

First, some important points from the last class:

Theorem: If $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ is a function whose graph has a tangent plane at the point $(x_0, y_0, f(x_0, y_0))$ (in other words, f is differentiable at (x_0, y_0)), then the tangent plane is the graph of the function

$$L(x, y) = \left(\frac{\partial f}{\partial x}(x_0, y_0) \right) (x - x_0) + \left(\frac{\partial f}{\partial y}(x_0, y_0) \right) (y - y_0) + f(x_0, y_0).$$

Definition: Tangent means that

$$L(x_0, y_0) = f(x_0, y_0)$$

and

$$\lim_{(x,y) \rightarrow (x_0,y_0)} \frac{f(x, y) - L(x, y)}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} = 0.$$

Theorem: If the partial derivatives of $f(x, y)$ are defined near (x_0, y_0) and continuous at (x_0, y_0) , then f is differentiable at (x_0, y_0) .

When f is differentiable at (x_0, y_0) , we can approximate $f(x, y)$ near (x_0, y_0) by

$$f(x, y) \approx L(x, y).$$

This is called the *linear approximation* or *tangent approximation* to f near (x_0, y_0) . The function $L(x, y)$ is called the *linearization* of f at (x_0, y_0) .

Definition: The differential of f is defined by

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy.$$

When f is differentiable, we can use the differential for making approximations:

$$\Delta f \approx \frac{\partial f}{\partial x} \Delta x + \frac{\partial f}{\partial y} \Delta y.$$

Preliminary Homework:

A surface S has the equation $z = f(x, y)$. At $(x, y) = (1, 2)$ we have

$$z = 45 \quad \frac{\partial z}{\partial x} = 2 \quad \frac{\partial z}{\partial y} = 1$$

A bug is crawling on the surface S , and a light shining directly down through S (which is transparent) casts the bug's shadow on the xy -plane; the position of the shadow is $\vec{r}(t)$. At time t_0 , the bug's shadow has position $\vec{r}(t_0) = \langle 1, 2 \rangle$ and velocity $\vec{r}'(t_0) = \langle 3, -1 \rangle$.

1. Find an equation for the tangent plane to S at the point $(1, 2, 45)$.

$$z = 45 + (2)(x - 1) + (1)(y - 2).$$

2. Use the tangent approximation

$$\vec{r}(t) \approx \vec{r}(t_0) + (t - t_0)\vec{r}'(t_0)$$

to approximate the shadow's position at time $t_0 + \Delta t$, where Δt is a very small change in t .

$$\vec{r}(t_0 + \Delta t) \approx \langle 1, 2 \rangle + \Delta t \langle 3, -1 \rangle = \langle 1 + 3\Delta t, 2 - \Delta t \rangle$$

3. Use the equation of the tangent plane to S to approximate the bug's new z -coordinate.

$$\begin{aligned} z &\approx 45 + 2(x - 1) + 1(y - 2) \approx 45 + 2((1 + 3\Delta t) - 1) + 1((2 - \Delta t) - 2) = \\ &45 + (2)(3)\Delta t + (1)(-1)\Delta t \end{aligned}$$

4. Show that your answer is

$$45 + \left(\left\langle \frac{\partial z}{\partial x}(\vec{r}(t_0)), \frac{\partial z}{\partial y}(\vec{r}(t_0)) \right\rangle \cdot \vec{r}'(t_0) \right) \Delta t.$$

$$45 + (2)(3)\Delta t + (1)(-1)\Delta t = \underbrace{45}_{f(\vec{r}(t_0))} + \left(\underbrace{\langle 2, 1 \rangle}_{\left\langle \frac{\partial f}{\partial x}(1, 2), \frac{\partial f}{\partial y}(1, 2) \right\rangle} \cdot \underbrace{\langle 3, -1 \rangle}_{\vec{r}'(t_0)} \right) \Delta t.$$

$$f(\vec{r}(t_0 + \Delta t)) \approx f(\vec{r}(t_0)) + \left(\left\langle \frac{\partial z}{\partial x}(\vec{r}(t_0)), \frac{\partial z}{\partial y}(\vec{r}(t_0)) \right\rangle \cdot \vec{r}'(t_0) \right) \Delta t.$$

$$\boxed{\frac{f(\vec{r}(t_0 + \Delta t)) - f(\vec{r}(t_0))}{\Delta t} \approx \left(\left\langle \frac{\partial z}{\partial x}(\vec{r}(t_0)), \frac{\partial z}{\partial y}(\vec{r}(t_0)) \right\rangle \cdot \vec{r}'(t_0) \right)}.$$

