Chapter 2

Rules of differentiation

There are many useful **rules** of differentiation. Let us investigate how the fundamental definitions presented above can lead to these rules.

2.1 Chain rule

It is entirely possible that the argument of the function we are considering is itself a function of the dependent variable, i.e. $f(x) = F[g(x)]$ for some functions F and g. For example, what is the derivative with respect to x of $f(x) = \sin(x^2 - x + 2)$? Here $F(X) = \sin(X)$, and $g(x) = x^2 - x + 2$. Let us apply what we have learnt.

Note that we want to work out the derivative of f with respect to x. From (1.1),

$$\frac{df}{dx} = \lim_{h \to 0} \frac{F[g(x+h)] - F[g(x)]}{h},$$

and then using (1.4), the right-hand side can be written as

$$\frac{df}{dx} = \lim_{h \to 0} \frac{F\left[g(x) + h\frac{dg}{dx} + o(h)\right] - F[g(x)]}{h}, \tag{2.1}$$

where we understand that dg/dx is evaluated at x, and we assume that g is differentiable. We also remember that $o(h)$ represents a term $T_h(h)$ such that, by definition

$$\lim_{h \to 0} \frac{T_h(h)}{h} = 0.$$

Now, let us relabel

$$g(x) + h\frac{dg}{dx} + o(h) = X + H; \ \ X = g(x), \ \ H = h\frac{dg}{dx} + o(h).$$

Then, using (1.4),

$$F(X+H) = F(X) + H\frac{dF}{dX} + o(H), \tag{2.2}$$

where here now $o(H)$ represents a term $T_H(H)$ such that, by definition

$$\lim_{H\to 0} \frac{T_H(H)}{H} = 0.$$

Of course, by the definition of H, T_H may also be thought of as a function of h. Now let us consider this term $T_H(H)$ which is $o(H)$. Importantly, in this particular context, since it appears in (2.1), we are only interested in how this term behaves as $h \to 0$. There are two different cases to consider.

1. If $dg/dx = 0$, then H is also a term which is $o(h)$ as $h \to 0$, and so clearly a term that is $o(H)$ as $H \to 0$ (i.e. a term that goes to zero "faster" than H as H goes to zero) is also $o(h)$ as $h \to 0$, (i.e. goes to zero "faster" than h as $h \to 0$) since H itself goes to zero "faster" than h as $h \to 0$.

2. If dg/dx is not equal to zero, we can make two further observations.

   - H is a term which is $O(h)$, but is most definitely not a term which is $o(h)$ as $h \to 0$.
   - Therefore, in this particular context, H is effectively linearly proportional to h.

   Therefore, here, if a term is $o(H)$ as $H \to 0$, it must also be $o(h)$ as $h \to 0$, which is what we require.

Armed with this result, (2.2) becomes

$$\begin{aligned} F(X+H) &= F(X) + \left[h\frac{dg}{dx} + o(h)\right]\frac{dF}{dX} + o(H), \\ &= F[g(x)] + \left[h\frac{dg}{dx}\right]\frac{dF}{dg} + o(h), \end{aligned} \tag{2.3}$$

as $h \to 0$. Here, we write the derivative of F with respect to its argument $X = g(x)$ as dF/dg. We could have written it as $F'[g(x)]$ if we'd rather follow Lagrange of course. Also, we've exploited the property of the order parameter o that any constant multiple of a term which is $o(h)$ is also $o(h)$ by definition.

If we now substitute (2.3) for the first term in the numerator back into (2.1), we obtain

$$\begin{aligned}\frac{df}{dx} &= \lim_{h\to 0}\frac{F\left[g(x)\right]+h\frac{dg}{dx}\frac{dF}{dg}+o(h)-F[g(x)]}{h},\\ &= \lim_{h\to 0}\frac{dF}{dg}\frac{dg}{dx}+\frac{o(h)}{h}.\end{aligned}$$

Finally, taking the limit, we have the **chain rule**

$$\frac{df}{dx}=\frac{dF}{dg}\frac{dg}{dx}=\frac{df}{dg}\frac{dg}{dx}, \tag{2.4}$$

where we have made a final notational observation that $F(g)$ and $f(g)$ are the same function by definition. For our specific example (of course):

$$\frac{d}{dx}\left[\sin(x^2-x+2)\right]=\left[\cos(x^2-x+2)\right](2x-1).$$

2.2 Product rule

Consider the situation where $f(x)=u(x)v(x)$, i.e. f can be written as the product of two other functions u and v. Then, as you are asked to establish on the example sheet from the definition of a derivative:

$$\frac{df}{dx}=v\frac{du}{dx}+u\frac{dv}{dx}. \tag{2.5}$$

Indeed, the so-called quotient rule is just a special case of the product rule:

$$f=\frac{u}{v}\to f'=\frac{u'v-v'u}{v^2}.$$

2.3 (General) Leibniz rule

The product rule can be generalized to higher order derivatives very straightforwardly by recursive application.

$$\begin{aligned}f &= uv,\\ f' &= u'v+uv',\\ f'' &= u''v+u'v'+u'v'+uv'',\\ &= u''v+2u'v'+uv'',\\ f''' &= u'''v+u''v'+2u''v'+2u'v''+u'v''+uv''',\\ &= u'''v+3u''v'+3u'v''+uv'''.\end{aligned}$$

Hopefully, this is reminiscent of Pascal's triangle and the binomial theorem. The general Leibniz rule is:

$$
\begin{aligned}
f^{(n)}(x) &= u^{(n)}v + nu^{(n-1)}v' + \frac{n(n-1)}{2!}u^{(n-2)}v'' + \ldots \\
&\quad \cdots + \begin{pmatrix} n \\ r \end{pmatrix} u^{(n-r)}v^{(r)} + \ldots uv^{(n)},
\end{aligned}
$$

where a superscript (n) denotes the n^{th} derivative, and the binomial coefficient

$$
\begin{pmatrix} n \\ r \end{pmatrix}
$$

denotes the number of combinations of r elements that can be taken from n elements without replacement. Of course, this "rule" relies on the function f and its first $n-1$ derivatives all being differentiable.

