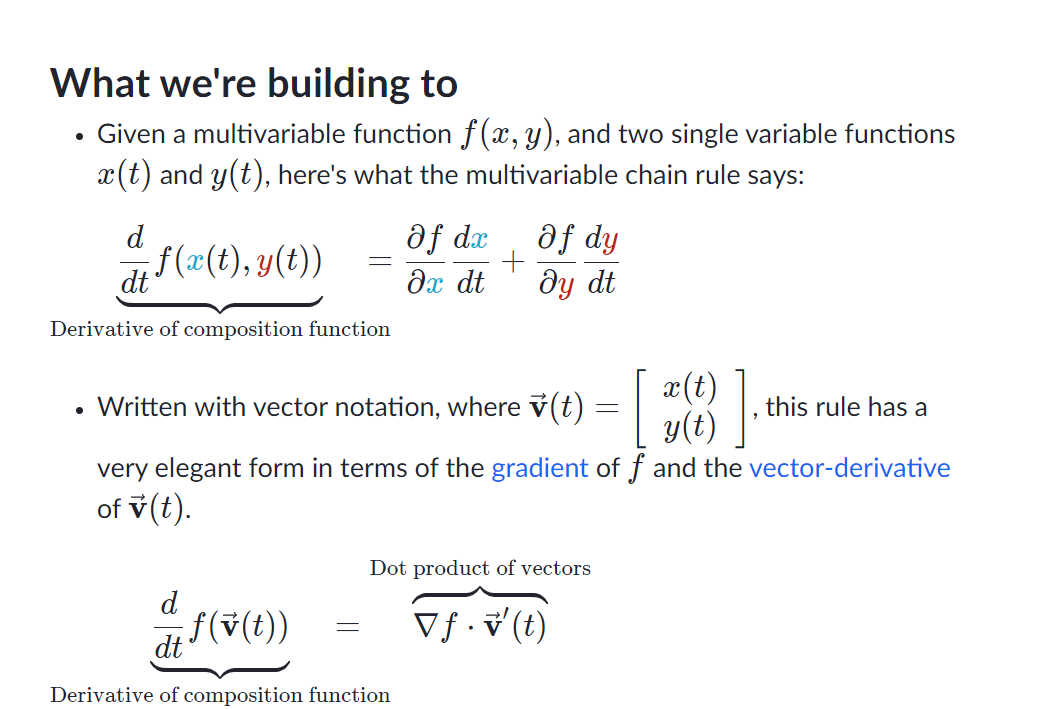
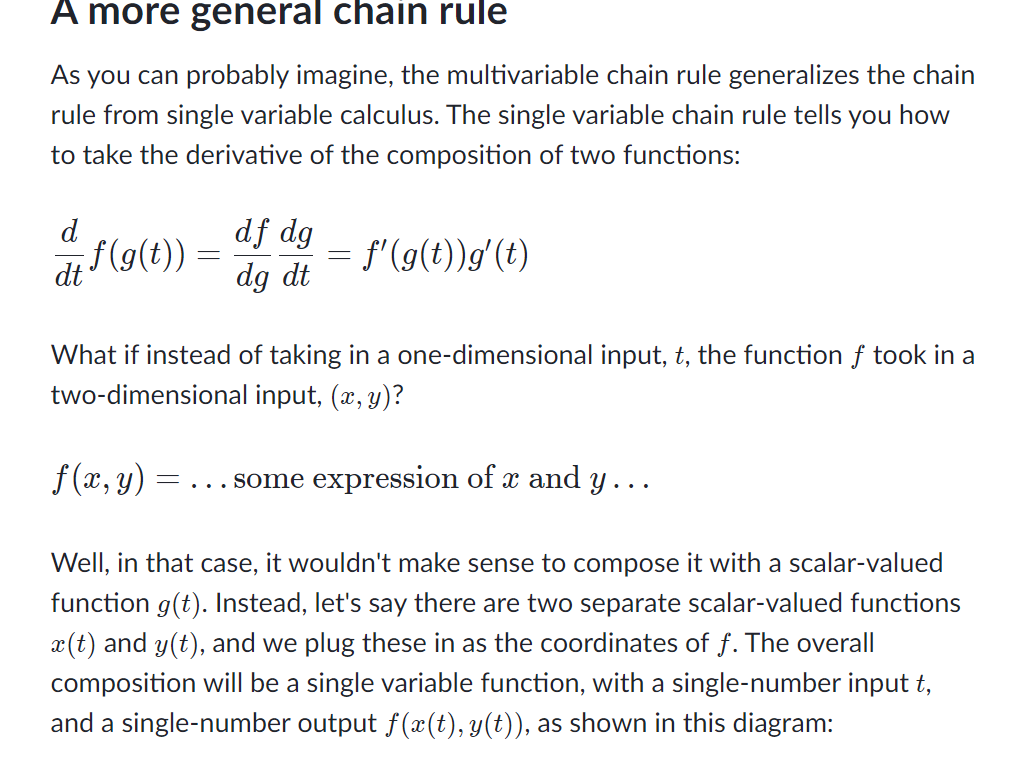
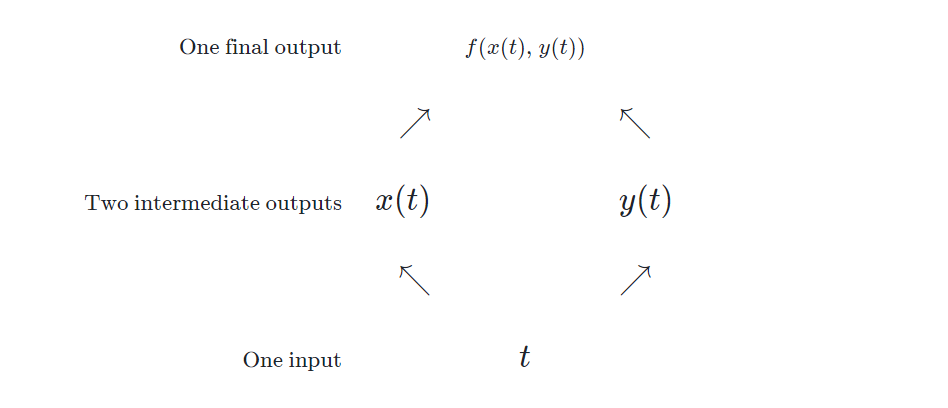
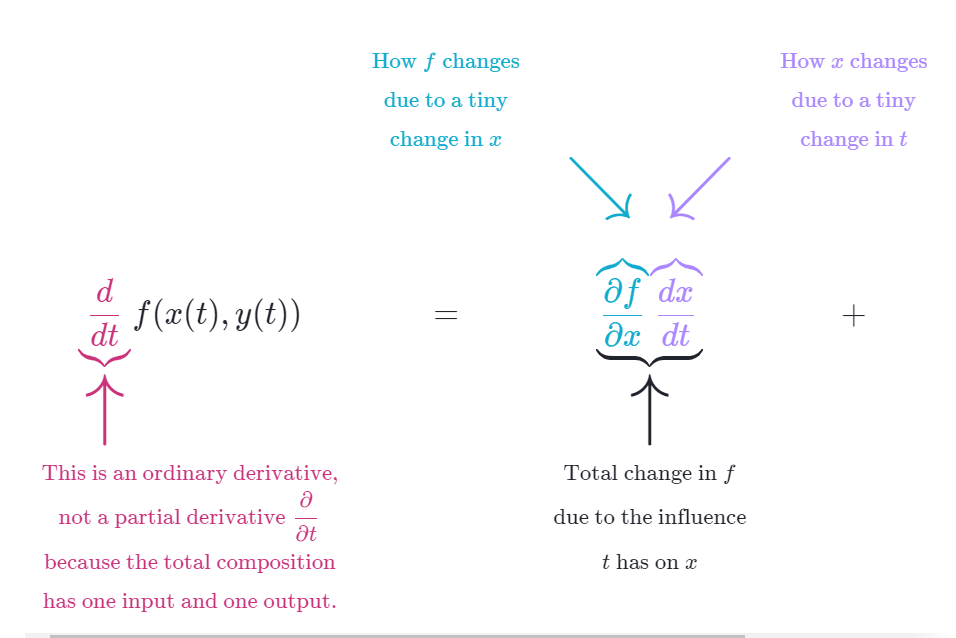
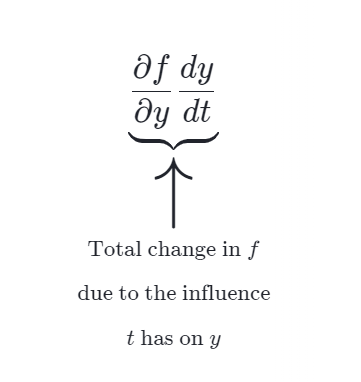
Multivariable Chain Rule, directional derivative, and gradient