Definition: If $f : \mathbb{R}^n \rightarrow \mathbb{R}$, the *gradient* of f is the vector whose components are its partial derivatives:

$$\nabla f(x, y, z) = \left\langle \frac{\partial f}{\partial x}(x, y, z), \frac{\partial f}{\partial y}(x, y, z), \frac{\partial f}{\partial z}(x, y, z) \right\rangle.$$

If f is differentiable, we may also call ∇f the (total) derivative of f and write it f' .

Theorem (the chain rule): If $\vec{r}(t)$ is differentiable at t_0 , and $f(x, y, z)$ is differentiable at $\vec{r}(t_0)$, then

$$\frac{d}{dt}(f(\vec{r}(t))) = f'(\vec{r}(t)) \cdot \vec{r}'(t) = \nabla f(\vec{r}(t)) \cdot \vec{r}'(t).$$

If you want to picture the chain rule geometrically, the preliminary homework gives a way to think about it. Let $\vec{r} : \mathbb{R} \rightarrow \mathbb{R}^2$ be the position function of a point moving in the xy -plane. (In the preliminary homework, this is the shadow of a bug.) Imagine the xy -plane sitting inside \mathbb{R}^3 .

Let $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ be a function whose graph is a surface S , so $f(x, y)$ is the height of the surface at (x, y) . Now imagine a point moving on S directly above (or below) the moving point in the xy -plane. (In the preliminary homework, this is the bug itself.) The height of that point at time t is given by the composition

$$z = f(\vec{r}(t)) = (f \circ \vec{r})(t).$$

To find how fast this height is changing, we compute

$$\frac{dz}{dt} = (f \circ \vec{r})'(t) = f'(\vec{r}(t)) \cdot \vec{r}'(t) = \nabla f(\vec{r}(t)) \cdot \vec{r}'(t).$$

Example: Suppose our bug is crawling around the surface $z = x^2 + y^2$, so that its shadow is moving in the xy -plane with (x, y) -coordinates at time t given by $\vec{r}(t) = \langle t^2, t^3 \rangle$. When the bug is at the point $(1, 1, 2)$, how fast is its height increasing?

The bug's height is $z = f(x, y) = x^2 + y^2$ when its shadow has position (x, y) , and the shadow's position at time t is $(x, y) = \vec{r}(t) = \langle t^2, t^3 \rangle$. When the bug is at $(1, 1, 2)$ we have $\vec{r}(t) = \langle 1, 1 \rangle$ and $t = 1$. By the Chain Rule,

$$\frac{dz}{dt} = \frac{d}{dt}(f(\vec{r}(t))) = \nabla f(\vec{r}(t)) \cdot \vec{r}'(t).$$

$$\vec{r}'(t) = \langle 2t, 3t^2 \rangle \quad \vec{r}'(1) = \langle 2, 3 \rangle \quad \nabla f(x, y) = \langle 2x, 2y \rangle \quad \nabla f(1, 1) = \langle 2, 2 \rangle$$

$$\left. \frac{dz}{dt} \right|_{t=1} = \nabla f(\vec{r}(1)) \cdot \vec{r}'(1) = \nabla f(1, 1) \cdot \langle 2, 3 \rangle = \langle 2, 2 \rangle \cdot \langle 2, 3 \rangle = 10.$$

