Mathematics Department Stanford University
Math 51H – Product rule

The product rule for real valued functions is a good exercise in differentiability. The proof given below is not streamlined (to minimize the number of estimates); part of the point is to show that one need not have a clever idea in order to get the conclusion.

Theorem 1 *Suppose $U \subset \mathbb{R}^n$ is open, $a \in U$, $f, g : U \to \mathbb{R}$ differentiable at a. Then fg is differentiable at a, with $D(fg)(a) = g(a)(Df)(a) + f(a)(Dg)(a)$.*

Note that here $Df(a), Dg(a), D(fg)(a) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$.

Proof: Suppose $B_\rho(a) \subset U$, $\rho > 0$; below all δs will be taken $< \rho$ even if not explicitly stated.

By the definition of the differentiability of f at a, for any $\varepsilon_f > 0$ there is $\delta_f > 0$ such that

$$\|h\| < \delta_f \Rightarrow f(a+h) = f(a) + (Df)(a)h + R_f(a,h),\ |R_f(a,h)| \le \varepsilon_f \|h\|.$$

By the definition of the differentiability of g at a, for any $\varepsilon_g > 0$ there is $\delta_g > 0$ such that

$$\|h\| < \delta_g \Rightarrow g(a+h) = g(a) + (Dg)(a)h + R_g(a,h),\ |R_g(a,h)| \le \varepsilon_g \|h\|.$$

Now,

$$\begin{aligned}(fg)(a+h) &= f(a+h)g(a+h) = (f(a) + (Df)(a)h + R_f(a,h))(g(a) + (Dg)(a)h + R_g(a,h)) \\ &= f(a)g(a) + (f(a)(Dg)(a) + g(a)(Df)(a))h + R_{fg}(a,h),\end{aligned}$$

where

$$\begin{aligned}R_{fg}(a,h) =& f(a)R_g(a,h) + g(a)R_f(a,h) + ((Df)(a)h)((Dg)(a)h) \\ &+ ((Df)(a)h)R_g(a,h) + R_f(a,h)((Dg)(a)h) + R_f(a,h)R_g(a,h).\end{aligned}$$

The theorem thus follows if we show that for all $\varepsilon > 0$ there exists $\delta > 0$ such that $\|h\| < \delta$ implies $|R_{fg}(a,h)| \le \varepsilon \|h\|$. To show this, let's make each of the six terms $\le \frac{\varepsilon}{6}\|h\|$.

First, let $\delta_1 > 0$ be δ_g in the definition of differentiability of g corresponding to $\varepsilon_g = \frac{\varepsilon}{6(|f(a)|+1)} > 0$. Then for $\|h\| < \delta_1$,

$$|f(a)R_g(a,h)| \le |f(a)| \frac{\varepsilon}{6(|f(a)|+1)} \|h\| \le \frac{\varepsilon}{6}\|h\|,$$

as desired. Similarly, let $\delta_2 > 0$ be $\delta_f > 0$ in the definition of the differentiability of f corresponding to $\varepsilon_f = \frac{\varepsilon}{6(|g(a)|+1)} > 0$; then $\|h\| < \delta_2$ gives

$$|g(a)R_g(a,h)| \le \frac{\varepsilon}{6}\|h\|.$$

Next,

$$|((Df)(a)h)((Dg)(a)h)| = |(Df)(a)h|\,|((Dg)(a)h| \le \|(Df)(a)\|\|h\|\,\|(Dg)(a)\|\|h\|,$$

so with $\delta_3 = \frac{\varepsilon}{6\|(Df)(a)\|\|(Dg)(a)\|+1}$, $\|h\| < \delta_3$ implies

$$\begin{aligned}|((Df)(a)h)((Dg)(a)h)| &\le \|(Df)(a)\|\,\|(Dg)(a)\|\|h\|\,\|h\| \\ &\le \|(Df)(a)\|\,\|(Dg)(a)\| \frac{\varepsilon}{6\|(Df)(a)\|\|(Dg)(a)\|+1} \|h\| \le \frac{\varepsilon}{6}\|h\|.\end{aligned}$$

For the fourth term, $((Df)(a)h)R_g(a,h)$, we have

$$|((Df)(a)h)R_g(a,h)| \le \|Df(a)\|\|h\|\,|R_g(a,h)|,$$

so this will be bounded by a small multiple of $\|h\|$ if we just make $R_g(a,h)$ small: so let δ_g be given by the differentiability of g at a with $\varepsilon_g = 1$, and and let $\delta_4 = \min(\delta_g, \frac{\varepsilon}{6(\|Df(a)\|+1)})$, so for $\|h\| < \delta_4$, and thus in particular $\|h\| < \delta_g$, $|R_g(a,h)| \le \|h\|$, and

$$|(Df)(a)h)R_g(a,h)| \le \|Df(a)\|\|h\|\,\|h\| \le \|Df(a)\| \frac{\varepsilon}{6(\|Df(a)\|+1)} \|h\| \le \frac{\varepsilon}{6}\|h\|,$$

as desired. Note that we had quite a bit we could give up here: we only needed to use $\varepsilon_g = 1$ (and not some small quantity depending on ε) in the definition of the differentiability of g. Reversing the roles of f and g gives $\delta_5 > 0$ such that $\|h\| < \delta_5$ implies

$$|(Dg)(a)h)R_f(a,h)| \le \frac{\varepsilon}{6}\|h\|.$$

Finally, taking $\varepsilon_f = \varepsilon_g = 1$ in the definition of differentiability of f, g gives $\delta_f, \delta_g > 0$ such that $\|h\| < \min(\delta_f, \delta_g)$ gives $|R_f(a,h)| \le \|h\|$, $R_g(a,h)| \le \|h\|$, and thus

$$|R_f(a,h)R_g(a,h)| \le \|h\|^2.$$

Letting finally $\delta_6 = \min(\delta_f, \delta_g, \frac{\varepsilon}{6})$ gives that for $\|h\| < \delta_6$,

$$|R_f(a,h)R_g(a,h)| \le \|h\|^2 \le \frac{\varepsilon}{6}\|h\|.$$

In combination, with $\delta = \min\{\delta_j : \ j = 1, \dots, 6\}$,

$$\|h\| < \delta \Rightarrow |R_{fg}(a,h)| \le \varepsilon\|h\|,$$

completing the proof. $\square$

A somewhat more streamlined version would use the standard rewriting

$$f(a+h)g(a+h) - f(a)g(a) = f(a+h)(g(a+h) - g(a)) + (f(a+h) - f(a))g(a),$$

and use the definition of differentiability of g, resp. f, in the two terms, as well as $\lim_{h\to 0} f(a+h) = f(a)$ since differentiability implies continuity.