1. Differentiation and Tangency.

Definition 1.1. Suppose X and Y are normed vector spaces, A is a subset of X and $f : A \to Y$.

For each $v \in X$ we let

$$\partial_v f = \{(a, w) \in \mathbf{int}\, A \times Y : w = \lim_{t \to 0} \frac{1}{t}[f(a + tv) - f(a)]\}.$$

Owing to the uniqueness of limits we find that $\partial_v f$ is a function with values in Y which we call the **partial derivative of f with respect to v**. In case $X = \mathbf{R}^n$ and $i \in \{1, \ldots, n\}$ we let

$$\partial_i f = \partial_{\mathbf{e}_i} f$$

and call this function the ***i*-th partial derivative of f**; here $\mathbf{e}_i$ is the i-th standard basis vector in $\mathbb{R}^n$.

Proposition 1.1. Suppose $a \in \mathbf{int}\, A$. Then

(i) $a \in \mathbf{dmn}\, \partial_0 f(a)$ and $\partial_0 f(a) = 0$;
(ii) if $v \in X$, $a \in \mathbf{dmn}\, \partial_v f$ and $c \in \mathbf{R}$ then $a \in \mathbf{dmn}\, \partial_{cv} f$ and $\partial_{cv} f(a) = c\partial_v f(a)$;
(iii) if $X = \mathbb{R}$ then $a \in \mathbf{dmn}\, \partial_1 f(a)$ if and only if f is f is differentiable at a in which case $f'(a) = \partial_1 f(a)$.

Proof. (i) and (iii) are immediate. To prove (ii) we suppose $v \in X$, $a \in \mathbf{dmn}\, \partial_v f$,$c \in \mathbf{R} \sim \{0\}$ and $\epsilon > 0$. Choose $\delta > 0$ such that

$$|\frac{1}{t}[f(a + tv) - f(a)] - \partial_v f(a)]| \le \frac{\epsilon}{|c|}$$

whenever $a + tv \in A$ and $0 < |t| < \epsilon$. Then

$$|\frac{1}{u}[f(a + ucv) - f(a)] - c\partial_v f(a)]| = |c||\frac{1}{cu}[f(a + ucv) - f(a)] - \partial_v f(a)]| \le |c|\epsilon = \epsilon$$

whenever $a + ucv \in A$ and $0 < |u| < \frac{\delta}{|c|}$. □

We let

$$\partial f = \{(a, L) \in \mathbf{int}\, A \times \mathbf{B}(X; Y) : \lim_{x \to a} \frac{|f(x) - f(a) - L(x - a)|}{|x - a|} = 0\}.$$

Owing to the uniqueness of limits we find that ∂f is a function with values in $\mathbf{B}(X; Y)$ which we call the **differential of f**. We say f **is differentiable at** a if $a \in \mathbf{dmn}\, \partial f$.

Proposition 1.2. Suppose f is diffenentiable at a. Then

(i) $v \in X$ then $a \in \mathbf{dmn}\, \partial_v f$ and $\partial f(a)(v) = \partial_v f(a)$ for any $v \in X$ and
(ii) f is continuous at a.

Proof. (i) holds trivially if $v = 0$ so suppose $v \in X \sim \{0\}$. For any $t \in \mathbb{R} \sim \{0\}$ we use the linearity of $\partial f(a)$ to write

$$|\frac{1}{t}[f(a + tv) - f(a)] - \partial f(a)(v)| = |v|\frac{|f(a + tv) - f(a) - \partial f(a)(tv)|}{|tv|}$$

which should make (i) evident.

For any $x \in A$ we have

$$\begin{aligned}|f(x)-f(a)| &\le |f(x)-f(a)-\partial f(a)(x-a)| + |\partial f(a)(x-a)| \\ &\le |x-a|\left(\frac{|f(x)-f(a)-\partial f(a)(x-a)|}{|x-a|}\right) + ||\partial f(a)|||x-a|\end{aligned}$$

which readily implies that

$$\lim_{x\to a} f(x) = f(a)$$

which gives (ii). □

Proposition 1.3. Suppose f is locally constant. Then, for any $a \in \mathbf{int}\, A$, f is differentiable at a and $\partial f(a) = 0$.

