\[
\lim_{|h| \to 0} \sum_{i=1}^{n} \frac{|D_i f(a) - D_i f(a)|}{|h|} \leq 0
\]

since \(D_i f\) is continuous at \(a\).

Here is a very useful application of the chain rule:

**Theorem:** \(g_1, g_2, \ldots, g_m : \mathbb{R}^n \to \mathbb{R}\) continuously diff at \(a\), and let \(F : \mathbb{R}^n \to \mathbb{R}\) be diff at 

\((g_1(a), \ldots, g_m(a))\). Define \(F : \mathbb{R}^n \to \mathbb{R}\),

\[F(x) = f(g_1(x), \ldots, g_m(x))\]. Then

\[D_i F(a) = \sum_{j=1}^{m} D_i f(g_j(a)) \cdot D_i g_j(a)\]

**Proof:** \(F = f \circ g, \ g = (g_1, g_2, \ldots, g_m)\) diff at \(a\).

\[F'(a) = f'(g(a)) \circ g'(a) = (D_i f(g(a)), \ldots, D_m f(g(a))) \cdot \begin{pmatrix} D_i g_1(a) & \cdots & D_i g_m(a) \\ \vdots & \ddots & \vdots \\ D_m g_1(a) & \cdots & D_m g_m(a) \end{pmatrix}\]

and we are done! (Why?)
Let's recall a basic fact about inverses from the one-variable setting. Let \( f : \mathbb{R} \to \mathbb{R} \), \( f'(a) \neq 0 \), \( f' \) exists in an open set containing \( a \). 

\( f \) is increasing in a neighborhood of \( a \), so \( f \) is 1-1 there and hence \( f^{-1} \) exists. It follows that \( f \circ f^{-1}(x) = x \) in a small neighborhood of \( a \), so

\[
(f^{-1})'(x) = \frac{1}{f'(f(x))}
\]

Our goal is a higher-dimensional version of this theory.

Lemma: \( A \subseteq \mathbb{R}^n \) rectangle and let \( f : A \to \mathbb{R}^n \) be continuously differentiable. If there is \( M \in \mathbb{R}^+ \)

\[
\|Df(x)\| \leq M \quad \forall x \in \text{interior of } A
\]

then

\[
\|f(y) - f(x)\| \leq n^2 M |x - y| \quad \forall x, y \in A.
\]

Proof: This has mean value theorem written all over it.

\[
g'(y) - g'(x) = \sum_{i=1}^{n} \left[ g_i'(y_{i-1}, y_{i}, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) - g_i'(y_{i-1}, y_{i}, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \right]
\]

Write it out for \( n = 2 \)

\[
\left| g_1'(y_{-1}, y_1, \ldots, x_{i+1}, \ldots, x_n) - g_1'(y_{-1}, y_1, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \right| = \left| (y_i - x_i) \cdot Dz f'(z_{ij}) \right| \leq M |y_i - x_i|.
\]
Thus, $|f(y) - f(x)| \leq \sum_{j=1}^{n} |y_j - x_j| M \leq n M |y - x|$

Note: $\sum_{j=1}^{n} |x_j| \leq n |x|$ since $|x_j| \leq |x|$.  

But $\sum_{j=1}^{n} |x_j|^2 \leq \left( \sum_{j=1}^{n} |x_j|^2 \right)^{\frac{1}{2}} = \sqrt{n} |x|$

Cauchy-Schwarz

It follows that $|f(y) - f(x)| \leq \sum_{i=1}^{n} |f(y) - f(x)| \leq n^2 M |y - x|$

can replace $n^2$ by $n^2$.

Can we save a bit more?

$f(y) = (f_1(y), \ldots, f_n(y))$, so

$|f(y) - f(x)| = \left( \sum_{i=1}^{n} (f_i(y) - f_i(x))^2 \right)^{\frac{1}{2}}$

$\leq \left( \sum_{i=1}^{n} n M^2 |y_i - x_i|^2 \right)^{\frac{1}{2}} = M |y - x| \cdot n$, so

yes...
Inverse function theorem: \( f: \mathbb{R}^n \rightarrow \mathbb{R}^n \) continuously differentiable in an open set containing \( a \), and \( \det f'(a) \neq 0 \). Then there is an open set \( V \) containing \( a \) and an open set \( W \) containing \( f(a) \) such that \( f: V \rightarrow W \) has a continuous inverse \( f^{-1}: W \rightarrow V \) which is differentiable for all \( y \in W \) satisfies

\[
(f^{-1})'(y) = \left[f'(f^{-1}(y))\right]^{-1}
\]

Proof: Let \( \lambda = Df(a) \). By assumption, \( \lambda \) is non-singular.

\[
D(\lambda^{-1} \circ f)(a) = D(\lambda^{-1})(f(a)) \circ Df(a) = \lambda^{-1} \circ Df(a)
\]

\( \lambda^{-1} \circ Df(a) \) is the identity. By precomposing \( w \) with \( \lambda^{-1} \), we may always assume that \( \lambda = \text{identity} \).

