1. Topology and vector spaces

**Definition 1.1.** A topological group is a group $G$ with a topology $\tau$ such that $\cdot : G \times G \to G$, $(g, h) = gh$ and $I : G \to G$, $I(g) = g^{-1}$ is continuous.

**Example 1.2.**

1. Let $G$ be a group and $d(g, h) =\begin{cases} 1 & \text{if } g \neq h \\ 0 & \text{else} \end{cases}$. This induces the discrete metric. In the induced topology every set is open and hence $G$ is a topological group.

2. Let us consider $\mathbb{R}^\mathbb{R}$ with the topology of pointwise convergence. Then $(\mathbb{R}^\mathbb{R}, +)$ is a commutative topological group.

**Definition 1.3.** A topological vector space is a vector space over $K \in \{\mathbb{R}, \mathbb{C}\}$ with a topology on $V$ such that $(V, +)$ is a topological group and $\cdot : K \times V \to V$ is continuous.

Why topology? For differentiation. We need even more.

**Definition 1.4.** Let $V$ be vector spaces with a metric. Let $\Omega \subset V$ be an open set and $f : \Omega \to \mathbb{R}$ be a map. $f$ is called differentiable at $x_0 \in \Omega$ if there exists a continuous linear map $T : V \to \mathbb{R}$ such that

$$
\lim_{d(x_0, x) \to 0} \frac{|f(x) - f(x_0) + T(x - x_0)|}{d(x, x_0)} = 0
$$

**Remark 1.5.** Let $V = \mathbb{R}(\mathbb{N})$ equipped with the pointwise topology. Then a linear map $T : V \to \mathbb{R}$ is continuous if and only if $\sum_k |T(e_k)| < \infty$. Indeed, $C = \{ (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) : \varepsilon \in \{-1, 0, 1\} \} \cap V$ is compact. Thus

$$
\sum_k |T(e_k)| = \sup_{x \in C} |T(x)|
$$

Differentiation is usually done in normed vector spaces.
Definition 1.6. \((V, \| \|)\) is called a normed vector space if \(V\) is a vector space and
\[ \| \| : V \to [0, \infty) \]
satisfies.
\begin{enumerate}
  \item \(\|x\| = 0 \iff x = 0\),
  \item \(\|\lambda x\| = |\lambda|\|x\|\),
  \item \(\|x + y\| \leq \|x\| + \|y\|\).
\end{enumerate}
for all \(x, y \in V\), \(\lambda \in K\). The associated metric on \((V, \| \|)\) is defined by
\[ d_{\|}(x, y) = \|x - y\| .\]

Remark 1.7. For a normed vector space \((V, +)\) is a topological group.

Lemma 1.8. A normed vector space \((V, \| \|)\) is complete if and only if every absolutely convergent series is convergent.

Proof. Let us assume that \(V\) is complete and that
\[ \sum_{n} \|x_n\| < \infty . \]
Using the Cauchy criterion in \(\mathbb{R}\), we find for every \(\varepsilon > 0\) an natural number \(n_0\) such that for \(m \geq n \geq n_0\)
\[ \sum_{k=n+1}^{m} \|x_k\| < \varepsilon . \]
This shows that \(y_n = \sum_{k=1}^{n} x_k\) satisfies
\[ \|y_m - y_n\| = \| \sum_{k=n+1}^{m} x_k\| \leq \sum_{k=n+1}^{m} \|x_k\| < \varepsilon . \]
Thus \((y_n)\) is Cauchy. Since \(V\) is complete we find a limit \(y = \lim_{n} y_n\). For the converse we assume that \((y_n)\) is Cauchy. Passing to a subsequence (if necessary) we may assume \(\|y_{n+1} - y_n\| = d(y_{n+1}, y_n) < 2^{-n}\). Then the series \(x_0 = y_0, x_n = y_n - y_{n-1}\) satisfies
\[ \sum_{n} \|x_n\| < \infty . \]
By assumption, the partial sums
\[ z_n = x_0 + x_1 - x_0 + x_2 - x_1 + \cdots x_n - x_{n-1} = x_n \]
converge. Thus \((x_n)\) is convergent.
Definition 1.9. \( V \) and \( W \) be normed vector spaces. Let \( \Omega \subset V \) be an open set and \( f : \Omega \to W \) be a map. \( f \) is called differentiable at \( x_0 \in \Omega \) if there exists a linear map \( T \) such that
\[
\lim_{\|v\| \to 0} \frac{\|f(x_0 + v) - f(x_0) + T(v)\|}{\|v\|} = 0.
\]

