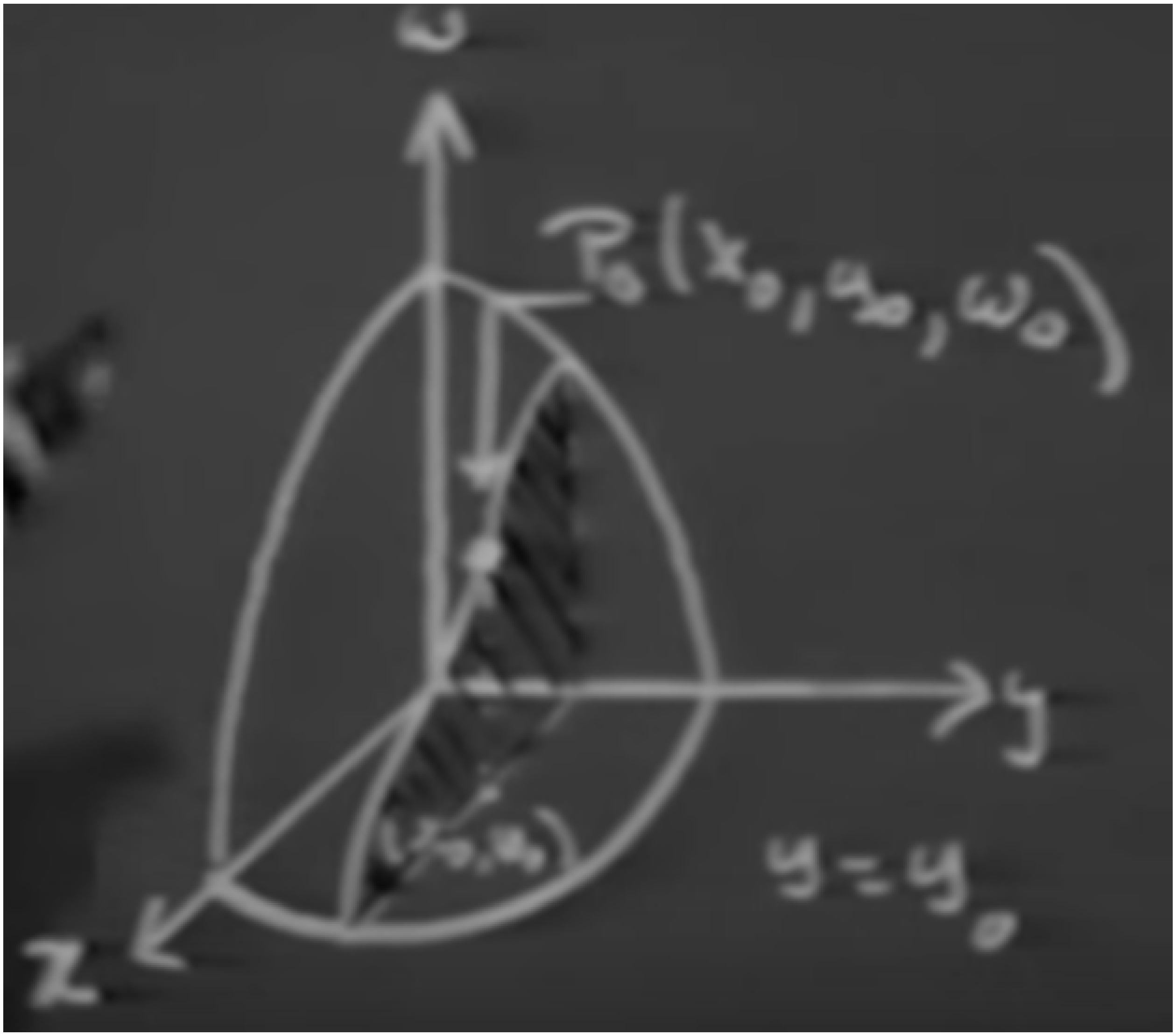
partial derivatives

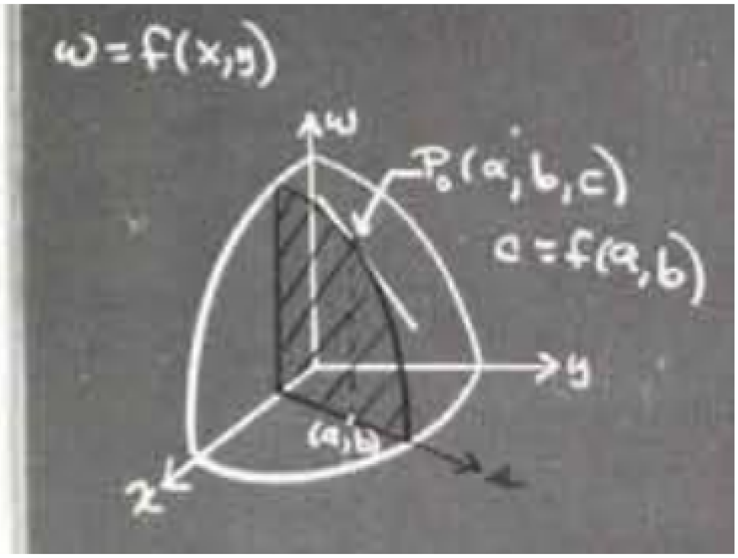
- Now, lets move our attention to functions like this, which map a vector to a scalar:
  $f(\vec{V}) = x^2 + y$ where $\vec{V} = (x, y)$
- Notice that an alternative way to think about the above function is just a function that takes in an ordered pair (x, y) and not necessarily as a vector $\vec{V} = (x, y)$.
- More generally, these are n-tuples. An n-tuple is a ordered list of numbers. Order matters. Note how these are basically vectors.
- structure (the rules of the game) of n-tuples (n-dimensional vector space): these are basically inspired by vectors but are extended to n-dimensional space where n is a positive integer.
  - $(a_1, a_2, ...a_n) = (b_1, b_2, ...b_n)$ if and only if $a_k = b_k$.
  - $(a_1, a_2, ...a_n) + (b_1, b_2, ...b_n) = (a_1 + b_1, a_2 + b_2, ...a_k + b_k)$
  - $c(a_1, ...a_n) = (ca_1, ...ca_n)$

  These all exist even though they cannot be visualized as arrows. Notice that this definition is a generalization.
- limit definition:
  $\lim_{(x_1, x_2, ...x_n) \to (a_1, a_2, ...a_n)} f(x_1, ..x_n) = L$ means that given $\epsilon > 0$, we can find a $\delta > 0$ such that $0 < \sqrt{(x_1 - a_1)^2 + ..(x_n - a_n)^2} < \delta$ → $|f(x_1...x_2) - L| < \epsilon$.
