1 Recap

- Defined differentiability (after examples indicating things can go wrong).
- Computed the derivative and proved continuity.
- Meaning of gradient.

2 Differentiability

A sufficient condition for differentiability: Differentiability seems like a pain in the neck to check. Fortunately, we have a *sufficient* (but not *necessary*) condition that helps us. Theorem: Suppose $f : S \subset \mathbb{R}^n \to \mathbb{R}$ is a scalar field and $\vec{a} \in S$ is an interior point. Suppose the partials $f_{x_1}, f_{x_2}, \ldots, f_{x_n}$ *exist* in an open ball $B(\vec{a}, r) \subset S$ and they are continuous at $\vec{a}$. Then f is differentiable in the multivariable sense at $\vec{a}$. Such functions are said to be *continuously differentiable* or $\mathcal{C}^1$.

Examples:

- If $f(x), g(y)$ are differentiable functions on $\mathbb{R}$ with continuous derivatives then $h(x, y) = f(x)g(y)$ is differentiable in the multivariable sense. Indeed, h_x, h_y exist and by continuity laws, they are continuous.
- By the one-variable chain rule and continuity laws, a linear combination of functions like $f(x)^k g(y)^l$ is also differentiable.
- As a consequence, polynomials are differentiable on all of $\mathbb{R}^n$.
- Rational functions are differentiable wherever their denominator is non-zero.

Proof: The only candidate for the total derivative at $\vec{a}$ is surely the linear map $\vec{v} \to \langle \nabla f(\vec{a}), \vec{v} \rangle$. Let us prove for the special case of $f(x, y)$ first. $f(a + h, b + k) - f(a, b)$ must be proved to be $\nabla_{\vec{v}} f(a, b) + \|(h, k)\| E$ where $E \to 0$ as $(h, k) \to (0, 0)$. $f(a + h, b + k) - f(a, b) = f(a + h, b + k) - f(a, b + k) + f(a, b + k) - f(a, b) = I + II$. I: By the Lagrange MVT $I = \frac{\partial f}{\partial x}(a + \theta_1, b + k)h$ and $II = \frac{\partial f}{\partial y}(a, b + \theta_2)k$, where $\theta_1 \in (0, h)$ and $\theta_2 \in (0, k)$. Roughly speaking, when h, k are small, I is almost $f_x(a, b)h$ and II is almost $f_y(a, b)k$ by the assumption of continuity of the partial derivatives.
More rigorously, $f(a + h, b + k) - f(a, b) - f_x(a, b)h - f_y(a, b)k = (I - f_x(a, b)h) + (II - f_y(a, b)k)$. Hence, when $\|(h, k)\| < \delta$ (which immediately implies that $|h| < \delta, |k| < \delta$), then by continuity of f_x, f_y, $|f_x(a + \theta_1, b + k) - f_x(a, b)| < \frac{\epsilon}{2}$ and $|f_y(a, b + \theta_2) - f_y(a, b)| < \frac{\epsilon}{2}$. Thus $|(I - f_x(a, b)h)| < |h|\frac{\epsilon}{2}$ and $|II - f_y(a, b)k| < |k|\frac{\epsilon}{2}$. Thus, $\frac{|f(a+h,b+k) - f(a,b) - f_x(a,b)h - f_y(a,b)k|}{\|(h,k)\|} < \epsilon$. This implies the result in this case.
When we have n variables $x_1, \ldots, x_n$, the proof is similar. Indeed, write $f(a_1 + h_1, \ldots) - f(a, b)$ as a sum $I + II + \ldots$ where $I = f(a_1 + h_1, \ldots) - f(a_1, \ldots)$, etc. For each of the n summands, use Lagrange's MVT to get partials into the picture. For each of the partials, we can replace them by their values at $\vec{a}$ at the cost of an error $\frac{\epsilon}{n}$ provided $\vec{h}$ is small enough. The same manipulations as before show what we need.
Recall that if $h(x) = \sin(x^2)$ then $f'(x) = \cos(x^2)2x$. That is, if $f : \mathbb{R} \to \mathbb{R}$ is differentiable and $g : \mathbb{R} \to \mathbb{R}$ is differentiable, then $f \circ g : \mathbb{R} \to \mathbb{R}$ is differentiable and