Remark 1.10. If \( f \) is differentiable at \( x_0 \), then \( f \) is continuous at \( x_0 \).

Funny, the derivative is a linear map! In order to understand what it means to be continuously differentiable we need a norm on \( L(X, Y) \).

Proposition 1.11. Let \( X \) be a normed space and \( Y \) be a Banach space. We define \( L(X, Y) \) as the space of map \( T : X \to Y \) which are linear, i.e.
\[
T(x + \lambda y) = T(x) + \lambda T(y).
\]
and continuous. The norm on \( L(X, Y) \) is given by
\[
\|T\|_{op} = \sup_{\|x\| \leq 1} \|T(x)\|.
\]

Then \( L(X, Y) \) is a Banach space.

Proof. Let us first show that a linear map \( T : X \to Y \) is continuous iff \( \|T\| < \infty \). Indeed, if \( \|T\| \) is finite, then
\[
\|T(x) - T(y)\| = \|T(x - y)\| \leq \|T\|_{op}\|x - y\|
\]
holds for all \( x, y \in V \). Thus \( T \) is Lipschitz and thus continuous. For the converse, we assume that \( T \) is continuous. Then \( T^{-1}(B(0,1)) \) is open and henceforth contains \( B(0,\varepsilon) \) for some \( \varepsilon > 0 \). Now let \( \|x\| \leq 1 \) and \( 0 < \delta < \varepsilon \). Then \( \|(\varepsilon - \delta)x\| < \varepsilon \) and hence
\[
\|T(x)\| = (\varepsilon - \delta)^{-1}\|T(\varepsilon - \delta)(x)\| < (\varepsilon - \delta)^{-1}.
\]
This shows that \( \|T\|_{op} \leq (\varepsilon - \delta)^{-1} \) for every \( \delta > 0 \) and thus \( \|T\|_{op} \leq \varepsilon^{-1} \). Now, we observe that \( \|\| \) is a norm. We only check the triangle inequality. Indeed,
\[
\|T + S\|_{op} = \sup_{\|x\| \leq 1} \|(T + S)(x)\| = \sup_{\|x\| \leq 1} \|T(x) + S(x)\| \leq \sup_{\|x\| \leq 1} \|T(x)\| + \|S(x)\| \leq \|T\|_{op} + \|S\|_{op}.
\]
Finally we have to show that \( L(X, Y) \) is complete. Let \( (T_n) \) be a Cauchy sequence of linear maps. For fixed \( x \in X \), we have
\[
\|T_n(x) - T_m(x)\| \leq \|T_n - T_m\|\|x\|.
\]
Thus \((T_n(x))\) is Cauchy and we may define
\[
T(x) = \lim_n T_n(x).
\]
Then we have
\[
T(x + \lambda y) = \lim_n T_n(x + \lambda y) = \lim_n T_n(x) + \alpha T_n(y) = T(x) + \lambda T(y).
\]
Thus \(T\) is linear. Let us show that
\[
\lim_n \|T - T_n\|_{op} = 0. \tag{1.1}
\]
Indeed, let \(x \in X\) with \(\|x\| \leq 1\). Then we have
\[
\|T(x) - T_n(x)\| = \|\lim_m T_m(x) - T_n(x)\| \leq \limsup_{m \geq n} \|T_m(x) - T_n(x)\|
\]
\[
\leq \sup_{m \geq n} \|T_m - T_n\| \|x\| \leq \sup_{m \geq n} \|T_m - T_n\|.
\]
In particular \(\|T\|_{op} \leq \|T - T_1\|_{op} + \|T_1\|_{op}\) is finite and \(T\) is continuous. Moreover, 
\[
\lim_n d(T, T_n) = 0 \implies \lim_n T_n = T.
\]

