

Example: Imagine you're calculating the **mass** of a can using its:

- Radius r
- Height h
- Wall thickness t
- Material density ρ

The formula for mass:

$$m = \left(2\pi r^2 + 2\pi r h\right) \cdot t \cdot \rho$$

Now, if you want to know **how mass changes when you only increase height** h, keep the others constant. That's called a **partial derivative**:

$$\frac{\partial m}{\partial h} = 2\pi r t \rho$$

This tells you: for every unit increase in height, the mass increases by this much (assuming r, t, ρ stay the same).

📌 Key Takeaways

- **Context matters**: Constants, variables, and parameters shift depending on the situation.

- **Partial differentiation** means taking the derivative with respect to one variable while treating others as constants.
- It's just like normal derivatives, but in higher dimensions.
- Use the ∂ **(curly d)** symbol to show it's a partial derivative.

1. A Trickier Example Function

We're given a multivariable function:

$$f(x, y, z) = \sin(x) \cdot e^{yz^2}$$

Let's find partial derivatives one at a time:

🔷 Partial Derivative with respect to x:

Only $\sin(x)$ involves x. Treat e^{yz^2} as constant.

$$\frac{\partial f}{\partial x} = \cos(x) \cdot e^{yz^2}$$

🔷 Partial Derivative with respect to y:

Only yz^2 has y. Use the **chain rule** on the exponential.

$$\frac{\partial f}{\partial y} = \sin(x) \cdot e^{yz^2} \cdot z^2$$

🔷 Partial Derivative with respect to z:

Again, use the **chain rule**, since yz^2 involves z.

$$\frac{\partial f}{\partial z} = \sin(x) \cdot e^{yz^2} \cdot 2yz$$

🧮 Introducing the Total Derivative

Now imagine:

The variables x, y, z are each functions of a **new variable** t:

- $x = t - 1$
- $y = t^2$
- $z = \frac{1}{t}$

We want to find how the entire function f changes as t changes — this is the **total derivative**:

$$\frac{df}{dt}$$

🧠 Why Not Just Plug In and Differentiate?

Yes, you *could* substitute all of $x(t), y(t), z(t)$ into the original f and then differentiate. But if the function is **huge**, that can get messy.

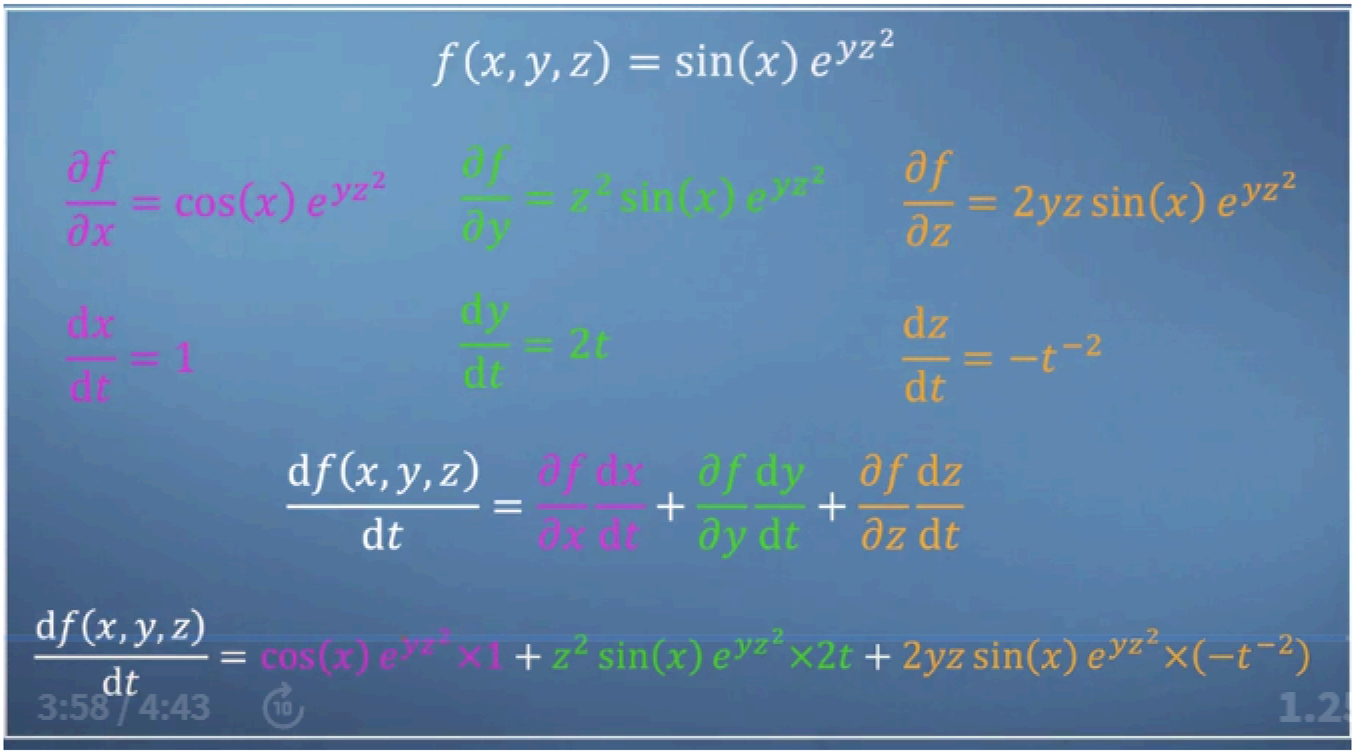
Instead, use the **chain rule** across all variables.

