# Higher–Dimensional Chain Rules

**I. Introduction.** The one–dimensional Chain Rule tells us how to find the derivative of a composed function $w(t) = k\big(c(t)\big)$ in terms of $k'$, $c$, and $c'$, in the case where $k$ and $c$ (and therefore $h$) are real-valued functions of one variable. As you are well aware, the Chain Rule says: If $k'\big(c(t)\big)$ and $c'(t)$ both exist then $w'(t)$ exists and is given by

$$\frac{d}{dt}\Big[k\big(c(t)\big)\Big] = w'(t) = k'\big(c(t)\big)c'(t).$$

Now, the notion of composing functions extends to functions with more general domains and ranges. For example, if

$$\vec{C}(t) = \langle\, \cos(t), \sin(t)\, \rangle$$

and

$$k(x, y) = x^2 e^y,$$

then

$$w(t) = k\big(\vec{C}(t)\big) = \cos^2(t)e^{\sin(t)}, \tag{$*$}$$

and you might guess that there exists another chain rule, similar to the first one, for this situation. That guess would be correct: there is indeed a chain rule which gives $\frac{d}{dt}\Big[k\big(\vec{C}(t)\big)\Big]$ in terms of the functions $k$ and $\vec{C}$ and their derivatives. In fact, there is a whole family of chain rules, one for each case

$$\mathbf{R}^n \overset{C}{\longrightarrow} \mathbf{R}^k \overset{k}{\longrightarrow} \mathbf{R}^\ell.$$

This handout will state and (almost) prove the first and simplest of these, the case in which $n = \ell = 1$ and $k = 2$:

$$\mathbf{R} \overset{\vec{C}}{\longrightarrow} \mathbf{R}^2 \overset{k}{\longrightarrow} \mathbf{R}.$$

Note that ($*$) above is an instance of this case.

**II. The (almost) proof.** First, let me set the stage and introduce some notation.

- **[a]:** Let $\vec{C} = \vec{C}(t) = \langle\, f(t), g(t)\, \rangle$ be a planar curve, with $(x_0, y_0) = \vec{C}(t_0) = \langle\, f(t_0), g(t_0)\, \rangle$. I will assume that $\vec{C}'(t_0) = \langle\, f'(t_0), g'(t_0)\, \rangle$ exists.
- **[b]:** Let $k = k(x, y)$ be a real–valued function of two variables, defined in a neighborhood of $(x_0, y_0) = \vec{C}(t_0)$. I will assume that $k$ is differentiable at $(x_0, y_0)$, so that

$$\lim_{(x,y)\to(x_0,y_0)} \frac{|f(x, y) - L(x, y)|}{\|\,\langle\, x - x_0, y - y_0\, \rangle\,\|} = 0, \tag{1}$$

  where $L(x, y) = k(x_0, y_0) + k_x(x_0, y_0)(x - x_0) + k_y(x_0, y_0)(y - y_0)$.
- **[c]:** For any $h \neq 0$, let $\left\{ \begin{array}{l} x_h = f(t_0 + h) \\ y_h = g(t_0 + h) \end{array} \right\}$.

Now I can state and prove the theorem.

**Theorem.** *Let $k$ and $\vec{C}$ be as above, and let $w(t) = k\big(\vec{C}(t)\big)$. Then $w$ is differentiable at $t_0$, and*

$$w'(t_0) = k_x(x_0, y_0)f'(t_0) + k_y(x_0, y_0)g'(t_0). \tag{2}$$

*Proof*. We begin with the definition of the (one–dimensional) derivative:

$$w'(t_0) = \lim_{h\to 0} \frac{k\big(\vec{C}(t_0 + h)\big) - k\big(\vec{C}(t_0)\big)}{h} = \lim_{h\to 0} \frac{k\big(f(t_0 + h), g(t_0 + h)\big) - k\big(f(t_0), g(t_0)\big)}{h}$$

$$= \lim_{h\to 0} \frac{k\big(x_h, y_h\big) - k\big(x_0, y_0\big)}{h}. \tag{3}$$

I will break the problem of evaluating this limit into several steps.

**Step 1.** Add and subtract $L(x_h, y_h)$ in the numerator of (3), so as to break (3) into two limit problems:

$$(3) = \lim_{h\to 0} \underbrace{\frac{k(x_h, y_h) - L(x_h, y_h)}{h}}_{(B)} + \lim_{h\to 0} \underbrace{\frac{L(x_h, y_h) - k(x_0, y_0)}{h}}_{(A)},$$

if both of these limits exist.