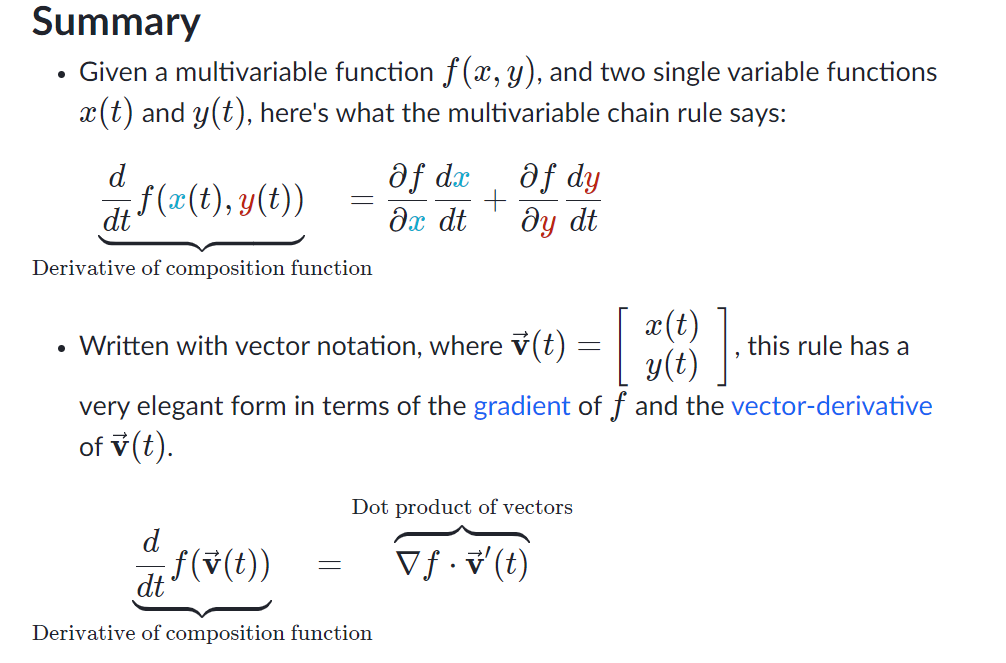
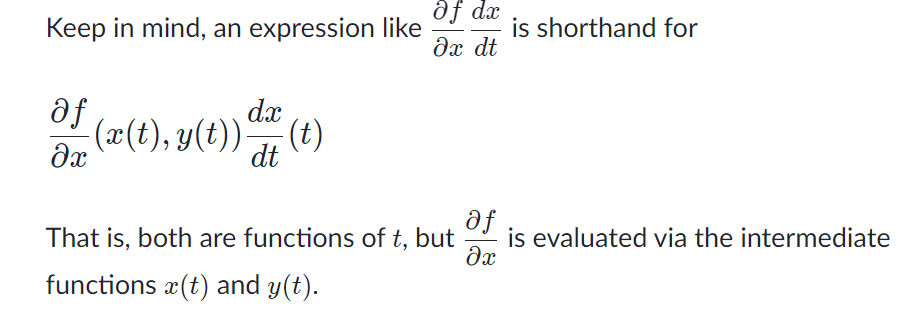