**Proposition 1.12.** (Chain rule) \(\Omega \subset V\) open, \(\bar{\Omega} \subset W\) open. \(f : \Omega \to W, g : \bar{\Omega} \to Z\), \(x_0 \in \omega, y_0 = f(x_0) \in \bar{\Omega}\). If \(f\) is differentiable at \(x_0\) and \(g\) is differentiable at \(y_0\), then \(g \circ f\) is differentiable and
\[
(g \circ f)'(x_0) = g'(f(x_0))f'(x_0).
\]

**Proof.** Let us introduce the error functions
\[
\varepsilon_f(v) = f(x_0 + v) - f(x_0) - T(v)
\]
and
\[
\varepsilon_g(q) = g(y_0 + w) - g(y_0) - S(w).
\]
Note that
\[
\lim_{\|v\| \to 0} \frac{\|\varepsilon_f(v)\|}{\|v\|} = 0 = \lim_{\|w\| \to 0} \frac{\|\varepsilon_g(w)\|}{\|w\|}.
\]
For \(v \in V\) we introduce \(w = f(x_0 + v) - f(x_0) = T(v) + \varepsilon_f(v)\). Moreover, \(R = S \circ T\). Then we have
\[
g(f(x_0 + v)) - g(f(x_0)) - R(v) = g(f(x_0) + w) - g(f(x_0)) - R(v)
\]
\[
= S(w) + \varepsilon_g(w) - R(v) = S(\varepsilon_f(v)) + \varepsilon_g(w).
\]
Since \(S\) is continuous, we have
\[
\lim_{\|v\| \to 0} \frac{\|S(\varepsilon_f(v))\|}{\|v\|} \leq \|S\| \frac{\|\varepsilon_f(v)\|}{\|v\|} = 0.
\]
Now, we use the standard cancellation trick
\[
\frac{\|\varepsilon_g(w)\|}{\|v\|} = \frac{\|\varepsilon_g(w)\|}{\|w\|} \frac{\|w\|}{\|v\|} = \frac{\|f(x_0 + v) - f(x_0)\|}{\|w\|} \frac{\|w\|}{\|v\|} = \frac{\|\varepsilon_g(w)\|}{\|w\|} \frac{\|T(v) + \varepsilon_f(v)\|}{\|v\|}.
\]
Note that for \(\|w\| = 0\) we have \(\varepsilon_g(w) = 0\) and thus we may assume here that \(\|w\| \neq 0\). Since \(f\) is continuous we have
\[
\lim_{\|v\|\to 0} \|w\| = \lim_{\|v\|\to 0} \|f(x_0 + v) - f(x_0)\| = 0.
\]
Therefore it suffices to show that
\[
\{\frac{\|T(v) + \varepsilon_f(v)\|}{\|v\|} : \|v\| < \delta_1\}
\]
is bounded for \(\delta_1\) small enough. However, we know that there exists a \(\delta_1 > 0\) such that
\[
\frac{\|\varepsilon_f(v)\|}{\|v\|} \leq 1
\]
holds for \(\|v\| \leq \delta_1\). Thus we get
\[
\frac{\|T(v) + \varepsilon_f(v)\|}{\|v\|} \leq \|T\| \|v\| + \|v\| \leq (\|T\| + 1)\|v\|
\]
for \(\|v\| \leq \delta_1\). This implies
\[
\lim_{\|v\|\to 0} \frac{\|\varepsilon_g(w)\|}{\|v\|} = 0.
\]
This clearly implies \((g \circ f)'(x_0) = R = S \circ T\) and the assertion is proved. 