🔗 Total Derivative Formula:

$$\frac{df}{dt} = \frac{\partial f}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dt} + \frac{\partial f}{\partial z} \cdot \frac{dz}{dt}$$

This means:

Add up how much f changes with each variable, **multiplied** by how much that variable changes with t.



✅ Key Takeaways:

- **Partial derivatives** look at one variable at a time.
- The **total derivative** looks at how a function changes when **all its inputs** are changing.
- It uses the **chain rule** across multiple variables.
- Very handy when variables depend on another variable like time, space, etc.

🌟 What Is the Jacobian?

The **Jacobian** is a fancy name for something simple:

> It's a **vector of all partial derivatives** of a function with respect to its input variables.

If you have a function like:

$$f(x_1, x_2, x_3, \dots)$$

Then the **Jacobian** is just:

$$J = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3}, \dots \right)$$

🔷 **By convention**, we write this as a **row vector**.

📘 Example: A Simple Function

Given:

$$f(x, y, z) = x^2y + 3z$$

Let's build the **Jacobian** step by step:

1. $\frac{\partial f}{\partial x} = 2xy$
2. $\frac{\partial f}{\partial y} = x^2$
3. $\frac{\partial f}{\partial z} = 3$

So the **Jacobian** is:

$$J = (2xy, x^2, 3)$$

🔍 What Does the Jacobian Tell Us?

Think of the Jacobian as:

- A **gradient vector**,
- That points in the direction of the **steepest increase** in the function,
- With each component showing how **sensitive** the function is to changes in that particular variable.

✨ Example:

At point (0,0,0) the Jacobian is:

$$J(0,0,0)=(0,0,3)$$

This means the function **only increases in the z direction** at that point — the x and y directions have no effect there.

🧭 Visualization and Intuition

To help build intuition, the video walks through a more **complex 2D function**. Here's how they visualized it:

1. Color plot

Shows high values (bright) and low values (dark) of the function.

2. 3D plot

Adds **height** for the function value (z) — a landscape with hills and valleys.

3. Contour plot

Like a topographic map. Lines represent **equal height** (constant z-values). Close lines mean steep slope.