It follows that whenever \( f(a+h) = f(a) \),

\[
\left\| \frac{f(a+h) - f(a) - \lambda(h)}{h} \right\| = \frac{1}{\|h\|} = 1 \quad \text{yet} \quad \frac{1}{\|h\|}
\]

\[
lim_{h \to 0} \frac{f(a+h) - f(a) - \lambda(h)}{h} = 0.
\]

\[
\implies f(\xi) = f(a) \text{ impossible for } \xi \text{ arbitrarily close (but not equal to) to } a.
\]
It follows that if a closed rectangle $U$ containing $a$ in its interior such that

1. $f(x) \neq f(a)$ if $x \in U$ and $x \neq a$.
2. $|f'(x)| \neq 0$ for $x \in U$, $x \neq a$ (continuity).
3. $|f'(x) - f'(a)| < \frac{1}{2n^2}$ for $x \in U$.

$\Rightarrow$ Lipschitz lemma $\Rightarrow$

$$|f(x_1) - f(x_2)| \leq \frac{1}{2} |x_1 - x_2|$$

$\forall$

$$|x_1 - x_2| - |f(x_1) - f(x_2)|$$

$\Rightarrow$ (4) $|x_1 - x_2| \leq 2|f(x_1) - f(x_2)|$

$\forall x, x_2 \in U.$

Note that $f(\partial U)$ is a compact set which does not contain $f(a)$ (by 1.)

$\Rightarrow$ $|f(a) - f(x)| \geq d > 0$ some $d$ for $x \in \partial U.$
Let \( W = \{ y : |y - f(a)| < \frac{1}{2} \} \).

If \( y \in W \) and \( x \in \partial W \), then

5. \( |y - f(a)| < |y - f(x)| \)

To see this, note that

\[
d \leq |f(a) - f(x)| = |f(a) - y + y - f(x)| \\
\leq |f(a) - y| + |y - f(x)|
\]

Since \( |f(a) - y| < \frac{1}{2} \), \( \Rightarrow |y - f(x)| > \frac{1}{2} \)

and we are done.

We will show that for any \( y \in W \) \( \exists ! x \in \text{int} \text{ of } U \)

\( \Rightarrow f(x) = y \).

To see this, define \( U \rightarrow \mathbb{R} : \)

\[ g(x) = |y - f(x)|^2 = \sum_{i=1}^{n} |y_i - f_i(x)|^2 \]

continuous! \( \Rightarrow \) has a minimum on \( U \).
If \( x \in \partial U \), \( g(a) < g(x) \) by \( \circ \).

\( \Rightarrow \) min occurs on the interior of \( U \).

\( \Rightarrow \) \( D_y g(x) = 0 \) \( \forall i \), i.e.

\[ \sum_{i=1}^{n} 2(y^i - f^i(x)) \cdot D_j f^i(x) = 0 \quad \forall j \]

But! \( \{D_j f^i(x)\} \) has a non-zero determinant by \( \circ \).

\( \Rightarrow \) \( y^i - f^i(x) = 0 \) \( \forall i \), i.e.

\[ y = f(x) \]  

proving the existence of \( x \); \( w \)

uniqueness implied by \( \circ \).

If \( V = U \cap f^{-1}(W) \) we have shown that \( f: V \to W \) has an inverse \( f^{-1}: W \to V \).

But what about differentiability?
Rewrite \( \Phi \) in the form

\[
\left| \hat{g}'(y_1) - \hat{g}'(y_2) \right| \leq 2 |y_1 - y_2| \quad \forall y_1, y_2 \in W
\]

\( \Rightarrow \hat{g}' \) is continuous.

Let \( \mu = D \Phi(x) \). We will see in a moment that \( \hat{g}' \) is diff at \( y = \hat{g}(x) \) w/ derivative \( \mu' \).

We have \( \hat{g}(x_1) = \hat{g}(x) + \mu(x_1 - x) + o(x_1 - x) \)

where \( \lim_{x_1 \to x} \frac{o(x_1 - x)}{|x_1 - x|} = 0 \)

\( \Rightarrow \mu'^{-1}(\hat{g}(x_1) - \hat{g}(x)) = x_1 - x + \mu'^{-1}(o(x_1 - x)) \)

(\textbf{Note:} Every \( y \in W = \hat{g}(x_1) \) for some \( x_1 \in V \),

so \( \hat{g}'(y_1) = \hat{g}'(y) + \mu'(y_1 - y) - o(\hat{g}'(y_1) - \hat{g}'(y)) \)

so it suffices to show that

\[
\lim_{y \to y_1} \frac{1}{|y - y_1|} |\mu(\hat{g}'(y_1) - \hat{g}'(y))| = 0.
\]
so, it suffices (why?) to prove that
\[
\lim_{y_1 \to y} \frac{|f(g'(y_1)) - f'(y)|}{|y_1 - y|} = 0
\]

So, \[
\frac{|f(g'(y_1)) - f'(y)|}{|y_1 - y|} = \frac{|f(g(y_1)) - f'(y_1)||g'(y)}{||g'(y_1)| - |g'(y)||}
\]

so we are done!