(*) **The chain rule in several variables**

Given a function of two variable $z = f(x, y)$, where both of the variables x, y are themselves functions of another variable,

$$x = x(t) \quad \text{and} \quad y = y(t),$$

it follows that z is (indirectly) a function of t as well. I.e.,

$$z = f(x, y) = f(x(t), y(t)) = \tilde{f}(t),$$

so it makes sense to compute dz/dt, which is done using the several-variables version of the chain rule:

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

More generally (but only slightly), suppose that $w = g(x, y, z)$, where x, y, z are themselves functions of the variables α and β:

$$x = x(\alpha, \beta),\ y = y(\alpha, \beta) \ \text{ and } \ z = z(\alpha, \beta).$$

Then w is (indirectly) a function of α and β, and the partial derivatives $\partial w/\partial\alpha$ and $\partial w/\partial\beta$ are given by

$$\frac{\partial w}{\partial \alpha} = \frac{\partial w}{\partial x} \cdot \frac{\partial x}{\partial \alpha} + \frac{\partial w}{\partial y} \cdot \frac{\partial y}{\partial \alpha} + \frac{\partial w}{\partial z} \cdot \frac{\partial z}{\partial \alpha} \quad \text{and} \quad \frac{\partial w}{\partial \beta} = \frac{\partial w}{\partial x} \cdot \frac{\partial x}{\partial \beta} + \frac{\partial w}{\partial y} \cdot \frac{\partial y}{\partial \beta} + \frac{\partial w}{\partial z} \cdot \frac{\partial z}{\partial \beta}.$$

(*) **Explanation**

I'll explain the first half of the second case above using linear approximation.[†] The same explanation works in general.

Suppose that the variable α changes by a small amount $\Delta\alpha$ and that the variable β is held fixed. In this case, the three variables x, y and z will change by the amounts Δx, Δy and Δz, and if $\Delta\alpha$ is small then these changes can be approximated using linear approximation:

$$\Delta x \approx \frac{\partial x}{\partial \alpha} \cdot \Delta\alpha, \quad \Delta y \approx \frac{\partial y}{\partial \alpha} \cdot \Delta\alpha \quad \text{and} \quad \Delta z \approx \frac{\partial z}{\partial \alpha} \cdot \Delta\alpha. \tag{1}$$

Now when x, y and z all change, then w will change by some amount Δw. If the change $\Delta\alpha$ is small enough, then the resulting changes Δx, Δy and Δz will also be small, which means that Δw can also be approximated using (the more general form of) linear approximation:

$$\Delta w \approx \frac{\partial w}{\partial x} \cdot \Delta x + \frac{\partial w}{\partial y} \cdot \Delta y + \frac{\partial w}{\partial z} \cdot \Delta z. \tag{2}$$

Substituting the approximations from Equation (1) into the approximation Equation (2), we see that if $\Delta\alpha$ is small enough, then

$$\Delta w \approx \frac{\partial w}{\partial x} \cdot \frac{\partial x}{\partial \alpha} \cdot \Delta\alpha + \frac{\partial w}{\partial y} \cdot \frac{\partial y}{\partial \alpha} \cdot \Delta\alpha + \frac{\partial w}{\partial z} \cdot \frac{\partial z}{\partial \alpha} \cdot \Delta\alpha, \tag{3}$$

[†]See the Lecture Notes from 11/13/17 on the course website.

and dividing both sides of (3) by $\Delta\alpha$ shows that

$$\frac{\Delta w}{\Delta \alpha} \approx \frac{\partial w}{\partial x} \cdot \frac{\partial x}{\partial \alpha} + \frac{\partial w}{\partial y} \cdot \frac{\partial y}{\partial \alpha} + \frac{\partial w}{\partial z} \cdot \frac{\partial z}{\partial \alpha}$$

and furthermore, this approximation becomes increasingly accurate as $\Delta\alpha$ shrinks to 0. In other words,

$$\frac{\partial w}{\partial \alpha} = \lim_{\Delta\alpha \to 0} \frac{\Delta w}{\Delta \alpha} = \frac{\partial w}{\partial x} \cdot \frac{\partial x}{\partial \alpha} + \frac{\partial w}{\partial y} \cdot \frac{\partial y}{\partial \alpha} + \frac{\partial w}{\partial z} \cdot \frac{\partial z}{\partial \alpha},$$

as the chain rule states.

(*) **The Envelope Theorem**

Read Supplementary Note #4.[‡]

[‡]Again, right? Because you already read it at least once, as suggested in the syllabus.