Here is another useful corollary of the chain rule. A function f on $\mathbb{R}^n$ is called (positively) **homogeneous** of degree a ($a \in \mathbb{R}$) if $f(t\mathbf{x}) = t^a f(\mathbf{x})$ for all $t > 0$ and $\mathbf{x} \neq \mathbf{0}$.

$$f(x,y) = x^2 + 3xy + 7y^2$$

$$f(tx,ty) = t^2x^2 + 3tx\,ty + 7t^2y^2 = t^2 f(x,y)$$

2.36 Theorem (Euler's Theorem). *If f is homogeneous of degree a, then at any point $\mathbf{x}$ where f is differentiable we have*

$$\boxed{x_1\partial_1 f(\mathbf{x}) + x_2\partial_2 f(\mathbf{x}) + \cdots + x_n\partial_n f(\mathbf{x}) = af(\mathbf{x}).}$$

Case $a=2$

$$\nabla f(x,y) = \left(\partial_1 f\begin{bmatrix}x\\y\end{bmatrix}, \partial_2 f\begin{bmatrix}x\\y\end{bmatrix}\right) = (2x+3y,\ 3x+14y)$$

$$\begin{bmatrix}x\\y\end{bmatrix}\cdot\nabla f\left(\begin{bmatrix}x\\y\end{bmatrix}\right) = \begin{bmatrix}x\\y\end{bmatrix}\cdot\begin{bmatrix}2x+3y\\3x+14y\end{bmatrix} = x(2x+3y)+y(3x+14y)$$

$$= 2x^2+3xy+3xy+14y^2 = 2(x^2+3xy+7y^2) = 2f(x,y)$$

Proof: By the chain rule.

$$\frac{\partial\varphi}{\partial t_k} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial t_k} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial t_k},$$

Let $\varphi(t) = f(tx)$

$$\varphi'(t) = \nabla f(tx)\cdot\frac{d\,tx}{dt} = \nabla f(tx)\cdot x$$

Because f is homogeneous of degree a then $f(tx) = t^a f(x)$

$$\varphi'(t) = \frac{d}{dt}t^a f(x) = a t^{a-1} f(x) = \frac{a}{t}t^a f(x) = \frac{a}{t}f(tx).$$

Thus $\quad a t^{a-1} f(x) = \nabla f(tx)\cdot x$

Now set $t=1$ so

$$\boxed{af(x) = \nabla f(x)\cdot x}$$

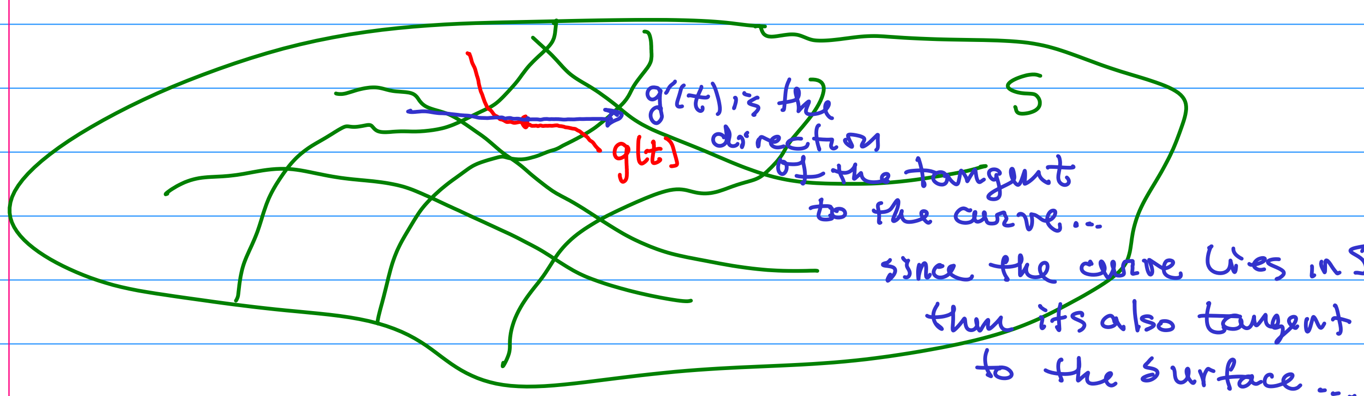
$$\boxed{\underbrace{x_1\partial_1 f(\mathbf{x}) + x_2\partial_2 f(\mathbf{x}) + \cdots + x_n\partial_n f(\mathbf{x})}_{\nabla f(x)\cdot x} = af(\mathbf{x}).}$$