$(f \circ g)'(x) = f'(g(x))g'(x)$.
The rough idea of the proof is as follows. $g(x+h) \approx g(x) + hg'(x)$ when h is small. $f(y+k) \approx f(y) + kf'(y)$ when k is small. So $f(g(x+h)) \approx f(g(x) + hg'(x))$ which is $f(g(x)) + hg'(x)f'(g(x))$ when h is small. Thus $\frac{f(g(x+h))-f(g(x))}{h} \approx g'(x)f'(g(x))$ when h is small. Of course, one has to make the above rigorous using δs and ϵs.
There is a genuine need for a higher-variable chain rule. Here are two examples where such a rule might help.
Suppose a particle is moving along a path $\vec{r}(t)$ in a room. One question is what rate of temperature rise will the particle experience? That is, suppose $T(x,y,z)$ is the temperature (presumably an infinitely differentiable function) and $\vec{r}(t) = (x(t), y(t), z(t))$ is the trajectory (again presumably highly differentiable), then what is $\frac{dT(x(t),y(t),z(t))}{dt}$?
Consider the polar coordinates $x = r\cos(\theta), y = r\sin(\theta)$, i.e., $r^2 = x^2+y^2$ and $\tan(\theta) = \frac{y}{x}$. (By the way, they make sense only away from the positive x-axis and the origin.) Again, let's assume $T(x,y)$ is the temperature of a hot circular plate. So $\tilde{T}(r,\theta) = T(x(r,\theta), y(r,\theta))$ is a function. We want $\frac{\partial \tilde{T}}{\partial r}, \frac{\partial \tilde{T}}{\partial \theta}$ in terms of $\frac{\partial T}{\partial x}, \frac{\partial T}{\partial y}$. What we want is $\lim_{h\to 0} \frac{T(\vec{r}(t+h))-T(\vec{r}(t))}{h}$. Note that $x(t+h) \approx x(t) + hx'(t)$ when h is small. Likewise for $y(t), z(t)$, i.e., $\vec{r}(t+h) \approx \vec{r}(t) + h\vec{r'}(t)$. Now $T(x+\Delta x, y+\Delta y, z+\Delta z) \approx T(x,y,z) + \Delta x T_x + \Delta y T_y + \Delta z T_z$ (by definition of differentiability). Taking $\Delta x = x'(t)h$ and likewise for y, z, we see that $T(\vec{r}(t+h)) \approx T(\vec{r}(t)) + h(x'(t)T_x + y'(t)T_y + z'(t)T_z)$, i.e., $\frac{T(\vec{r}(t+h))-T(\vec{r}(t))}{h} \approx x'(t)T_x + y'(t)T_y + z'(t)T_z = \langle \nabla T, \vec{r'}(t)\rangle = \nabla_{\vec{r'}(t)} T$.

Theorem: Let $f(\vec{r}) : S \subset \mathbb{R}^n \to \mathbb{R}$ be a scalar field. Let $\vec{r}(t) : (a,b) \in \mathbb{R} \to \mathbb{R}^n$ be a vector-valued function. Define the composition $h(t) : (a,b) \to \mathbb{R}$ as $h(t) = f(\vec{r}(t))$. Suppose $t_0 \in (a,b)$ is a point where $x_1(t), x_2(t), \ldots$ are differentiable functions and f is differentiable at $\vec{r}(t_0)$. Then $h(t)$ is differentiable at t_0 and $h'(t_0) = \langle \nabla f(\vec{r}(t_0)), \vec{r'}(t_0)\rangle = \nabla_{\vec{r'}(t_0)} f(\vec{r}(t))$.

Examples:

- If a path is a *regular* path, i.e., $\vec{r'}(t) \neq 0 \ \forall\ t$, then $\frac{1}{\|\vec{r'}(t)\|}\langle \nabla f(\vec{r}(t_0)), \vec{r'}(t_0)\rangle$ is called the *directional derivative along the curve* and denoted as $\frac{df}{ds}$ (the change in f per metre of the curve). For instance, if $f(x,y) = x^2 - 3xy$ and the path is $(t, t^2 - t + 2)$, and we want to find $\frac{df}{ds}|_{t=1}$, then we calculate as follows. $\nabla f = (f_x, f_y) = (2x - 3y, -3x)$ which at $t = 1$ is $\nabla f(\vec{r}(1)) = (-4, -3)$ and $\vec{r'}(t) = (1, 2t-1)$ which leads to $\vec{r'}(1) = (1,1)$. Thus $\frac{df}{ds}|_{t=1} = \frac{1}{\sqrt{2}}\langle \nabla f, \vec{r'}(t)\rangle = \frac{1}{\sqrt{2}}(-4,-3).(1,1) = \frac{-7}{\sqrt{2}}$.