\section*{Example 1.13} \hspace*{1cm} Let \(u\) be continuously differentiable function in two variables. We are looking for the derivative of
\[
h(t) = \int_0^t u(t, s)ds.
\]

\textbf{Solution:} We define \(g(t, r) = \int_0^t u(r, s)ds\). The derivative is given by the gradient
\[
T(x, y) = \nabla g(t, r) \begin{pmatrix} x \\ y \end{pmatrix} = \frac{\partial g}{\partial t}(t, r)x + \frac{\partial u}{\partial r}(t, r)y = u(r, t)x + \left( \int_0^t \frac{\partial u}{\partial r}u(r, s)ds \right)y
\]
We define \( f(t) = (t, t) \). The derivative is \( f'(t) = \begin{pmatrix} 1 & 1 \end{pmatrix} \). Hence, we get

\[
h'(t) = F'(t, t)f'(t) = u(t, t) + \int_0^t \frac{\partial u}{\partial r} u(t, s)ds.
\]

**Remark 1.14.** Let \( T : V \rightarrow W \) be a linear map. Then \( T'(x) = T \) for all \( x \in V \).

**Example 1.15.** Consider \( \det : M_n(\mathbb{R}) \rightarrow \mathbb{R} \). Then

\[
(\det)'(A)(B) = \sum_{k=1}^n \det(A(e_1), ..., B(e_k), ..., A(e_n)).
\]

Here \( B(e_j) \) is the \( j \)-th column and for \( k = 1, \ldots, m \) we replace the \( k \)'th column of \( B \) by the \( k \)-th column of \( A \). Moreover, the derivative of the function

\[
g(t) = \det(1 + tA)
\]

is given by

\[
g'(0) = tr(A) = \sum_{i=1}^n a_{ii}.
\]

Indeed, we note

\[
det(A + B) = \sum_{\sigma} \varepsilon(\sigma) \prod_{i=1}^n (A + B)_{i, \sigma(i)}
\]

\[
= \sum_{\sigma} \varepsilon(\sigma) \prod_{i=1}^n (a_{i, \sigma(i)} + b_{i, \sigma(i)})
\]

\[
= det(A) + \sum_{\sigma} \varepsilon(\sigma) \sum_{k=1} b_{k, \sigma(k)} \prod_{i \neq k} a_{i, \sigma(i)}
\]

+ higher monomials in the coefficient \( b_{ij} \).

This yields the first assertion. For the second that if \( a_{ij} = \delta_{ij} \) only the term \( \sigma = id \) survives and we get

\[
det(1 + tA) = 1 + t \, tr(A) + \text{ higher monomials in } t.
\]

2. **Taylor formula**

We will first consider the Taylor formula for functions with values in \( \mathbb{R} \). Recall the scalar Taylor formula
Theorem 2.1. (Taylor formula with integral remainder) Let $f : I \to \mathbb{R}$ a $(n+1)$-times continuously differentiable. Then

$$f(x_0 + t) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!} t^k + R_n(x_0, t)$$

where

$$R_n(x_0, t) = \frac{1}{n!} \int_0^t f^{(n+1)}(x_0 + s)(t - s)^n ds.$$ 

Lemma 2.2. Let $f : \Omega \to \mathbb{R}$ be $n$ times differentiable in $\Omega$. Let $v \in V$ such that $[x_0, x_0 + v] = \{x_0 + tv : 0 \leq t \leq v\} \subset \Omega$. Define $h(t) = x_0 + tv$. Then the scalar function $g(t) = f(x_0 + tv)$ satisfies

$$g^{(n)}(t) = f^{(n)}(x_0 + tv, \ldots, x_0 + tv).$$

Proof. $n = 1$: By the chain rule we know that

$$g'(t) = f'(x_0 + tv)(v).$$

Now, we consider $f' : \Omega \to L(V, \mathbb{R})$ and an the function $F : \Omega \to \mathbb{R}$ defined by

$$F(x) = f'(x)(v)$$

We apply the chain rule again and get

$$g''(t) = \frac{d}{dt}F(x_0 + tv) = F'(x_0 + tv)(v).$$

In order to calculate this derivative we write $F(x) = e_v(f'(x))$, were $e_v : L(V, \mathbb{R}) \to \mathbb{R}$ is given by $e_v(T) = T(v)$. Since $e_v$ is linear, we deduce from the chain rule

$$F'(x)(w) = e_v \circ f''(x) = f''(x)(v)(w).$$

The general case is proved by induction following the same arguments.