Theorem: (the chain rule): If $\vec{r}(t)$ is differentiable at t_0 , and $f(x, y, z)$ is differentiable at $\vec{r}(t_0)$, then

$$\frac{d}{dt} (f(\vec{r}(t))) = \nabla f(\vec{r}(t)) \cdot \vec{r}'(t).$$

Rephrasing this, if w is a function of x, y, z , and x, y, z are all functions of t , then

$$\begin{aligned} \frac{dw}{dt} &= \left\langle \frac{\partial w}{\partial x}, \frac{\partial w}{\partial y}, \frac{\partial w}{\partial z} \right\rangle \cdot \left\langle \frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} \right\rangle = \frac{\partial w}{\partial x} \frac{dx}{dt} + \frac{\partial w}{\partial y} \frac{dy}{dt} + \frac{\partial w}{\partial z} \frac{dz}{dt} \\ \Delta w &\approx \frac{\partial w}{\partial x} \Delta x + \frac{\partial w}{\partial y} \Delta y + \frac{\partial w}{\partial z} \Delta z \approx \frac{\partial w}{\partial x} \frac{dx}{dt} \Delta t + \frac{\partial w}{\partial y} \frac{dy}{dt} \Delta t + \frac{\partial w}{\partial z} \frac{dz}{dt} \Delta t = \\ &\quad \left(\frac{\partial w}{\partial x} \frac{dx}{dt} + \frac{\partial w}{\partial y} \frac{dy}{dt} + \frac{\partial w}{\partial z} \frac{dz}{dt} \right) \Delta t = \left(\frac{dw}{dt} \right) \Delta t \end{aligned}$$

Here's another way to envision the chain rule physically: Suppose (x, y, z) is the position of a moving object, and $w = f(x, y, z)$ is the temperature at point (x, y, z) . To find the rate of change of the temperature of the moving object with respect to time, we have

$$\frac{dw}{dt} = \frac{\partial w}{\partial x} \frac{dx}{dt} + \frac{\partial w}{\partial y} \frac{dy}{dt} + \frac{\partial w}{\partial z} \frac{dz}{dt}.$$

We can approximate the change in temperature over a small interval of time Δt by

$$\Delta w \approx \frac{\partial w}{\partial x} \underbrace{\frac{dx}{dt} \Delta t}_{\approx \Delta x} + \frac{\partial w}{\partial y} \underbrace{\frac{dy}{dt} \Delta t}_{\approx \Delta y} + \frac{\partial w}{\partial z} \underbrace{\frac{dz}{dt} \Delta t}_{\approx \Delta z}$$

Example: If $w = x^2 y^2$, $x = \sin(t)$, and $y = \cos(t)$, find $\frac{dw}{dt}$ at $t = \frac{\pi}{3}$.

$$\begin{aligned} t = \frac{\pi}{3} \quad x &= \sin\left(\frac{\pi}{3}\right) = \frac{\sqrt{3}}{2} \quad y = \cos\left(\frac{\pi}{3}\right) = \frac{1}{2} \\ \frac{\partial w}{\partial x} &= 2xy^2 = \frac{\sqrt{3}}{4} \quad \frac{\partial w}{\partial y} = 2x^2y = \frac{3}{4} \quad \frac{dx}{dt} = \cos(t) = \frac{1}{2} \quad \frac{dy}{dt} = -\sin(t) = -\frac{\sqrt{3}}{2} \\ \frac{dw}{dt} &= \frac{\partial w}{\partial x} \frac{dx}{dt} + \frac{\partial w}{\partial y} \frac{dy}{dt} = \left(\frac{\sqrt{3}}{4}\right) \left(\frac{1}{2}\right) + \left(\frac{3}{4}\right) \left(-\frac{\sqrt{3}}{2}\right) = -\frac{\sqrt{3}}{4} \end{aligned}$$