  Notice how we defined a notion of a distance here for n-dimensional space.
- in the above definition, in the 2-D case, note how the neighborhood is really a region and not a linear neighborhood.
- the number of independent variables is the same as the number of degrees of freedom. notice that independent variables can be varied without changing other independent variables. this leads us to define something called a partial derivative.
- partial derivative definition:
  $f_{x_1}(x_1, ..x_n) = lim_{\Delta x_1 \to 0} \frac{f(x_1 + \Delta x_1, x_2..x_n) - f(x_1, x_2..x_n)}{\Delta x_1}$
  Notice the notation f_{x_1} instead of f' for the derivative.
- Note that this can only be done provided the variables are independent. If varying x_1 varied any of $x_2, x_3...x_n$, then how can we vary x_1 without varying another variable.
- We write f_{x_1} as $\frac{\partial f}{\partial x_1}$. Notice the funny d → it indicates partial derivative.
- Because our structure (rules of the game) for partial derivatives so closely parallels that of the “usual” derivative, the same properties hold like product rule, quotient rule, chain rule etc.

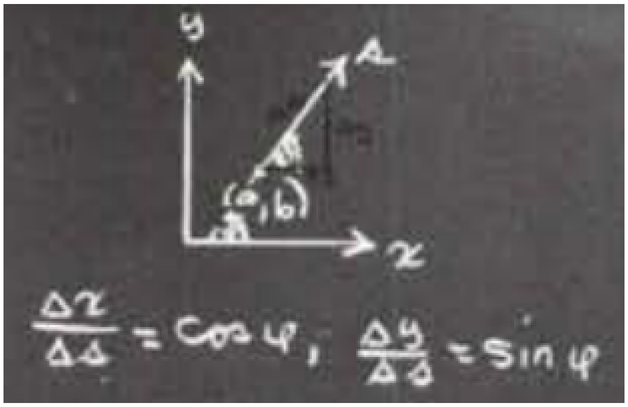


Taking the derivative w.r.t x at a fixed y is the slope of the curve formed by the intersection of the surface and the plane y = k