Proof. This is trivial. □

Proposition 1.4. Suppose $A = X$ and $f \in \mathbf{B}(X, Y)$. Then, for any $a \in X$, f is differentiable at a and

$$\partial f(a) = f.$$

Proof. Suppose $a, x \in X$. Then

$$\frac{|f(x)-f(a)-f(x-a)|}{|x-a|} = 0.$$

□

Example 1.1. Let $f : \mathbf{R}^2 \to \mathbf{R}$ be such that

$$f(x,y) = \begin{cases} 1 & \text{if } y = x^2, \\ x \neq 0 & \text{else.} \end{cases}$$

I claim that

$$\partial_{(u,v)} f(0,0) = 0 \text{ whenever } (u,v) \in \mathbb{R}2.$$

To verify this, note that $f(0,0) = 0$ and that if $(u,v) \in \mathbf{R}^2$, $u^2 + v^2 = 1$ and $t \neq 0$ then

$$f(t(u,v)) = \begin{cases} 0 & \text{if } v = 0, \\ 0 & \text{if } v \neq 0 \text{ and } |t| < u^2/|v|. \end{cases}$$

But f is not differentiable at $(0,0)$ because f is not continuous at $(0,0)$.

We will elaborate on this example later.

Example 1.2. Let $f(x) = |x|$ for $x \in \mathbf{R}^n$ and suppose $a \in \mathbf{R}^n \sim \{0\}$. I claim that f is differentiable at a and that

$$\partial f(a)(v) = \frac{v \bullet a}{|a|}, \quad v \in \mathbf{R}^n. \tag{1}$$

This will follow directly from the inequality

$$\big||x| - |a| - \frac{(x-a)\bullet a}{|a|}\big| \le 2\frac{|x-a|^2}{|x|+|a|}, \quad x \in \mathbf{R}^n. \tag{2}$$

Indeed, if (2) holds, given $\epsilon > 0$ we let $\delta = \frac{\epsilon}{2|a|}$ and find that if $|x-a| < \delta$ then

$$\frac{|f(x) - f(a) - |a|^{-1}(x-a) \bullet a|}{|x-a|} \le 2\frac{|x-a|}{|x|+|a|} \le \frac{2\delta}{|a|} \le \epsilon.$$

To prove (2) we multiply it by $|a|(|x|+|a|)$ and obtain the equivalent inequality

$$(3) \qquad \big||a|(|x|+|a|)(|x|-|a|) - (|x|+|a|)(x-a) \bullet a\big| \le 2|a||x-a|^2, \quad x \in \mathbf{R}^n.$$

We have

$$\begin{aligned} &|a|(|x|+|a|)(|x|-|a|) - (|x|+|a|)(x-a) \bullet a \\ &\qquad = |a|(x+a) \bullet (x-a) - (|x|+|a|)(x-a) \bullet a \\ &\qquad = \big(|a|(x+a) - (|x|+|a|)a\big) \bullet (x-a) \\ &\qquad = \big(|a|x - |x|a\big) \bullet (x-a). \end{aligned}$$

Moreover,

$$\big||a|x - |x|a\big| = \big||a|(x-a) + (|a|-|x|)a\big| \le |a|\big(|x-a| + \big||x|-|a|\big|\big) \le 2|a||x-a|,$$

establishing (3).

Definition 1.2. Suppose X and Y are metric spaces, A is a subset of X and $f : A \to Y$. We let

$$\mathbf{Lip}(f) = \sup\{\frac{\rho_Y(f(x), f(y))}{\rho_X(x,y)} : x, y \in A \text{ and } x \ne a\}$$

and call this extended real number the **Lipschitz constant of** f.

Lemma 1.1. Suppose $\gamma : [0,1] \to Y$; γ is continuous; $0 \le N < \infty$ and γ is differentiable at each point of $(0,1)$ with

$$|\gamma'(t)| \le N \quad \text{whenever } t \in (0,1).$$

Then

$$(1) \qquad |\gamma(1) - \gamma(0)| \le N.$$

Proof. Suppose $N < \tilde{N} < \infty$ and let

$$T = \{t \in [0,1] : |\gamma(s) - \gamma(0)| \le \tilde{N}s \text{ whenever } 0 \le s \le t\}.$$

Evidently, T is an interval containing 0, $T \subset [0,1]$ and T is closed since γ is continuous. Let

$$(2) \qquad t_1 = \sup T \in T.$$

Were it the case that $t_1 < 1$, we could choose $t_2 \in (t_1, 1]$ such that if $t_1 < t \le t_2$ then $|\gamma(t) - \gamma(t_1)| \le \tilde{N}(t_1 - t)$ because $\gamma'(t_1)$ exists and has norm not exceeding N. But this implies that

$$|\gamma(s)| \le |\gamma(s) - \gamma(t_1)| + |\gamma(t_1)| \le \tilde{N}(s - t_1) + \tilde{N}t_1 = \tilde{N}s$$

whenever $t_1 < s \le t_2$. So $t_2 \in T$ which is incompatible with (2). So $1 = t_1 \in T$ and therefore $|\gamma(1) - \gamma(0)| \le \tilde{N}$, which, owing to the arbitrariness of $\tilde{N}$, implies (1). $\square$

