Proof of the Chain Rule

This is a proof of the Chain Rule that is based on a Lemma that should be thought of as as base case of Taylor's Theorem. This proof has the advantage of generalising easily to many variables which will be needed in Analysis II.

Lemma: Let f be a function defined on a neighbourhood of p. If f is differentiable at p then there is a function $E(h)$ defined on a neighbourhood of 0 such that

$$f(p+h) = f(p) + f'(p)h + hE(h) \quad \text{and} \quad \lim_{h\to 0} E(h) = 0. \tag{0.1}$$

Conversely suppose $E(h)$ is a scalar function defined on an neighbourhood of zero and α be a real number such that

$$f(p+h) = f(p) + \alpha h + hE(h) \quad \text{and} \quad \lim_{h\to 0} E(h) = 0. \tag{0.2}$$

Then f is differentiable at p and $f'(p) = \alpha$.

Proof: Suppose f is differentiable and set

$$E(h) = \begin{cases} \frac{f(p+h)-f(p)}{h} - f'(p) & \text{if } h \neq 0 \\ 0 & \text{if } h = 0 \end{cases}$$

Clearly from the definition of the derivative as a limit it is clear that (0.1) holds. Conversly if (0.2) holds then

$$\lim_{h\to 0} \frac{f(p+h)-f(p)}{h} = \lim_{h\to 0} (\alpha + E(h)) = \alpha$$

so f is differentiable at p and $f'(p) = \alpha$. ∎

Theorem (Chain Rule): Suppose f be a real valued function defined in a neighbourhood of p and suppose f is differentiable at p. Suppose also that g is a real valued function defined in a neighbourhood of $f(p)$ and that g is differentiable at $f(p)$. Then the composition $g \circ f$ is differentiable at p and $(g \circ f)'(p) = g'(f(p))f'(p)$.

Proof: Set $q = f(p)$. Since f is differentiable at p and g is differentiable at q the previous lemma says there are functions E_1 and E_2 defined on a neighbourhood of zero such that

$$f(p+h) = f(p) + f'(p)h + hE_1(h) \quad \text{and} \quad \lim_{h\to 0} E_1(h) = 0. \tag{0.3}$$

$$g(q+k) = g(q) + g'(q)k + kE_2(k) \quad \text{and} \quad \lim_{k\to 0} E_2(k) = 0. \tag{0.4}$$

We shall think of k as a function of h given by $k = f(p+h) - f(p)$. Then

$$\begin{aligned} g(f(p+h)) &= g(f(p)+k) = g(q+k) \\ &= g(q) + g'(q)k + kE_2(k) \\ &= g(f(p)) + hg'(f(p))\frac{k}{h} + h\frac{k}{h}E_2(k) \\ &= g(f(p)) + hg'(f(p))f'(p) + h\left[g'(f(p))\left(\frac{k}{h} - f'(p)\right) + \frac{k}{h}E_2(k)\right] \\ &= g(f(p)) + hg'(f(p))f'(p) + hE_3(h) \end{aligned}$$

where

$$E_3(h) = g'(f(p))\left(\frac{k}{h} - f'(p)\right) + \frac{k}{h}E_2(k).$$

Now

$$\lim_{h\to 0} \frac{k}{h} = \lim_{h\to 0} \frac{f(p+h)-f(p)}{h} = f'(p).$$

Also notice that as f is differentiable at p it is continuous at p so $\lim_{h\to 0} k = \lim_{h\to 0} f(p+h) - f(p) = 0$. Thus $\lim_{h\to 0} E_2(k) = 0$ and so putting these all together we get $\lim_{h\to 0} E_3(h) = 0$. Hence the lemma gives $(g \circ f)$ is differentiable at p with derivative $g'(f(p))f'(p)$. ∎