2.4 Taylor Series

2.4.1 Taylor's Theorem

Remember (1.4), which can be rewritten as

$$
f(x_0 + h) = f(x_0) + h \left. \frac{df}{dx} \right|_{x=x_0} + o(h).
$$

Provided the first n derivatives of f exist, this can be generalized to

$$
\begin{aligned}
f(x_0 + h) &= f(x_0) + h \left. \frac{df}{dx} \right|_{x=x_0} + \frac{h^2}{2!} \left. \frac{d^2 f}{dx^2} \right|_{x=x_0} + \ldots \\
&\quad \cdots + \frac{h^n}{n!} \left. \frac{d^n f}{dx^n} \right|_{x=x_0} + E_n, \qquad (2.6)
\end{aligned}
$$

where $E_n = o(h^n)$ as $h \to 0$. In fact, if $f^{(n+1)}$ exists $\forall x \in (x_0, x_0 + h)$ and f^n is continuous on $[x_0, x_0 + h]$, then for some $x_0 < x_n < x_0 + h$,

$$
E_n = \frac{f^{(n+1)}(x_n)}{(n+1)!} h^{n+1},
$$

and so $E_n = O(h^{n+1})$ as $h \to 0$. This is a statement of **Taylor's Theorem**. Note that it is an exact statement which expresses the value of a function f at a point $x_0 + h$ in terms of the value of the function at x_0, its derivatives at x_0, and an error term E_n whose behaviour we know as h gets smaller.

2.4.2 Taylor Polynomials

Assume that $x = x_0 + h$. Then, (2.6) can be rewritten as

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + \cdots + \frac{(x - x_0)^n}{n!}f^{(n)}(x_0) + E_n.$$

The first n terms on the right hand side are the first n terms of the **Taylor series**, or alternatively the nth order Taylor polynomial of $f(x)$ about the point $x = x_0$. It gives a **local** approximation to the function, and interestingly, in a sense, it is the best possible local approximation to the function.

2.4.3 Determination of coefficients

The coefficients of the Taylor series are determined from evaluations of higher and higher order derivatives of the underlying function. Without loss of generality (wlog) consider an expansion of $f(x)$ about $x = 0$. Then the "obvious" series representation of a sufficiently smooth function (i.e. one which has sufficiently many higher derivatives) $f(x)$ is

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n + \ldots$$

It is hopefully straightforward to establish that:

$$f(0) = a_0,\ f'(0) = a_1,\ f''(0) = 2a_2,\ \ldots\ f^{(n)}(0) = n!a_n.$$

2.5 L'Hôpital's Rule

Taylor series representations are very useful to understand L'Hôpital's rule, which can be used to determine the value of **indeterminate forms** such as ∞/∞ and $0/0$. The simplest version of L'Hôpital's rule is for the case where $f(x)$ and $g(x)$ are both differentiable at $x = x_0$ and

$$\lim_{x \to x_0} f(x) = \lim_{x \to x_0} g(x) = 0.$$

Then if $g'(x_0) \neq 0$,

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \lim_{x \to x_0} \frac{f'(x)}{g'(x)}, \tag{2.7}$$

provided the limit exists (as is often the case, there can be quite scary pathological cases beloved of clear-thinking analysts...)

A proof for this special case (the rule applies in much more general circumstances) is quite short. From the Taylor series representations:

$$\begin{aligned} f(x) &= f(x_0) + (x - x_0)f'(x_0) + o(x - x_0); \\ &= 0 + (x - x_0)f'(x_0) + o(x - x_0); \\ g(x) &= g(x_0) + (x - x_0)g'(x_0) + o(x - x_0); \\ &= 0 + (x - x_0)g'(x_0) + o(x - x_0); \\ \rightarrow \frac{f}{g} &= \frac{f' + \frac{o(x-x_0)}{(x-x_0)}}{g' + \frac{o(x-x_0)}{(x-x_0)}}. \end{aligned} \tag{2.8}$$

Therefore, taking the limit:

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \lim_{x\to x_0} \frac{f' + \frac{o(x-x_0)}{(x-x_0)}}{g' + \frac{o(x-x_0)}{(x-x_0)}} = \frac{f'(x_0)}{g'(x_0)}. \tag{2.9}$$

Interestingly, the rule can be generalised to higher orders. For example, consider

$$f(x) = 3\sin(x) - \sin(3x); \;\; g(x) = 2x - \sin(2x).$$

For these functions, $f(0) = g(0) = f'(0) = g'(0) = f''(0) = g''(0) = 0$. As an exercise, show that

$$\lim_{x\to 0} \frac{f(x)}{g(x)} = 3 = \lim_{x\to 0} \frac{f'''(x)}{g'''(x)}.$$