Corollary 2.3. Let $\Omega \subset V$ be an open set and $f : \Omega \to \mathbb{R}$ be $(n+1)$-times continuously differentiable. Let $x_0 \in \Omega$ and $v \in V$ such $[x_0, x_0 + v] \subset \Omega$. Then

$$f(x_0 + v) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!} (v, \ldots, v) + R_n(x_0, v)$$

where

$$R_n(x_0, v) = \frac{1}{n!} \int_0^t f^{(n+1)}(x_0 + sv, \ldots, x_0 + sv)(t - s)^n ds.$$ 

Now, we want to consider function $f : I \to X$ where $X$ is a Banach space.
**Lemma 2.4.** Let $f : [a, b] \to X$ be continuous function. For a partition $\pi = \{a = x_0, \ldots, x_n = b\}$ and $\xi_1, \ldots, \xi_n$ with $\xi_i \in [x_{i-1}, x_i]$ the Riemann sum is given by

$$S(\pi, \xi) = \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}) .$$

Then

$$\int_{a}^{b} f(s)ds = \lim_{\text{mesh}(\pi) \to 0} S(\pi, \xi)$$

exists. Moreover,

$$F(t) = \int_{a}^{t} f(s)ds$$

satisfies $F'(t) = f(t)$.

**Proof.** Since $f$ is uniformly continuous, we can find for $\varepsilon > 0$ a $\delta > 0$ such that $|t - s| < \delta$ implies

$$\|f(t) - f(s)\| < \varepsilon .$$

We consider two partitions such that $\pi_1$ and $\pi_2$ with $\text{mesh}(\pi_1) < \delta$ and $\text{mesh}(\pi_2) < \delta$. Let $\pi = \pi_1 \cup \pi_2$. Let us assume that $\pi_1$ has $n + 1$ points and $\xi = (\xi_1, \ldots, \xi_n)$ is an intermediate vector (i.e. such that $\xi_i \in (x_{i-1}, x_i)$). We assume that $\pi$ has $m + 1$ points $\eta$ is an intermediate vector for $\pi$. Let us agree to use the $\xi$’s for every interval of $\pi$ which is contained in $[x_{i-1}, x_i]$. Then, we get

$$\|S(\pi_1, \xi) - S(\pi, \eta)\| = \|\sum_{i=1}^{n} (f(x_i)(x_i - x_{i-1}) - \sum_{[y_{j-1}, y_j] \subset [x_{i-1}, x_i]} f(\eta_j)(y_j - y_{j-1})\|$$

$$\leq \sum_{i=1}^{n} \sum_{[y_{j-1}, y_j] \subset [x_{i-1}, x_i]} \|f(\xi_i) - f(\eta_j)\|(y_j - y_{j-1}) \leq \varepsilon (b - a) .$$

Thus we get

$$\|S(\pi_1, \xi) - S(\pi_2, \tilde{\xi})\| \leq 2\varepsilon (b - a) .$$

This proves the first assertion. Let us also note the immediate consequences

$$\|\int_{a}^{b} f(s)ds\| \leq \int_{a}^{b} \|f(s)\|ds$$

and

$$\int_{a}^{b} (f(s) + g(s))ds = \int_{a}^{b} f(s)ds + \int_{a}^{b} g(s)ds .$$
For the prove of the second assertion, we observe that
\[
\| \int_a^{b+t} f(s)ds - \int_a^b f(s)ds - f(b)t \| = \| \int_b^{b+t} (f(s) - f(b))ds \|
\leq \int_b^{b+t} \| f(s) - f(b) \| ds .
\]
Given \( \varepsilon > 0 \) we may find \( \delta > 0 \) such that \( |s| < \delta \) implies \( \| f(s) - f(b) \| < \varepsilon \). This yields
\[
\| \varepsilon(t) \| = \| \int_a^{b+t} f(s)ds - \int_a^b f(s)ds - f(b)t \| \leq \varepsilon |t| .
\]
The assertion follows.