The chain rule in different settings:

$$w = f(x) = f(g(t))$$

$$t \rightarrow x \rightarrow w$$

$$\frac{dw}{dt} = \frac{dw}{dx} \frac{dx}{dt}$$

$$w = f(x, y, z) = f(\vec{g}(t))$$

$$t \rightarrow (x, y, z) \rightarrow w$$

$$\frac{dw}{dt} = \underbrace{\frac{\partial w}{\partial x} \frac{dx}{dt} + \frac{\partial w}{\partial y} \frac{dy}{dt} + \frac{\partial w}{\partial z} \frac{dz}{dt}}_{\text{one term for each intermediate variable}}$$

$$w = f(x, y, z) = f(G(s, t))$$

$$(s, t) \rightarrow (x, y, z) \rightarrow w$$

$$\frac{\partial w}{\partial t} = \frac{\partial w}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial w}{\partial y} \frac{\partial y}{\partial t} + \frac{\partial w}{\partial z} \frac{\partial z}{\partial t}$$

$$\frac{\partial w}{\partial s} = \frac{\partial w}{\partial x} \frac{\partial x}{\partial s} + \frac{\partial w}{\partial y} \frac{\partial y}{\partial s} + \frac{\partial w}{\partial z} \frac{\partial z}{\partial s}$$

Example: We can identify points on the cone $x^2 + y^2 = z^2$, $z \geq 0$, using two coordinates, r and θ , by setting

$$x = r \cos(\theta) \quad y = r \sin(\theta) \quad z = r \quad 0 \leq \theta \leq 2\pi \quad 0 \leq r.$$

Define w on the cone by

$$w = xy - xz^2.$$

Find $\frac{\partial w}{\partial r}$ at the point $(x, y, z) = (-2, 0, 2)$.

At the point $(x, y) = (-2, 0)$ we have

$$\begin{aligned} r = 2 \quad \theta = \pi \quad x = -2 \quad y = 0 \quad z = 2 \\ \frac{\partial w}{\partial x} = y - z^2 = -4 \quad \frac{\partial w}{\partial y} = x = -2 \quad \frac{\partial w}{\partial z} = -2xz = 8 \\ \frac{\partial x}{\partial r} = \cos(\theta) = -1 \quad \frac{\partial y}{\partial r} = \sin(\theta) = 0 \quad \frac{\partial z}{\partial r} = 1 \end{aligned}$$

We treat θ as a constant and differentiate with respect to r , using the chain rule:

$$\frac{\partial w}{\partial r} = \frac{\partial w}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial w}{\partial y} \frac{\partial y}{\partial r} + \frac{\partial w}{\partial z} \frac{\partial z}{\partial r} = (-4)(-1) + (-2)(0) + (8)(1) = 12$$

At a general point, we have

$$\begin{aligned} \frac{\partial w}{\partial r} &= \frac{\partial w}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial w}{\partial y} \frac{\partial y}{\partial r} + \frac{\partial w}{\partial z} \frac{\partial z}{\partial r} = (y - z^2)(\cos(\theta)) + (x)(\sin(\theta)) + (-2xz)(1) = \\ &= (r \sin(\theta) - r^2)(\cos(\theta)) + (r \cos(\theta))(\sin(\theta)) + (-2r^2 \cos(\theta))(1) = \\ &= 2r \sin(\theta) \cos(\theta) - 3r^2 \cos(\theta). \end{aligned}$$

What does this mean? We define w as a function of (r, θ) by looking at the point on the cone $(x, y, z) = (r \cos(\theta), r \sin(\theta), r)$, then computing $w = xy - xz^2$. We want to know, when $(x, y) = (-2, 0)$, the rate of change of w with respect to r .

For example, suppose w denotes the temperature at a given point on the cone. Consider the ubiquitous bug crawling on the cone, with its shadow moving in the xy -plane. The bug's temperature is w . When the bug's shadow is where $(r, \theta) = (2, -\pi)$, and the bug moves so its shadow's new location is where $(r, \theta) = (2 + \Delta r, -\pi)$ (that is, θ remains constant and r changes by Δr), the bug's temperature will have changed by

$$\Delta w \approx \frac{\partial w}{\partial r} \Delta r = 12 \Delta r.$$

Recall implicitly defined functions: An equation $f(x, y, z) = 0$ defining a surface S can be thought of as implicitly defining z as a function of x and y near a point on S . If we want to find $\frac{\partial z}{\partial x}$ at that point, we can treat y as a constant and z as a function of x , and differentiate the equation with respect to x :