Theorem 1.1. Suppose X and Y are normed vector space, A is a open subset of X, $f : A \to Y$. Suppose that $\partial_v f(a)$ exists for each $(a, v) \in A \times X$ and let

$$M = \sup\{|\partial_v f(a)| : a \in A,\ v \in X \text{ and } |v| = 1\}.$$

Then

$$M \leq \mathbf{Lip}(f)$$

with equality if A is convex.

Proof. Suppose $a \in A$. Choose $\delta > 0$ such that $\{x \in X : |x - a| < \delta\} \subset A$. For any $v \in X \sim \{0\}$ and any real number t with $|t| < \delta/|v|$ we have

$$|t^{-1}(f(a + tv) - f(a))| \leq t^{-1}\mathbf{Lip}(f)|(a + tv) - a| = \mathbf{Lip}(f).$$

Thus $M \leq \mathbf{Lip}(f)$.

Now suppose A is convex, $a, b \in A$ and $M < \infty$. Let $\gamma(t) = f(a + t(b - a)) - f(a),\ 0 \leq t \leq 1$. Because $\partial_{b-a} f(a + t(b - a))$ exists and has norm not exceeding M for each $t \in [0, 1]$ we find that γ is continuous on $[0, 1]$ and differentiable on $(0, 1)$ with

$$|\gamma'(t)| = |\partial_{b-a} f(a + t(b - a))||b - a| \leq M|b - a|,\ 0 < t < 1.$$

Suppose $M < \tilde{M} < \infty$. The Theorem now follows from the previous Lemma. □

Example 1.3. In this example we will see that equality need not hold in the previous Theorem if A is not convex. Let

$$\gamma(t) \ =\ \frac{e^{it}}{t} \in \mathbf{C}, \quad 1 < t < \infty;$$

let

$$L(t) \ =\ \int_1^t |\gamma'(\tau)|\, d\tau;$$

let

$$A \ =\ \{e^{is}\gamma(t) : 1 < t < \infty \text{ and } |s| < \frac{\pi}{2}\} \subset \mathbf{C};$$

note that

$$(-\frac{\pi}{2}, \frac{\pi}{2}) \times (1, \infty) \ni (s, t) \mapsto e^{is}\gamma(t)$$

is univalent; and define $f : A \to \mathbf{R}$ by requiring that

$$f(e^{is}\gamma(t)) \ =\ L(t) \quad \text{if } |s| < \pi \text{ and } 1 < t < \infty.$$

Then A is open, f is differentiable at each point of A and

$$\sup\{||\partial f(\underline{a})|| : a \in A\} < \infty.$$

(Proof?) But, as $L(t) \uparrow \infty$ and as $e^{is}\gamma(t) \to \mathbf{0}$ as $t \uparrow \infty$ we find that $\mathbf{Lip}(f) = \infty$.

Theorem 1.2. Suppose Y is a normed vector space, n is a positive integer, A is a subset of $\mathbf{R}^n$, $f : A \to Y$ and $a \in \mathbf{int}\, A$. Suppose $\partial_i f$ exists in a neighborhood of a and is continuous at a for each $j = 1, \ldots, n$. Then f is differentiable at a and

$$\partial f(a) = \sum_{j=1}^{n} e^j\, \partial_j f(a).$$

Proof. Let $\epsilon > 0$. Choose $\delta > 0$ such that if $B = \Pi_{i=1}(a_i - \delta, a_i + \delta)$ then $B \subset A \cap_{i=1}^n \mathbf{dmn}\, \partial_i f$ and such that if $x \in B$ then

$$|\partial_j f(x) - \partial_j f(a)| \le \epsilon/\sqrt{n} \quad \text{whenever } j = 1, \ldots, n.$$

To prove the Theorem it will suffice to show that

$$(1) \qquad |f(x) - f(a) - \sum_{i=1}^n (x_i - a_i)\partial_i f(a)| \le \epsilon |x - a| \quad \text{whenever } x \in B.$$