Lemma 2.5. Let \([a, b] \subset \bigcup_{x \in [a, b]} B(x, \delta_x)\) be an open cover for \([a, b]\). Then there exists a partition \( \pi = \{ a = t_0 < t_1 < \cdots < t_m = b \} \) such that for every \( i = 1, \ldots, m \) we find \( x_i \in [t_i, t_{i+1}] \) and \( \max\{|x_i - t_i|, |t_{i+1} - x_i|\} < \delta_x \).

Proof. Let us denote by \( S \) the set of all points \( s \in [a, b] \) such that there are \( t_0 = a < t_1 < \cdots < t_m \) with \( x_i \in [t_{i-1}, t_i] \), \( \max\{|x_i - t_i|, |t_{i+1} - x_i|\} < \delta_x \) and \( x_{m+1} \in [t_m, y] \) such that
\[
\max\{|x_{m+1} - t_m|, |y - x_{m+1}|\} < \delta_{x_{m+1}} .
\]
Note that \( S \) is not empty because \( a \in S \). Let \( s = \sup S \). We claim that \( s \in S \).
Indeed, let we may find \( y \in S \) such that \( y > s - \delta_s \). Then we find \( t_0 = a < \cdots t_m < y \) and \( t_m \leq x_{m+1} \leq y \). We may define \( t_{m+1} = y \) and \( x_{m+2} = s \). Also, we must have \( s = b \). Indeed, assume \( s < b \). \( s \in S \) implies that \( s - x_{m+1} < \delta_{x_{m+1}} \). Let \( \rho < \delta_{x_{m+1}} - (s - x_{m+1}) \) such that also \( s + \rho < b \). Then \( s + \rho \in S \) yields a contradiction. Thus \( s = b \) and the assertion is proved.

Lemma 2.6. Let \( f : [a, b] \to X \) be continuously differentiable. Then
\[
\int_a^b f(s)ds = f(b) - f(a) .
\]

Proof. By continuity it suffices to assume that \( f \) is differentiable on an open subset of \([a, b]\). Let \( \varepsilon > 0 \) and \( \delta > 0 \). For every \( x \in [a, b] \) we may find \( \delta_x < \delta \) such that
\[
|t - x| < 2\delta_x \Rightarrow \| f(t) - f(x) - f'(x)(t - x) \| \leq \varepsilon|t - x| .
\]
Then we have \([a, b] \subset B(x, \delta_x)\). By compactness, we may find a finite subset \(\xi_1, \ldots, \xi_n\) such that
\[
[a, b] \subset \bigcup_{i=1}^{n} B(\xi_i, \delta_{\xi_i}) .
\]
We apply Lemma \(\text{Lemma 2.3}^{\text{comp}}\) and find a partition \(\pi = \{a = t_0 < t_1 < \cdots < t_m = b\}\) such that for every \(i = 1, \ldots, m\) we find \(\xi_i \in [t_i, t_{i+1}]\) and \(\max\{|\xi_i - t_i|, |t_{i+1} - \xi_i|\} < \delta_{\xi_i}\).