$$\begin{aligned}\frac{\partial}{\partial x}(f(x, y, z)) &= \frac{\partial}{\partial x}(0) \\ \frac{\partial f}{\partial x} \underbrace{\frac{\partial x}{\partial x}}_{=1} + \frac{\partial f}{\partial y} \underbrace{\frac{\partial y}{\partial x}}_{=0} + \frac{\partial f}{\partial z} \underbrace{\frac{\partial z}{\partial x}}_{\text{unknown}} &= 0 \\ \frac{\partial f}{\partial x} + \frac{\partial f}{\partial z} \frac{\partial z}{\partial x} &= 0 \\ \frac{\partial z}{\partial x} &= -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial z}}\end{aligned}$$

Theorem: (the implicit function theorem) If $f(x, y, z) = 0$ implicitly defines z as a function of x and y , then

$$\frac{\partial z}{\partial x} = -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial z}} \quad \frac{\partial z}{\partial y} = -\frac{\frac{\partial f}{\partial y}}{\frac{\partial f}{\partial z}}.$$

If $f(x, y) = 0$ implicitly defines y as a function of x , then

$$\frac{dy}{dx} = -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}}.$$

Example: Earlier, we looked at the surface

$$ax^2 + by^2 + cz^2 = d$$

and used implicit differentiation:

$$\begin{aligned}2ax + 2cz \frac{\partial z}{\partial x} &= 0 \\ \frac{\partial z}{\partial x} &= -\frac{ax}{cz}.\end{aligned}$$

Now we can use the implicit function theorem:

$$\begin{aligned}f(x, y, z) &= ax^2 + by^2 + cz^2 - d & f(x, y, z) &= 0 \\ \frac{\partial z}{\partial x} &= -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial z}} = -\frac{2ax}{2cz} = -\frac{ax}{cz}.\end{aligned}$$

You do not have to memorize the implicit function theorem, but you may use it if you wish.

Exercise: A surface S has the equation $z = f(x, y)$. At $(x, y) = (1, 2)$ we have

$$z = 45 \quad \frac{\partial z}{\partial x} = 2 \quad \frac{\partial z}{\partial y} = 1$$

A bug is crawling on the surface S , and a light shining directly down through S (which is transparent) casts the bug's shadow on the xy -plane; the position of the shadow is $\vec{r}(t)$. At time t_0 , the bug's shadow has position $\vec{r}(t_0) = \langle 1, 2 \rangle$ and velocity $\vec{r}'(t_0) = \langle 3, -1 \rangle$.

Use the chain rule to find the rate of change of the bug's altitude with respect to time at the time t_0 .

Exercise: A radiation source at the origin subjects an object located at point (x, y, z) to radiation of intensity $I = \frac{k}{x^2 + y^2 + z^2}$, where k is a constant. At time $t = 0$, an object located at point $(1, 2, 1)$ is moving toward the point $(4, 6, 13)$ at a speed of 2.

Find the object's velocity at time $t = 0$.

At time $t = 0$, at what speed is the radiation intensity experienced by the object changing?

Exercise: We can identify points on the cone $x^2 + y^2 = z^2$, $z \geq 0$, using two coordinates, r and θ , by setting

$$x = r \cos(\theta) \quad y = r \sin \theta \quad z = r \quad 0 \leq \theta \leq 2\pi \quad 0 \leq r.$$

Define w by

$$w = xy - xz^2.$$

Find $\frac{\partial w}{\partial \theta}$ at the point $(-2, 0, 2)$.

(This is the same cone, the same function w , and the same point as in the example on page 6.)

Exercise: Suppose that S is a level surface $f(x, y, z) = k$ of a differentiable function f and $\vec{r}(t)$ is a regular parametrization of a path γ lying in S . Since the value of f equals k for all points on S , and all points $\vec{r}(t)$ are on S , we have

$$f(\vec{r}(t)) = k.$$

Start with this equation and differentiate both sides (using the chain rule for the left hand side) to show that

$$\nabla f(\vec{r}(t)) \perp \vec{r}'(t).$$

Remember that k is a constant!