- When taking partial derivatives w.r.t one variable, we consider a very special plane. Later we will look at derivatives that are slopes of curves given by intersection of the surface with planes in other directions.
- Given f_x and f_y the *tangent plane is the plane that these two vectors define*. to find this, just take the cross product of these two to find the normal. Once we have the normal and a point that lies on the plane, we know how to find the equation of the plane.
- For 3-D, with x and y as independent variables, following above method, we can show that the equation of the plane is $\frac{\partial f}{\partial x}(x - x_0) + \frac{\partial f}{\partial y}(y - y_0) + (z - z_0) = 0$ → $\Delta z = \frac{\partial f}{\partial x}(x - x_0) + \frac{\partial f}{\partial y}(y - y_0)$. Note that this Δz is for the plane, not for the surface. But for close enough points, this can be an approximator, as will be proved soon.
- We have seen f_x and f_y. Now we will look at derivatives that are slopes of curves given by intersection of the surface with planes in other directions. These are *not* called partial derivatives, just derivatives.



- $\frac{dw}{ds} = lim_{\Delta s \to 0} \frac{\Delta w}{\Delta s}$ is the derivative we are interested in. Note that Δs is the hypotenuse such that Δx and Δy are the two sides.



View from the top of x, y and s. s is the hypotenuse of the right triangle.

- Just as having a limit in the x- and y-coordinate directions does not imply the function itself has a limit at (x_0, y_0), so is it the case that the existence of both partial derivatives is not enough by itself to ensure derivatives exist for trace curves in other non-special vertical planes.
- A function $z = f(x, y)$ is differentiable at (x_0, y_0) if $f_x(x_0, y_0)$ and $f_y(x_0, y_0)$ exist and Δz satisfies an equation of the form

  $$\Delta z = f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y + \epsilon_1 \Delta x + \epsilon_2 \Delta y$$

  in which each of $\epsilon_1, \epsilon_2 \to 0$ as both $\Delta x, \Delta y \to 0$.
- We call f differentiable if it is differentiable at every point in its domain, and say that its graph is a smooth surface.
- Turns out, in order for it to be differentiable and for both ϵ_1, ϵ_2 to go to 0, we need f_x, f_y to exist in the neighborhood of (a, b) and f_x, f_y must be continuous at (a, b). To understand this, visualize how f_x and f_y existing in the neighborhood means that "smooth" approaches exist to get to (a, b) thru different paths.

