Proofs of Basic Theorems on Differentiable Functions

1. CHAIN RULE: When \( f : \mathbb{R}^n \mapsto \mathbb{R}^m \) is differentiable at \( a \in \mathbb{R}^n \) and \( g : \mathbb{R}^m \mapsto \mathbb{R}^p \) is differentiable at \( b = f(a) \), then the composite function \( h = g \circ f \) is differentiable at \( a \) with
\[
(dh)_a = (dg)_b \circ (df)_a.
\]

PROOF. By the definition of differentiability, for
\[
E_f(x) = \frac{f(x) - f(a) - (df)_a(x - a)}{||x - a||}
\]
and
\[
E_g(y) = \frac{g(y) - g(b) - (dg)_b(y - b)}{||y - b||},
\]
we have
\[
\lim_{x \to a} E_f(x) = 0 = \lim_{y \to b} E_g(y).
\]
Defining \((dh)_a\) to be \((dg)_b \circ (df)_a\), we need to show that
\[
E_h(x) = \frac{h(x) - h(a) - (dh)_a(x - a)}{||x - a||} \to 0 \text{ as } x \to a.
\]
Because linear transformations on finite dimensional vector spaces are continuous, there are positive constants \( C_f \) and \( C_g \) for which
\[
||(df)_a(x - a)|| \leq C_f ||x - a|| \quad \forall x \quad \text{and}
\]
\[
||(dg)_b(y - b)|| \leq C_g ||y - b|| \quad \forall y,
\]
Since \( f(x) - f(a) = (df)_a(x - a) + ||x - a||E_f(x), \)
we deduce that 
\[ \| f(x) - f(a) \| \leq (C_f + \|E_f(x)\|) \| x - a \| \forall x. \]

Using \( h(x) = g(f(x)) \) and \( h(a) = g(f(a)) = g(b) \), we can use these inequalities and the triangle inequality to obtain

\[ \|E_{h}(x)\| = \|g(f(x)) - g(f(a)) - (dg)_b((f(x)) - f(a)) \]
\[ + (dg)_b((f(x)) - f(a) - (df)_a(x - a))\|/\|x - a\| \]
\[ \leq \|E_{g}(f(x))\| \|f(x) - f(a)\|/\|x - a\| + C_g \|E_f(x)\| \]
\[ \leq \|E_{g}(f(x))\|(C_f + \|E_f(x)\|) + C_g \|E_f(x)\|. \]

Then, as \( x \to a \), \( \|E_{h}(x)\| \to 0 \) since \( \|E_f(x)\| \to 0 \), \( f(x) \to b \)
by continuity of \( f \) at \( a \), and thus \( \|E_{g}(f(x))\| \to 0 \) in view of
the fact that \( \|E_g(y)\| \to 0 \) as \( y \to b \). This completes the proof.

2. GENERALIZATION OF ROLLE'S THEOREM. Let
\( I = (a, b) \) be a possibly infinite interval and suppose
\( f : I \mapsto \mathbb{R} \) is a function which is differentiable on \( I \) and
for which \( \lim_{x \to a} f(x) = 0 = \lim_{x \to b} f(x) \). Then there is at least
one point \( c \in I \) for which \( f'(c) = 0 \).

PROOF. If \( f(x) = 0 \ \forall x \in I \), \( f'(x) = 0 \ \forall x \in I \).
Otherwise, replacing \( f \) by \( -f \) if need be, we can assume
there is a point \( x_1 \) in \( I \) for which \( f(x_1) > 0 \). By the assumptions on \( f \), we can choose \( a_1 \) and \( b_1 \) in \( I \) for which \( a_1 < x_1 < b_1 \) and \( |f(x)| < f(x_1) \) when either \( a < x \leq a_1 \) or \( b_1 \leq x < b \). Then, on the compact set \([a_1, b_1]\), \( f \) achieves a maximum value \( M \) at a point \( c \). Since \( M \geq f(x_1) > \max\{f(a_1), f(b_1)\}, \ c \in (a_1, b_1) \). Then \( f'(c) = 0 \) from the elementary calculus observation that \( f' \) vanishes at any local maximum or minimum point.

NOTE: Aside from the mild extension to possibly infinite intervals, this proof appears in most elementary calculus texts with "handwaving" over the existence of \( M \) since elementary calculus texts don't want to get into sups and infs, much less the properties of continuous functions on compact sets.

3. CAUCHY MEAN VALUE THEOREM. Let \( I \) be as in Rolle's Theorem with \( f(x) \) and \( g(x) \) two \( \mathbb{R} \)-valued differentiable functions on \( I \) having finite limits \( f(a), g(a) \) as \( x \to a \) and \( f(b), g(b) \) as \( x \to b \). Then there exists a point \( c \in I \) for which \( (f(b) - f(a))g'(c) = (g(b) - g(a))f'(c) \).

PROOF. Let
\[
h(x) = (f(b) - f(a))(g(x) - g(a)) \\
- (f(x) - f(a))(g(b) - g(a)).
\]
Then \( h \) satisfies the hypotheses of Rolle's Theorem so there is a point \( c \in I \) for which
\[
0 = h'(c) = (f(b) - f(a))g'(c) - (g(b) - g(a))f'(c).
\]
4. MEAN VALUE THEOREM. Let \([a, b]\) be a closed, bounded interval and \(f:[a, b] \mapsto \mathbb{R}\) a function which is differentiable on the open interval \((a, b)\) and continuous at both \(a\) and \(b\). Then \(f(b) - f(a) = (b - a)f'(c)\) for some \(c \in (a, b)\).

PROOF. Apply the Cauchy Mean Value Theorem with \(g(x) = x\), hence \(g'(c) = 1\).