Let $x \in B$. For each $j = 1, \ldots, n$ we define $\gamma_j : [0,1] \to A$ at t in $[0,1]$ by letting $\gamma_1(t) = a + t\,(x_1 - a_1)\,e_1$ and requiring that $\gamma_j(t) = \gamma_{j-1}(1) + t\,(x_j - a_j)\,e_j$ if $j > 1$. For each $j = 1, \ldots, n$ we define $g_j : [0,1] \to Y$ at t in $[0,1]$ by letting

$$g_j(t) = f(\gamma_j(t)) - f(\gamma_j(0)) - t\,(x_j - a_j)\,\partial_j f(a).$$

For any $j = 1, \ldots, n$ we have $g_j'(t) = (x_j - a_j)(\partial_j f(\gamma_j(t)) - \partial_j f(a))$ for t in $(0,1)$ which, by the previous Theorem, implies that

$$|g_j(1) - g_j(0)| \le |x_j - a_j|\,\epsilon/\sqrt{n}.$$

Thus

$$|f(x) - f(a) - \sum_{j=1}^n (x_j - a_j)\,\partial_j f(a)| \;=\; |\sum_{j=1}^n g_j(1) - g_j(0)| \;\le\; \epsilon/\sqrt{n} \sum_{j=1}^n |x_j - a_j| \;\le\; \epsilon\,|x - a|$$

so (1) holds. □

Theorem 1.3 (The Chain Rule.)**.** Suppose X, Y and Z are normed vector spaces, A is a subset of X, B is a subset of Y, $f : A \to B$, $g : B \to Z$, $a \in A$, f is differentiable at a and g is differentiable at $f(a)$. Then $g \circ f$ is differentiable at a and

$$\partial(g \circ f)(a) = \partial g(f(a)) \circ \partial f(a).$$

Proof. Suppose $\epsilon > 0$. Let

$$\epsilon_f = \min\{\frac{\epsilon}{2(1 + ||\partial g(f(a))||)}, 1\} \quad \text{and let} \quad \epsilon_g = \frac{\epsilon}{2(1 + ||\partial f(a)||)}.$$

Since f is differentiable at a we may choose $\delta_f > 0$ such that

$$|f(x) - f(a) - \partial f(a)(x - a)| \le \epsilon_f\,|x - a| \quad \text{whenever } x \in A \text{ and } |x - a| < \delta_f.$$

This together with the triangle inequality implies

$$|f(x) - f(a)| \le (1 + ||\partial f(a)||)\,|x - a| \quad \text{whenever } x \in A \text{ and } |x - a| < \delta_f.$$

Since g is differentiable at $f(a)$ we may choose $\delta_g > 0$ such that

$$|g(y) - g(f(a)) - \partial g(f(a))(y - f(a))| \le \epsilon_g\,|y - f(a)| \;\; \text{whenever } y \in B \text{ and } |y - f(a)| < \delta_g.$$

Let

$$\delta = \min\{\frac{\delta_g}{1 + ||\partial f(a)||}, \delta_f\}.$$

If $x \in A$ and $|x - a| < \delta$ then

$$|f(x) - f(a)| \le (1 + ||\partial f(a)||)\,|x - a| \le \delta_g$$

so

$$\begin{aligned}
|g(f(x)) - & g(f(a)) - \partial g(f(a))(\partial f(a)(x-a))| \\
&\le |g(f(x)) - g(f(a)) - \partial g(f(a))(f(x) - f(a))| + |\partial g(f(a))(f(x) - f(a) - \partial f(a)(x-a))| \\
&\le \epsilon_g \, |f(x) - f(a)| + ||\partial g(f(a))|| \, \epsilon_f \, |x-a| \\
&\le (\epsilon_g \, (1 + ||\partial f(a)||) + \epsilon_f \, ||\partial g(f(a))||) \, |x-a| \\
&\le \epsilon \, |x-a|.
\end{aligned}$$

□

The following Theorem is a generalization of the Leibniz rule.

Theorem 1.4. Suppose Y_1, Y_2, and Z are normed vector spaces, $\mu \in \mathbf{L}(Y_1, Y_2; Z)$ and $(b_1, b_2) \in Y_1, \times Y_2$. Then μ is differentiable at (b_1, b_2) and

$$\partial \mu(b_1, b_2)(v_1, v_2) = \mu(v_1, b_2) + \mu(b_1, v_2) \text{ whenever } (v_1, v_2) \in Y_1 \times Y_2.$$

Remark 1.1. Generalize this to $Y_1, Y_2, \ldots, Y_m$.

