Motivation for Chain Rule

If $f(x) = ax$ and $g(x) = bx$ then $f(g(x)) = abx$.

So $D(f(g)) = ab$ is the product of

$Df = a$ and $Dg = b$.

Proof of Chain Rule

Suppose f is differentiable at $g(x)$ and
g is differentiable at x.
Since g is differentiable, and also applying f, there is a number $Dg(x)$ with

$$f(g(x+h)) = f(g(x) + Dg(x)h + R_g h)$$

Now write $u = g(x)$ and $l = Dg(x)h + R_g h$ to get

$$f(g(x+h)) = f(u+l)$$

Note $l \to 0$ as $h \to 0$.

Proof concluded

Since f is differentiable there is a number $Df(u)$ with

$$f(g(x+h)) = f(u) + Df(u)l + R_f l$$

Expand $u = g(x)$ and $l = Dg(x)h + R_g h$ to get:

$$f(g(x+h)) = f(g(x)) + Df(g(x))Dg(x)h +$$

$$Df(g(x))R_g h + R_f Dg(x)h + R_f R_g h$$

Ignoring the h's, the second line $\longrightarrow 0$ as $h \longrightarrow 0$ so $Df(g(x))Dg(x)$ is the derivative of $f(g(x))$.