This yields
\[
\|f(b) - f(a) - S(\pi, \xi)\| = \left\| \sum_{i=1}^{n} f(x_{i+1}) - f(x_i) - f'(\xi_i)(x_{i+1} - x_i) \right\|
\leq \sum_{i=1}^{n} \left\| f(x_{i+1}) - f(\xi_i) - f'(\xi_i)(x_{i+1} - \xi_i) \right\|
+ \left\| f(\xi_i) - f(\xi_{i-1}) - f'(\xi_i)(\xi_i - x_{i-1}) \right\|
\leq \sum_{i=1}^{n} \varepsilon(|(x_{i+1} - \xi_i)| + |\xi_i - x_{i-1}|) \leq \varepsilon(b - a) .
\]

Thus for \(\delta\) small enough we find
\[
\|f(b) - f(a) - \int_{a}^{b} f'(s)ds\| \leq \|f(b) - f(a) - S(\pi, \xi)\| + \left\| \int_{a}^{b} f'(s)ds - S(\pi, \xi)\right\|
< \varepsilon(b - a) + \varepsilon .
\]

Since \(\varepsilon > 0\) is arbitrary, we deduce the assertion.

**Corollary 2.7.** Let \(X\) be a Banach space and \(f : (a, b) \to X\) be \((n + 1)\) times continuously differentiable. Then
\[
f(x_0 + t) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!} + R_n(x_0, t)
\]
where
\[
R_n(x_0, t) = \frac{1}{n!} \int_{0}^{t} f^{(n+1)}(x_0 + s)(t - s)^n ds .
\]

**Proof.** We use induction on \(n\). For \(n = 1\) we deduce from Lemma \(\text{Lemma 2.6}^{\text{Kund2}}\) that
\[
f(x_0 + t) - f(x_0) = \int_{0}^{t} f'(x_0 + s)ds .
\]
For \(n = 2\) we consider
\[
F'(t) = f'(x_0 + t) .
\]
Then we have
\[
F'(t) = f''(x_0 + t)t + f'(x_0 + t) .
\]
This yields
\[
\begin{align*}
  f(x_0 + t) - f(x_0) &= \int_0^t f'(x_0 + s)ds = \int_0^t (F'(s) - f''(x_0 + s)s)ds \\
  &= F(t) - F(0) - \int_0^t f''(x_0 + s)ds = f'(x_0 + t)t - \int_0^t f''(x_0 + s)ds \\
  &= \int_0^t f''(x_0 + s)(t - s)ds.
\end{align*}
\]

For general \( n \) it is best to use integration by parts (justified as above):
\[
\begin{align*}
  \frac{1}{(k-1)!} \int_0^t f^{(k)}(x_0 + s)(t - s)^{k-1}ds \\
  \leq \left[-\frac{1}{k!}f^{(k)}(x_0 + s)(t - s)\right]_0^t + \frac{1}{k!} \int_0^t f^{(k+1)}(x_0 + s)(t - s)^k ds \\
  = \frac{f^{(k)}(x_0)t^k}{k!} + \frac{1}{k!} \int_0^t f^{(k+1)}(x_0 + s)(t - s)^k ds
\end{align*}
\]
Iterating yields assertion.

**Corollary 2.8.** Let \( V \) be a normed space and \( X \) be a Banach space. Let \( \Omega \subset V \) be open, \( f : \Omega \to X \) be \((n + 1)\)-times (continuously) differentiable. Let \( x_0 \in \Omega \) such that \( B(x_0, \delta) \subset \Omega \).

Then
\[
  f(x_0 + v) = \sum_{k=0}^n \frac{1}{k!} f^{(k)}(x_0)(v, \ldots, v) + R_n(x_0, v)
\]
such that
\[
  \|R_n(x_0, v)\| \leq \frac{1}{(n+1)!} \sup_{y \in B(x_0, \delta)} \|f^{(n+1)}(y)\| \|v\|^{n+1}
\]
holds for all \( \|v\| \leq \delta \).