- From the above definition, dividing thru by Δs, $\frac{\Delta w}{\Delta s} = f_x \frac{\Delta x}{\Delta s} + f_y \frac{\Delta y}{\Delta s} + + \frac{\epsilon_1 \Delta x}{\Delta s} + \frac{\epsilon_2 \Delta y}{\Delta s}$. Taking $\lim_{\Delta s \to 0}$, we get $\frac{dw}{ds} = f_x cos(\phi) + f_y sin(\phi) + lim_{\Delta s \to 0} \epsilon_1 + lim_{\Delta s \to 0} \epsilon_2 =$
  which means
  $\frac{dw}{ds} = f_x cos(\phi) + f_y sin(\phi) = (f_x, f_y) \cdot (cos(\phi), sin(\phi))$
- Because the directional derivative is this dot product, it takes on its maximum value when the vector $cos(\phi), sin(\phi)$ is in the same direction as the (f_x, f_y) vector. Thus the gradient is maximum in the direction of this vector → and so this is called the gradient vector. This gradient vector is represented as $\nabla \vec{f}$.
- chain rule:
  suppose $w = f(x, y, z)$ and $x = x(r, s), y = y(r, s), z = z(r, s)$ and hence $w = g(r, s)$
  if these are all partially differentiable and their partial derivatives are continuous, we can write $\Delta w = \frac{\partial w}{\partial x} \Delta x + \frac{\partial w}{\partial y} \Delta y + \frac{\partial w}{\partial z} \Delta z + k_1 \Delta x + k_2 \Delta y + k_3 \Delta z$ where $k_1, k_2, k_3 \to 0$ as $\Delta x, \Delta y, \Delta z \to 0$.
  Dividing by Δr and taking the limit as $\Delta r \to 0$, we get:
  $\frac{\partial w}{\partial r} = \frac{\partial w}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial w}{\partial y} \frac{\partial y}{\partial r} + \frac{\partial w}{\partial z} \frac{\partial z}{\partial r}$
- Notice that in the derivation of the chain rule above, we need the condition that the partials exist and are continuous functions.
- More generally, $\frac{\partial w}{\partial x_1} = \sum_{k=1}^{n} \frac{\partial w}{\partial y_k} \frac{\partial y_k}{\partial x_1}$
- Some notation: $\frac{\partial^2 w}{\partial x \partial y} = \frac{\partial}{\partial x} \frac{\partial w}{\partial y}$. In general, $\frac{\partial^2 w}{\partial x \partial y} \neq \frac{\partial^2 w}{\partial y \partial x}$
- The word parametric means something like this: You have a variable but you assigned it a value and now its no more a variable. The following integral is an example of parametric integral:
  $\int_a^b f(x, y) dy$, at $x = x_0$
- Now, if we wanted to compute $\frac{d}{dx} \int_a^b f(x, y) dy$, at $x = x_0$, this would be $\lim_{h \to 0} \frac{\int_a^b f(x+h, y) dy - \int_a^b f(x, y) dy}{h}$ $= \lim_{h \to 0} \frac{(\int_a^b f(x+h, y) - f(x, y)) dy}{h} = \lim_{h \to 0} \int_a^b \frac{(f(x+h, y) - f(x, y))}{h} dy$.
- *If we can interchange the limit and the integral,* we get $\int_a^b f_x dy$. So assuming, we can do the interchange, we get
  $\frac{d}{dx} \int_a^b f(x, y) dy = \int_a^b f_x dy$
  To see why we can interchange the derivative and the integral, think about the fact that gradients are a linear function: the derivative of sums is the sum of derivatives. if you view the integral as a sum, using the linearity of derivatives you can see why we can exchange them!
- We call it variable limits of integration if we have something like
  $\int_{f(x)}^{g(x)} f(x, y) dy$. Notice that the integration limits are functions of x too. But remember, x is a parameter and its value gets fixed before you do the integration. So its a variable in the sense that it is a parameter and once you give it a value, the limits get fixed. But while doing $\frac{d}{dx} \int_{f(x)}^{g(x)} f(x, y) dy$, the limits of integration change since we take x and $x + h$.
  But if we write $\frac{d}{dx} \int_{f(x)}^{g(x)} f(x, y) dy = h(u, v, x)$ such that $u = g(x), v = f(x)$. Now
  $\frac{d}{dx} \int_{f(x)}^{g(x)} f(x, y) dy = \frac{d}{dx} h(u, v, x)$
  Now we can use the chain rule! Now,
  $\frac{d}{dx} h(u, v, x) = h_u \frac{du}{dx} + h_v \frac{dv}{dx} + \frac{\partial u}{\partial x}$. Now, h_u is like computing the definite integral as $F(x, u) - F(x, v)$ and then taking the derivative, which gives us back $f(x, u)$. Similarly, h_v is $-f(x, v)$.
- So now, we can write $\frac{d}{dx} \int_{f(x)}^{g(x)} f(x, y) dy$ as $h_u \frac{du}{dx} + h_v \frac{dv}{dx} + \frac{\partial u}{\partial x}$, which is just
  $f(x, u) u'(x) - f(x, v) v'(x) + \int_{f(x)}^{g(x)} f_x(x, y) dy$.

- The reason we did this derivation, even though there is nothing new here, is because Herb thinks this form of derivation shows up in a lot of places and he wanted us to see it here. What a nice guy <3
- Really, the above derivation is just the chain rule.
- We saw that if w is differentiable, by definition, we can write $dw = f_x dx + f_y dy$.
  Definition: $Mdx + Ndy$ is called an exact differential if and only if there exists $w = f(x, y)$ such that $dw = Mdx + Ndy$. That means we must have $f_x = M$ and $f_y = N$.
- So if we are given an exact differential $Mdx + Ndy$ and we want to find w, we can integrate M with respect to x (the constant of integration will be $g(y)$, some function of y. Now if we differentiate this result with respect to y (partial derivative), we should get N.
- Now notice that if w is “continuous enough”, then $w_{xy} = \frac{\partial^2 w}{\partial y \partial x}$ will equal $w_{yx} = \frac{\partial^2 w}{\partial x \partial y}$. Because of this, if w is “continuous enough”, given $Mdx + Ndy$, we can just check if $M_y = N_x$ to figure out if $Mdx + Ndy$ is an exact differential or not.