We conclude this section with an additional geometric insight into the meaning of the gradient of a function. If F is a differentiable function of $(x, y, z) \in \mathbb{R}^3$, the locus of the equation $F(x, y, z) = 0$ is typically a smooth two-dimensional surface S in $\mathbb{R}^3$. (We shall consider this matter more systematically in Chapter 3.) Suppose that $(x, y, z) = \mathbf{g}(t)$ is a parametric represention of a smooth curve on S. On the one hand, by the chain rule we have $(d/dt)F(\mathbf{g}(t)) = \nabla F(\mathbf{g}(t)) \cdot \mathbf{g}'(t)$. On the other hand, since the curve lies on S, we have $F(\mathbf{g}(t)) = 0$ for all t and hence $(d/dt)F(\mathbf{g}(t)) = 0$. Thus, for any curve on the S, the gradient of F is orthogonal to the tangent vector to the curve at each point on the curve. Since such curves can go in any direction on the surface, we conclude that at any point $\mathbf{a} \in S$, $\nabla F(\mathbf{a})$ is orthogonal to every vector that is tangent to S at $\mathbf{a}$. (Of course, this is interesting only if $\nabla F(\mathbf{a}) \neq \mathbf{0}$.) We summarize:

$S = \{x \in \mathbb{R}^3 : F(x) = 0\} \subseteq \mathbb{R}^3$

surface...

$g : \mathbb{R} \to S$



Note. $F(g(t)) = 0$ By chain rule

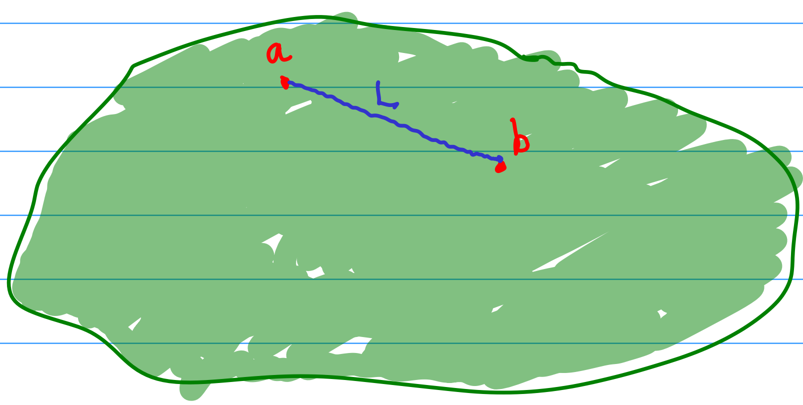
$$\frac{d}{dt} F(g(t)) = \underbrace{\nabla F(g(t))}_{\text{perpendicular to the tangents...}} \cdot \underbrace{g'(t)}_{\text{tangent vector of } S} = \frac{d}{dt} 0 = 0$$

this is normal vector to the surface...

2.4 The Mean Value Theorem

2.39 Theorem (Mean Value Theorem III). *Let S be a region in $\mathbb{R}^n$ that contains the points $\mathbf{a}$ and $\mathbf{b}$ as well as the line segment L that joins them. Suppose that f is a function defined on S that is continuous at each point of L and differentiable at each point of L except perhaps the endpoints $\mathbf{a}$ and $\mathbf{b}$. Then there is a point $\mathbf{c}$ on L such that*

$$f(\mathbf{b}) - f(\mathbf{a}) = \nabla f(\mathbf{c}) \cdot (\mathbf{b} - \mathbf{a}).$$



$f : S \to \mathbb{R}$

f is differentiable at each point of L.

Let $l : [0,1] \to S$ $\quad l(t) = a + t(b-a)$

$l(0) = 0$ and $l(1) = a + b - a = b$

$g(t) = f \circ l(t)$ $\quad l'(t) = b - a$

$$\varphi'(t) = \nabla f(l(t)) \cdot l'(t) = \nabla f(l(t)) \cdot (b-a)$$

Note $\varphi : [0,1] \to \mathbb{R}$ so by the mean value theorem for scalar functions there is $\tau \in (0,1)$ such that

$$\varphi(1) - \varphi(0) = \varphi'(\tau)(1-0)$$

$$f(b) - f(a) = \nabla f(\underbrace{l(\tau)}) \cdot (b-a)$$

Let $c = l(\tau)$ then c is on the line segment and

$$f(\mathbf{b}) - f(\mathbf{a}) = \nabla f(\mathbf{c}) \cdot (\mathbf{b} - \mathbf{a}).$$