s

Proof. We set $|(y_1, y_2)| = \max\{|y_1|, |y_2|\}$ for $(y_1, y_2) \in Y_1 \times Y_2$. Let $\epsilon > 0$. Then

$$\begin{aligned}
|\mu(y_1, y_2) - \mu(b_1, b_2) - & [\mu(y_1 - b_1, b_2) + \mu(b_1, y_2 - b_2)]| \\
&= |[\mu(y_1 - b_1, y_2) + \mu(b_1, y_2, -b_2)] - [\mu(y_1 - b_1, b_2) + \mu(b_1, y_2 - b_2)]| \\
&= |\mu(y_1 - b_1, y_2 - b_2) + \mu(b_1, y_2 - b_2)| \\
&\le ||\mu||(|y_1 - b_1| \, |y_2 - b_2| + |b_1| \, |y_2 - b_2|) \\
&\le ||\mu||(|b_1| + |(y_1, y_2) - (b_1, b_2)|)|(y_1, y_2) - (b_1, b_2)| \\
&\le \epsilon \, |(y_1, y_2) - (b_1, b_2)|
\end{aligned}$$

if

$$|(y_1, y_2) - (b_1, b_2)| < \delta = \min\{1, \epsilon/(||\mu||(1 + |b_2|)\}$$

where we have assumed $\mu \neq 0$ since Theorem holds trivially if $\mu = 0$. □

Theorem 1.5. Suppose X is a normed vector space, A is a subset of X and a is a point of A. Suppose Y_i is a normed vector space, A_i is a subset of Y_i, $f_i : A \to Y_i$ and f_i is differentiable at a for each $i = 1, 2$. Then (f_1, f_2) is differentiable at a and

$$\partial(f_1, f_2)(a) = (\partial f_1(a), \partial f_2(a)).$$

Proof. This follows immediately from the Definition. □

Example 1.4. Suppose $A \subset \mathbf{R}^n$, $f : A \to \mathbf{R}$, $g : A \to \mathbf{R}^m$, a is an interior point of A, and f and g are differentiable at a. Then fg is differentiable at a and

$$\partial(fg)(a)(v) = f(a)\partial g(a)(v) + \partial f(a)(v)g(a) \text{ whenever } v \in \mathbf{R}^n.$$

The point here is that $fg = \beta \circ (f, g)$ where β is the bilinear function whose value at $(c, v) \in \mathbf{R} \times \mathbf{R}^n$ is cv.

2. An example.

Let $\phi : \mathbb{R} \to \mathbb{R}$ be a smooth nonzero function such that $\mathbf{spt}\,\phi \subset (0,1)$. Let $c \in (0,1)$ be such that $\phi'(c) \neq 0$.

Let a and b be decreasing sequences of positive real numbers such that $\lim_{n\to\infty} b_n = 0$ and $b_{n+1} < a_n$ for $n \in \mathbb{N}$. Note that the family $\{[a_n, b_n] : n \in \mathbb{N}\}$ is disjointed. For each $n \in N$ let m_n be the midpoint of the interval (a_n, b_n), let $r_n = b_n - m_n$ and let $x_n = m_n + cr_n$. Let $F : \mathbb{R} \to \mathbb{R}$ be such that

$$F(x) = \begin{cases} x_n r_n \phi\left(\frac{x-m_n}{r_n}\right) & \text{if } n \in \mathbb{N} \text{ and } x \in (a_n, b_n), \\ 0 & \text{if } x \in \mathbb{R} \sim \cup_{n=0}^{\infty}(a_n, b_n). \end{cases}$$

Then F is smooth on $\mathbb{R} \sim \{0\}$, of class C^1 on $\mathbb{R}$ and $F'(0) = 0$. Moreover,

$$\lim_{x\to 0} \frac{|F(x)}{|x|^2} = 0.$$

However,

$$F'(a_n) = 0 \quad \text{for } n \in \mathbb{N}$$

and

$$\frac{F'(x_n) - F'(0)}{x_n - 0} = \phi'(c) \quad \text{for } n \in \mathbb{N}$$

so F' is *not* differentiable at 0.

3. A better example?

Suppose $q > 0$. Let

$$f(x) = \begin{cases} x^{q+2}\sin(x^{-q}) & \text{if } x > 0, \\ 0 & \text{if } x \leq 0. \end{cases}$$

$$f'(x) = (q+2)x^{q+1}\sin(x^{-q}) - qx\cos(x^{-q})' \quad \text{if } x > 0.$$

Thus f is of class C^1; $f'(x) = 0$ if $x < 0$;

$$\lim_{x\to 0} \frac{f(x)}{x^2} = 0;$$

$$\frac{f'(x) - f'(0)}{x - 0} = \frac{f'(x)}{x} = (q+1)x^{q+1}\sin(x^{-q}) - q\cos(x^{-q}) \quad \text{for } x > 0;$$

it follows that f' is it not differentiable at 0.