**Proof.** Since \( \Omega \) is open we may \( \delta > 0 \) such that \( B(x_0, \delta) \subset \Omega \). Let \( \|v\| \leq \delta \) and consider the function \( g(t) = f(x_0 + tv) \). Then \( g \) is \((n + 1)\) times continuously differentiable and we get
\[
  g(1) = \sum_{k=0}^n \frac{g^{(k)}(0)}{k!} + R_n(0, 1).
\]
As in Corollary 2.3 we find
\[
  g^k(t) = f^{(k)}(x_0 + tv)(v, \ldots, v).
\]
Therefore, we get
\[ \| R_n(0,1) \| = \left\| \frac{1}{n!} \int_0^1 g^{(n+1)}(s)(1-s)^n ds \right\| \]
\[ = \left\| \frac{1}{n!} \int_0^1 f^{(n+1)}(x_0 + sv)(v, \cdots, v)(1-s)^n ds \right\| \]
\[ \leq \frac{1}{n!} \sup_{y \in B(x_0, \delta)} \| f^{(n+1)}(y) \| \| v \|^{n+1} \int_0^1 (1-s)^n ds \]
\[ = \frac{1}{(n+1)!} \sup_{y \in B(x_0, \delta)} \| f^{(n+1)}(y) \| \| v \|^{n+1} . \]
This is exactly the estimate for the remainder claimed in the assertion. \[\square\]

A power series with values in a Banach space \( X \) is given by
\[ f(t) = \sum_{k=0}^{\infty} x_k (t - t_0)^k \]
where \( x_k \in X \). We will focus on \( t_0 = 0 \). The radius of convergence is given by
\[ R = (\limsup_n \| x_k \|^\frac{1}{2})^{-1} . \]

**Remark 2.9.** Let \( 0 \leq r < u < R \). Then for \( n \geq n_0 \)
\[ \sum_{k \geq n} |t|^k \| x_k \| \leq \frac{(\frac{r}{u})^n}{1 - r/u} . \]
Thus \( f \) is a convergent sum of continuous functions and hence continuous.

**Lemma 2.10.** On \((-R, R)\) \( f \) is infinitely often differentiable and
\[ f^{(n)}(t) = \sum_{k=n}^{\infty} \frac{k!}{n!} x_k (t - t_0)^{k-n} \]

**Proof.** We assume \( t_0 = 0 \). Let \( 0 < r < R \) and \( |t| < r \). Let \( |s| < r \). Let \( r < u \). Then we find \( n_0 \) such that \( \| x_k \| \leq u^{-k} \). First we observe that for \( k \geq 2 \)
\[ |s^k - t^k - k t^{k-1} (s-t) | = \left| \int_t^s k(v^{k-1} - t^{k-1}) dv \right| \]
\[ = | \int_t^s k(k-1) \int_t^v w^{k-2} dw dv | \leq k(k-1) \int_t^s \int_t^v w dv dw \]
\[ = \frac{k(k-1)}{2} \int_t^s (s-t)^2 \]
Then we get for all \( n \geq n_0 \)
\[ \left\| \sum_{k \geq n} s^k x_k - \sum_{k \geq n} t^k x_k - \sum_{k \geq n} t^{k-1} k x_k (s-t) \right\| \]
2. TAYLOR FORMULA

\[\begin{align*}
= & \left\| \sum_{k \geq n} (s^k - t^k - kt^{k-1}(s - t))x_k \right\| \\
\leq & \sum_{k \geq n} |s^k - t^k - kt^{k-1}(s - t)|\|x_k\| \\
\leq & \sum_{k \geq n} |t^k - s^k - kt^{k-1}(t - s)|u^{-k} \\
\leq & \sum_{k \geq n} \frac{k(k-1)}{2} t^{k-2}|s - t|^2 u^{-k} \\
\leq & \frac{|s - t|^2}{2} \sum_{k \geq n} k(k-1)(\frac{r}{u})^k.
\end{align*}\]

Note that the sum on the right hand side is convergent. This yields

\[\left( \sum_{k \geq n} t^k x_k \right)' = \sum_{k \geq n} t^{k-1} k x_k.\]

Differentiating the polynomial \(p(t) = \sum_{k=0}^{n_0} t^k x_k\) is no threat and the assertion follows for \(n = 1\). Induction yields the assertion in full generality. \(\blacksquare\)