4. Add Jacobian vectors

At every point, add an arrow showing:

- The direction of **steepest uphill**
- The **length** of the arrow represents how steep the slope is

---

✅ Key Observations from the Visualization

- Arrows **always point uphill**, toward higher function values
- Steeper regions (where contour lines are close together) → **longer Jacobian vectors**
- Flat regions or peaks/valleys → **short or zero Jacobian vectors**

🔚 Final Thoughts

- The **Jacobian** gives you **direction + magnitude** of steepest change
- It's **just partial derivatives** grouped into a vector — nothing too scary!
- Visually, it helps you **navigate a function's landscape**
- In high dimensions, we can't visualize it — but the **math still works**, and that's powerful

📌Jacobian as a Vector (Gradient)

The core idea here is that the **Jacobian** can describe not just how a **scalar function** changes with respect to its inputs, but also how a **vector-valued function** changes — and that's when the Jacobian becomes a **matrix**.

We begin with a **scalar function**:

$$f(x, y) = e^{-x^2+y^2}$$

To find the **Jacobian vector**, we calculate the partial derivatives:

$$\frac{\partial f}{\partial x}, \quad \frac{\partial f}{\partial y}$$

This gives a **2D vector** (a gradient), pointing in the direction of **steepest increase** of the function.

🔍 Example: Checking Specific Points

- At **(-1, 1)** → Jacobian points toward the origin
- At **(2, 2)** → Smaller vector, still points toward the origin
- At **(0, 0)** → Zero vector → function is flat (this could be a **maximum**, **minimum**, or **saddle point**)

When you look at the **whole vector field**, you see that vectors all point **towards the origin**, suggesting it is a **maximum**.

Visualized in 3D → function is like a hill, highest at the origin.

📌 Jacobian as a **Matrix** (For Vector-Valued Functions)

We now move on to **functions that return vectors**, not just single numbers.

Suppose:

- $u(x, y) = x + 2y$
- $v(x, y) = 3y - 2x$

This defines a new **vector-valued function**:

$$F(x, y) = \begin{bmatrix} u(x, y) \\ v(x, y) \end{bmatrix}$$

🧮 How to Build the Jacobian Matrix

For vector functions, the **Jacobian becomes a matrix** where:

- Each **row** is the gradient of one output function.
- Each **column** corresponds to an input variable.

So the **Jacobian matrix** is:

$$J = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix}$$

This matrix **transforms** a vector in the (x, y) space into one in the (u, v) space.

🧪 Test It Out:

Given $(x, y) = (2, 3)$, apply the transformation:

$$\begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \cdot 2 + 2 \cdot 3 = 8 \\ -2 \cdot 2 + 3 \cdot 3 = 5 \end{bmatrix} \Rightarrow (u, v) = (8, 5)$$

📌 Part 3: What If the Functions Are Nonlinear?

In real-world problems, the functions are usually **nonlinear**.

But if a function is **smooth** (no jumps or sharp corners), then even though it's nonlinear overall, we can:

> Zoom into a small region and **approximate it as linear** using the Jacobian.

This is **super important** in areas like optimization and machine learning — we repeatedly approximate nonlinear functions with linear ones during updates (like in gradient descent).

Part 4: Real Example – Polar to Cartesian

Switching coordinates is a common transformation. For instance:

- x = r cosθx
- y = r sinθy

We can build the **Jacobian matrix** of this transformation, and compute its **determinant**:

$$J = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}$$

The **determinant** of this matrix is:

$$\det(J) = r$$

This tells us that:

- As you move away from the origin (increasing rrr), **area scales by rrr**.
- The transformation is "stretching space" more as you move further out.

✅ Summary – What You Should Now Understand

Concept	Meaning
Jacobian Vector	For scalar-valued functions: points in direction of steepest slope (gradient)
Jacobian Matrix	For vector-valued functions: linear approximation of how output vectors change with inputs
Use of Jacobians	Crucial for optimization, coordinate transformations, and machine learning algorithms
Determinant of Jacobian	Tells how **space is scaled** under transformation (important in integration and probability)