Since this is true for any path γ in S , we can conclude that

$$\nabla f \perp S$$

You just proved the following **Theorem:** The gradient of a differentiable function f at a point is normal to the level surface (or level curve) of f containing that point.

Exercise: If $f(x, y, z) = 4x^2 - y^2 + z^2$, then the hyperboloid $4x^2 - y^2 + z^2 = 4$ is a level surface of f , so it should be perpendicular to the gradient of f at every point. The point $(1, 1, 1)$ is on this surface. Verify that the surface is perpendicular to the gradient of f at the point $(1, 1, 1)$ in the following way:

Use the implicit function theorem to compute $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ for the portion of the surface containing $(1, 1, 1)$, then evaluate those partial derivatives at the point $(1, 1, 1)$. Use these values to find the equation of a tangent plane to the surface at $(1, 1, 1)$.

Verify that the tangent plane is perpendicular to $\nabla f(1, 1, 1)$.

Theorem: (the chain rule): If $\vec{r}(t)$ is differentiable at t_0 , and $f(x, y)$ is differentiable at $\vec{r}(t_0)$, then

$$\frac{d}{dt} (f(\vec{r}(t))) = \nabla f(\vec{r}(t)) \cdot \vec{r}'(t).$$

Proof of the Chain Rule:

Let $\vec{r}(t) = (x_0, y_0)$. If f is differentiable at (x_0, y_0) , we can express $f(x, y)$ as

$$f(x, y) = \underbrace{a(x - x_0) + b(y - y_0) + f(x_0, y_0)}_{\text{tangent approximation } \mathcal{P}(x,y)} + \underbrace{E(x, y)}_{\text{error } f(x,y) - \mathcal{P}(x,y)}$$

where

$$\lim_{(x,y) \rightarrow (x_0,y_0)} \frac{E(x, y)}{|(x - x_0, y - y_0)|} = 0$$

$$a = \frac{\partial f}{\partial x}(x_0, y_0) \quad b = \frac{\partial f}{\partial y}(x_0, y_0) \quad \nabla f(x_0, y_0) = \langle a, b \rangle$$

We have $\vec{r}(t_0) = (x_0, y_0)$ and \vec{r} is differentiable at t_0 . We can write $\vec{r}(t) = \langle x(t), y(t) \rangle$ and compute the derivative of $f(\vec{r}(t))$ at $t = t_0$. We use the limit definition of derivative:

$$\frac{d}{dt} (f(\vec{r}(t_0))) = \lim_{t \rightarrow t_0} \frac{f(\vec{r}(t)) - f(\vec{r}(t_0))}{t - t_0}$$

Now we say $\vec{r}(t) = \langle x(t), y(t) \rangle$.

$$\lim_{t \rightarrow t_0} \frac{f(\vec{r}(t)) - f(\vec{r}(t_0))}{t - t_0} = \lim_{t \rightarrow t_0} \frac{f(x(t), y(t)) - f(x_0, y_0)}{t - t_0}$$

Now we express $f(x, y)$ as above.

$$\lim_{t \rightarrow t_0} \frac{f(x(t), y(t)) - f(x_0, y_0)}{t - t_0} =$$

$$\lim_{t \rightarrow t_0} \frac{a(x(t) - x_0) + b(y(t) - y_0) + f(x_0, y_0) + E(x(t), y(t)) - f(x_0, y_0)}{t - t_0}$$

We cancel some things out and do some regrouping:

$$\lim_{t \rightarrow t_0} \frac{a(x(t) - x_0) + b(y(t) - y_0) + f(x_0, y_0) + E(x(t), y(t)) - f(x_0, y_0)}{t - t_0} =$$

$$\lim_{t \rightarrow t_0} \frac{a(x(t) - x(t_0)) + b(y(t) - y(t_0)) + E(x(t), y(t))}{t - t_0} =$$

$$a \lim_{t \rightarrow t_0} \frac{(x(t) - x(t_0))}{t - t_0} + b \lim_{t \rightarrow t_0} \frac{(y(t) - y(t_0))}{t - t_0} + \lim_{t \rightarrow t_0} \frac{E(x(t), y(t))}{t - t_0}$$