Application of multivariable chain rule:

1. find the derivative of f(x) =(x-1)2 \* (x+5)2

Let u= (x-1), v= (x+5), respectively;

f(x) = f(u,v)= u2\*v2

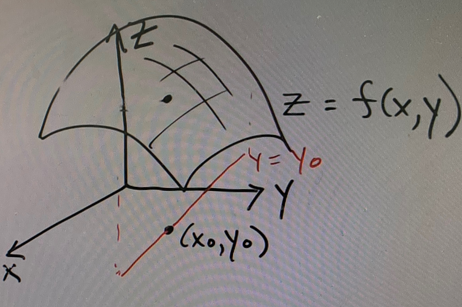
f’(x) = 2uv2 + 2u2 v=2(x-1)\*(x+5)2 + 2(x-1)2 \* (x+5)

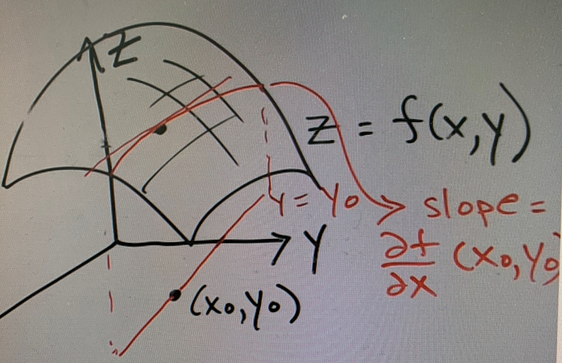
formula (f/g)’ = (f’g – fg’)/g2

Directional derivative and gradient:

[Gradient 2](https://www.youtube.com/watch?v=NomUbVmmyro)

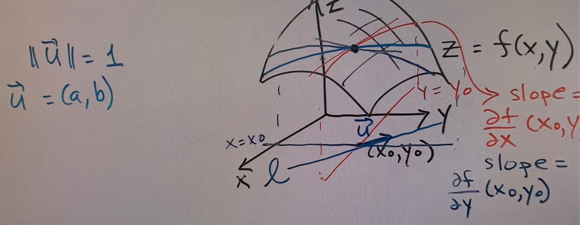
Z is a function of x, y. the derivative of z at point (x0, y0) along line y=y0 is the slope of the tangent line on z surface. And it is the rate of change of function z at (x0, y0) along line y=y0.