**Step 2: Evaluating the limit of (A).** Since

$$L(x_h, y_h) = k(x_0, y_0) + k_x(x_0, y_0)(x_h - x_0) + k_y(x_0, y_0)(y_h - y_0),$$

$$(A) = \frac{k_x(x_0, y_0)(x_h - x_0) + k_y(x_0, y_0)(y_h - y_0)}{h},$$

so that

$$\begin{aligned}
\lim_{h\to 0}(A) &= \lim_{h\to 0} k_x(x_0, y_0)\frac{x_h - x_0}{h} + \lim_{h\to 0} k_y(x_0, y_0)\frac{y_h - y_0}{h} \\
&= k_x(x_0, y_0)\lim_{h\to 0}\frac{x_h - x_0}{h} + k_y(x_0, y_0)\lim_{h\to 0}\frac{y_h - y_0}{h} \\
&= k_x(x_0, y_0)\lim_{h\to 0}\frac{f(t_0 + h) - f(t_0)}{h} + k_y(x_0, y_0)\lim_{h\to 0}\frac{g(t_0 + h) - g(t_0)}{h} \\
&= k_x(x_0, y_0)f'(t_0) + k_y(x_0, y_0)g'(t_0).
\end{aligned}$$

Thus, the limit of (A) is exactly $w'(t_0)$ (equation (2)). To prove the theorem, then, I must show that the limit of (B) is *zero*.

**Step 3: Evaluating the limit of (B).** Assume, for all $h \neq 0$, that $\langle\, x_h, y_h \,\rangle \neq \langle\, x_0, y_0 \,\rangle$[1], so that one can multiply and divide (B) by $\|\,\langle\, x_h, y_h \,\rangle - \langle\, x_0, y_0 \,\rangle\,\|$. Consider the absolute value of (B):

$$\big|(B)\big| = \underbrace{\frac{|k(x_h, y_h) - L(x_h, y_h)|}{\|\,\langle\, x_h, y_h \,\rangle - \langle\, x_0, y_0 \,\rangle\,\|}}_{(C)} \cdot \underbrace{\frac{\|\,\langle\, x_h, y_h \,\rangle - \langle\, x_0, y_0 \,\rangle\,\|}{|h|}}_{(D)}.$$

It follows that

$$\lim_{h\to 0}\big|(B)\big| = \lim_{h\to 0}\underbrace{\frac{|k(x_h, y_h) - L(x_h, y_h)|}{\|\,\langle\, x_h, y_h \,\rangle - \langle\, x_0, y_0 \,\rangle\,\|}}_{(C)} \cdot \lim_{h\to 0}\underbrace{\frac{\|\,\langle\, x_h, y_h \,\rangle - \langle\, x_0, y_0 \,\rangle\,\|}{|h|}}_{(D)}, \tag{4}$$

if the two limits on the right exist. Furthermore, the limit of (C) exists and equals *zero*; this follows immediately from equation (1). To finish the proof, I need only validate the step of distributing the limit across the product in (4)—that is, it will be enough simply to show that the limit of (D) exists. Its value will not matter.

**Step 4: The limit of (D).**

$$\begin{aligned}
\lim_{h\to 0}(D) &= \lim_{h\to 0}\frac{\|\,\langle\, x_h, y_h \,\rangle - \langle\, x_0, y_0 \,\rangle\,\|}{|h|} \\
&= \lim_{h\to 0}\left\|\, \frac{1}{h}\,(\,\langle\, x_h, y_h \,\rangle - \langle\, x_0, y_0 \,\rangle\,)\,\right\| \\
&= \left\|\, \lim_{h\to 0}\frac{1}{h}\big(\vec{C}(t_0 + h) - \vec{C}(t_0)\big)\,\right\| \\
&= \left\|\, \vec{C}\,'(t_0)\,\right\|. \quad \blacksquare
\end{aligned}$$

[1] This additional assumption is what makes this an *almost* proof. One needs a tricky argument involving properties of continuous functions to include the case in which the curve hits $(x_0, y_0)$ multiple times; this same problem arises in the proof of the one–dimensional Chain Rule. I will not bother you with the details.

**III. The gradient.** It is important to note that (2) can be written as a dot product:

$$w'(t_0) = k_x(x_0, y_0)f'(t_0) + k_y(x_0, y_0)g'(t_0) = \langle\, k_x(x_0, y_0), k_y(x_0, y_0)\,\rangle \cdot \langle\, f'(t_0), g'(t_0)\,\rangle\,.$$

The second vector in the dot product is an old friend, $\vec{C}'(t_0)$; this is the contribution of $\vec{C}$ to $w'(t_0)$. The contribution of $k$ is the vector on the left; this is something new.

**Definition.** *The* **gradient** *of $k$ at $(x_0, y_0)$, denoted "$\nabla k(x_0, y_0)$," is the vector $\langle\, k_x(x_0, y_0), k_y(x_0, y_0)\,\rangle$.*

With this definition, the chain rule can be stated:

$$w'(t_0) = \nabla k(\vec{C}(t_0)) \cdot \vec{C}'(t_0).$$

This way of viewing the chain rule is important; stay tuned for details.