In the first two limits we recognize the definition of the derivative.

$$\begin{aligned} a \lim_{t \rightarrow t_0} \frac{(x(t) - x(t_0))}{t - t_0} + b \lim_{t \rightarrow t_0} \frac{(y(t) - y(t_0))}{t - t_0} + \lim_{t \rightarrow t_0} \frac{E(x(t), y(t))}{t - t_0} &= \\ ax'(t_0) + by'(t_0) + \lim_{t \rightarrow t_0} \frac{E(x(t), y(t))}{t - t_0} &= \\ \langle a, b \rangle \cdot \langle x'(t_0), y'(t_0) \rangle + \lim_{t \rightarrow t_0} \frac{E(x(t), y(t))}{t - t_0} & \end{aligned}$$

Now we remember that $\langle a, b \rangle = \nabla f(x_0, y_0) = \nabla f(\vec{r}(t_0))$.

$$\langle a, b \rangle \cdot \langle x'(t_0), y'(t_0) \rangle + \lim_{t \rightarrow t_0} \frac{E(x(t), y(t))}{t - t_0} = \boxed{\nabla f(\vec{r}(t_0)) \cdot \vec{r}'(t_0)} + \lim_{t \rightarrow t_0} \frac{E(x(t), y(t))}{t - t_0}$$

The boxed part is what we want, so we have to show the remaining limit is zero.

We will assume for simplicity that $\vec{r}'(t_0) \neq \vec{0}$, so that for t near t_0 we have $\vec{r}(t) \neq \vec{r}(t_0)$, and we can safely divide by $|\vec{r}(t) - \vec{r}(t_0)|$. (This assumption can be eliminated by a small trick.)

$$\begin{aligned} \lim_{t \rightarrow t_0} \left| \frac{E(x(t), y(t))}{t - t_0} \right| &= \lim_{t \rightarrow t_0} \left| \frac{E(x(t), y(t))}{|\vec{r}(t) - \vec{r}(t_0)|} \right| \left| \frac{|\vec{r}(t) - \vec{r}(t_0)|}{t - t_0} \right| = \\ \lim_{t \rightarrow t_0} \left| \frac{E(x(t), y(t))}{|\vec{r}(t) - \vec{r}(t_0)|} \right| \lim_{t \rightarrow t_0} \left| \frac{|\vec{r}(t) - \vec{r}(t_0)|}{t - t_0} \right| & \end{aligned}$$

Now we recognize the definition of derivative in the right-hand term.

$$\lim_{t \rightarrow t_0} \left| \frac{E(x(t), y(t))}{|\vec{r}(t) - \vec{r}(t_0)|} \right| \lim_{t \rightarrow t_0} \left| \frac{|\vec{r}(t) - \vec{r}(t_0)|}{t - t_0} \right| = \lim_{t \rightarrow t_0} \left| \frac{E(x(t), y(t))}{|\vec{r}(t) - \vec{r}(t_0)|} \right| |\vec{r}'(t_0)|$$

We use the fact that as $t \rightarrow t_0$ we have $\vec{r}(t) \rightarrow \vec{r}(t_0)$, or $(x, y) \rightarrow (x_0, y_0)$.

$$\lim_{t \rightarrow t_0} \left| \frac{E(x(t), y(t))}{|\vec{r}(t) - \vec{r}(t_0)|} \right| |\vec{r}'(t_0)| = \lim_{(x, y) \rightarrow (x_0, y_0)} \left| \frac{E(x, y)}{|(x, y) - (x_0, y_0)|} \right| |\vec{r}'(t_0)|$$

Now we use the definition of tangent.

$$\left(\lim_{(x, y) \rightarrow (x_0, y_0)} \left| \frac{E(x, y)}{|(x, y) - (x_0, y_0)|} \right| \right) |\vec{r}'(t_0)| = (0) (|\vec{r}'(t_0)|) = 0.$$

This is what we wanted. We have shown

$$\boxed{\frac{d}{dt} (f(\vec{r}(t_0))) = \nabla f(\vec{r}(t_0)) \cdot \vec{r}'(t_0)}$$