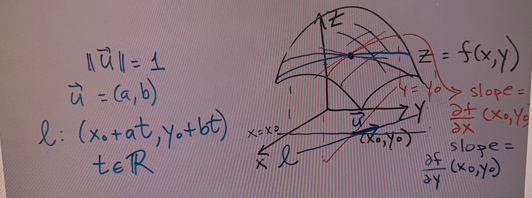
Similarly, the slope of the tangent line on z surface along line x=x0 at point (x0, y0) is the rate of change at (x0, y0) along line x=x0.

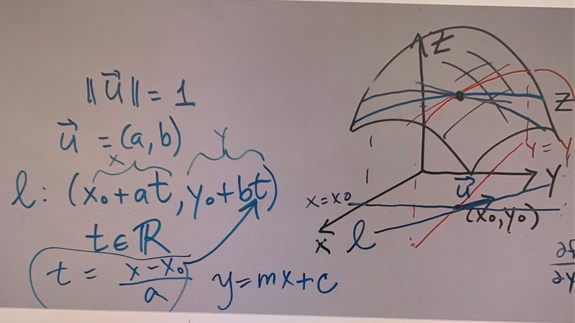
How about the slope (or rate of change) of z at (x0, y0) alone a general direction of u= (a, b)



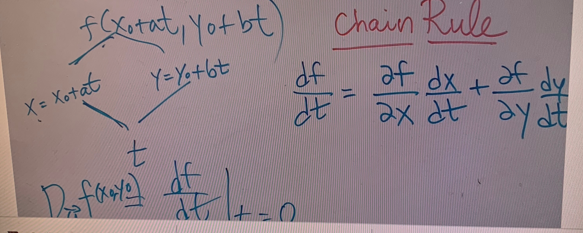
It is called the directional derivative of z alone the direction of **u**=(a, b) at point (x0, y0)

To find it, use the parametric line equation and the chain rule.

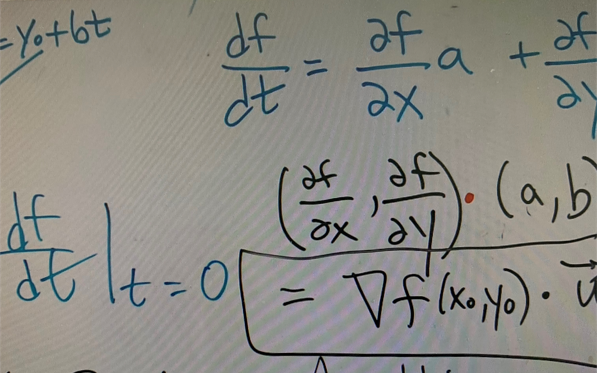




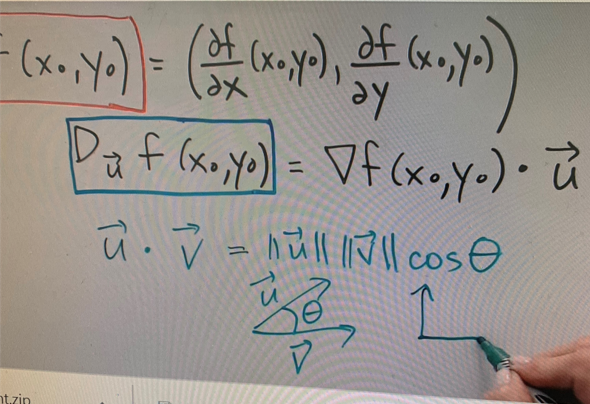
Chain rule:



Re-write:



Geometric meaning: if **u** coincide with the gradient vector ∇f(x0, y0), we shall have the greatest slope (rate of change), and the greatest rate of change is the length of the gradient vector. Therefore, at any point (x0, y0) on the X , Y plane, the direction to have the greatest rate of change (i.e., amount of increment on z due to a small unit of change on X, Y plane) is the direction of the gradient vector.



Summary

Therefore, if we want to maximize the value of function z=f(x,y) at (x0, y0), we move along the gradient vector. In contrast, if we want to minimize the z function, we move along the negative of the gradient vector.

The idea of gradient descent: move along the negative of gradient vector to minimize the cost (or Loss) function.

More examples: A **gradient is always perpendicular to the level curve.**

[Gradient 1](http://sites.science.oregonstate.edu/math/home/programs/undergrad/CalculusQuestStudyGuides/vcalc/grad/grad.html)