Definition

tion. A set $S \subset \mathbb{R}^n$ is called **convex** if whenever $\mathbf{a}, \mathbf{b} \in S$, the line segment from $\mathbf{a}$ to $\mathbf{b}$ also lies in S. Clearly every convex set is arcwise connected (line segments

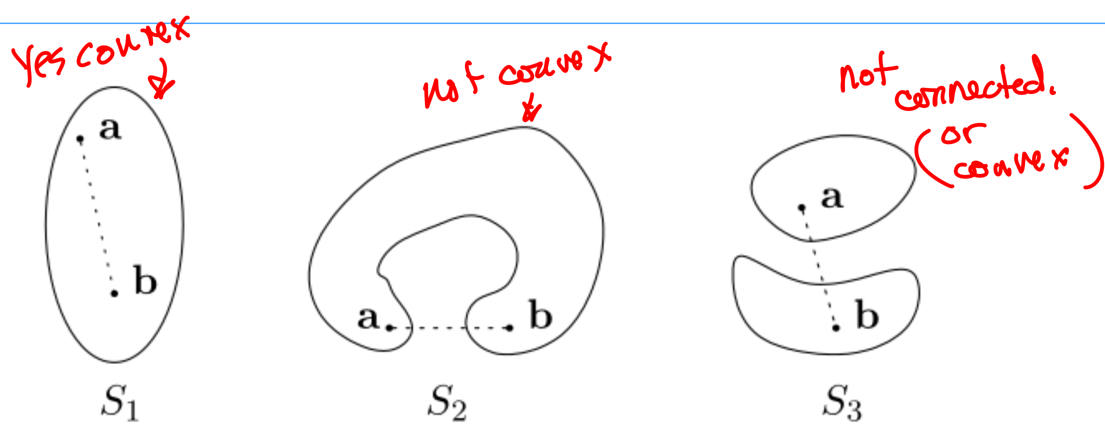


FIGURE 2.4: A convex set (S_1), a set that is connected but not convex (S_2), and a disconnected set (S_3).

allows to involve ∇f with difference between $f(a)$ and $f(b)$

2.40 Corollary. *Suppose that f is differentiable on an open convex set S and $|\nabla f(\mathbf{x})| \le M$ for every $\mathbf{x} \in S$. Then $|f(\mathbf{b}) - f(\mathbf{a})| \le M|\mathbf{b} - \mathbf{a}|$ for all $\mathbf{a}, \mathbf{b} \in S$.*

$$|f(\mathbf{b}) - f(\mathbf{a})| = |\nabla f(\mathbf{c}) \cdot (\mathbf{b} - \mathbf{a})| \le |\nabla f(c)||b-a| \le M|b-a|$$

Surprisingly useful...

Note convex implies connected.

2.41 Corollary. *Suppose f is differentiable on an open convex set S and $\nabla f(\mathbf{x}) = \mathbf{0}$ for all $\mathbf{x} \in S$. Then f is constant on S.*

$|f(b) - f(a)| \le 0|b-a|$ means $f(b) = f(a)$

actually connected is enough...

2.42 Theorem. *Suppose that f is differentiable on an open connected set S and $\nabla f(\mathbf{x}) = \mathbf{0}$ for all $\mathbf{x} \in S$. Then f is constant on S.*

Proof: Idea Fix $a \in S$ and

define $S_1 = \{x \in S : f(x) = f(a)\}$

$S_2 = S \setminus S_1 = \{x \in S : f(x) \ne f(a)\}$.

not 2d.

We know that S is connected so S_1, S_2 better not be a disconnection of S.

means $f(x) \ne f(a)$

Claim S_2 is open: Let $x \in S_2$. Since $S_2 \subseteq S$ then there is $\rho > 0$ such that $B(\rho, x) \subseteq S$.

Since f is continuous Given $\varepsilon = \dfrac{|f(x) - f(a)|}{2} > 0$ there is $\delta > 0$ such that $|x-y| < \delta$ and $y \in S$ implies that $|f(x) - f(y)| < \varepsilon$.

Consequently

$$|f(y)-f(a)| \geq |f(x)-f(a)| - |f(x)-f(y)| \geq 2\varepsilon - \varepsilon = \varepsilon > 0$$

so $y \in S_2$. $\quad r = \min(\delta, \rho)$

then $B(r, x) \subseteq S$ and every $y \in B(r, x)$ is also in S_2

so $B(r, x) \subseteq S_2$

This means S_2 is open... Thus $\overline{S_1} \cap S_2 = \emptyset$.

Next show $S_1 \cap \overline{S_2} = \emptyset$. ....