Notes on Calculus

by

Dinakar Ramakrishnan
253-37 Caltech
Pasadena, CA 91125
Fall 2001

Contents

0 Logical Background **2**
0.1 Sets 2
0.2 Functions 3
0.3 Cardinality 3
0.4 Equivalence Relations 4

1 Real and Complex Numbers **6**
1.1 Desired Properties 6
1.2 Natural Numbers, Well Ordering, and Induction 8
1.3 Integers 10
1.4 Rational Numbers 11
1.5 Ordered Fields 13
1.6 Real Numbers 14
1.7 Absolute Value 18
1.8 Complex Numbers 19

2 Sequences and Series **22**
2.1 Convergence of sequences 22
2.2 Cauchy's criterion 26
2.3 Construction of Real Numbers revisited 27
2.4 Infinite series 29
2.5 Tests for Convergence 31
2.6 Alternating series 33

3 Basics of Integration **36**
3.1 Open, closed and compact sets in $\mathbb{R}$ 36
3.2 Integrals of bounded functions 39
3.3 Integrability of monotone functions 42
3.4 Computation of $\int_a^b x^s dx$ 43
3.5 Example of a non-integrable, bounded function 45
3.6 Properties of integrals 46
3.7 The integral of x^m revisited, and polynomials 48

4 Continuous functions, Integrability **51**
4.1 Limits and Continuity 51
4.2 Some theorems on continuous functions 55
4.3 Integrability of continuous functions 57
4.4 Trigonometric functions 58
4.5 Functions with discontinuities 62

5 Improper Integrals, Areas, Polar Coordinates, Volumes **64**
5.1 Improper Integrals 64
5.2 Areas 67
5.3 Polar coordinates 69
5.4 Volumes 71
5.5 The integral test for infinite series 73

6 Differentiation, Properties, Tangents, Extrema **76**
6.1 Derivatives 76
6.2 Rules of differentiation, consequences 79
6.3 Proofs of the rules 82
6.4 Tangents 84
6.5 Extrema of differentiable functions 85
6.6 The mean value theorem 86

7 The Fundamental Theorems of Calculus, Methods of Integration **89**
7.1 The fundamental theorems 89
7.2 The indefinite integral 92
7.3 Integration by substitution 92
7.4 Integration by parts 95

6 Differentiation, Properties, Tangents, Extrema

6.1 Derivatives

Let a be a real number and f a function defined on an interval around a. One says that f is **differentiable at** a iff the following limit exists:

$$L := \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}.$$

When the limit exists, we will set

$$f'(a) = L.$$

We will also denote it sometimes by $\frac{df}{dx}(a)$.

Consider, for example, the case of a **linear function**

$$f(x) = mx + c,$$

whose graph is the **line of slope** m, passing through the point $(0, c)$. Since $f(x+h) = m(x+h) + c = f(x) + mh$, we have at any point a in $\mathbb{R}$,

$$\frac{f(x+h) - f(x)}{h} = \frac{mh}{h} = m,$$

which is independent of h. Hence it has the limit m as h approaches 0. Thus f is differentiable at $x = a$ with $f'(a)$ being the slope m.

In particular, when $m = 0$, the function f is just the **constant function** $x \mapsto c$, and the derivative is 0.

The next simple example to consider is the **quadratic function**

$$f(x) = \alpha x^2 + \beta x + \gamma,$$

with α, β, γ in $\mathbb{R}$. Then

$$\frac{f(x+h) - f(x)}{h} = \frac{(2\alpha x + \beta)h + \alpha h^2}{h} = 2\alpha + \beta + \alpha h,$$

which has the limit $2\alpha x + \beta$ as h tends to 0. Thus

$$\frac{df}{dx} = 2\alpha x + \beta.$$

This is L:eibnitz's notation, and it means that for any a,

$$f'(a) = \frac{df}{dx}(a) = 2\alpha a + \beta.$$

In particular, the **squaring function** $f(x) = x^2$ is differentiable everywhere with derivative $2x$. This is a superspecial case of the following important result on the **power function** x^t.

Proposition 6.1.1 *For any real number t, consider the function*

$$f(x) \;=\; x^t.$$

Then f is differentiable at any point a in $\mathbb{R}$, with

$$f'(a) \;=\; ta^{t-1}.$$

We will prove this in four stages. The first step is the following:

Proof when t is a positive integer m. Let $a, h \in \mathbb{R}$. By the binomial theorem,

$$(a+h)^m \;=\; \sum_{j=0}^{m} a^{m-j}h^j \;=\; a^m + ma^{m-1}h + \frac{m(m-1)}{2}a^{m-2}h^2 + \ldots h^m.$$

In other words,

$$\frac{(a+h)^m - a^m}{h} \;=\; ma^{m-1} + hg(a,h),$$

where

$$g(a,h) \;=\; \sum_{j=2}^{m} a^{m-j}h^{j-2} \;=\; \frac{m(m-1)}{2}a^{m-2} + \frac{m(m-1)(m-2)}{2}a^{m-3}h + \ldots + h^{m-2}.$$

Then $g(x,0) = \frac{m(m-1)}{2}a^{m-2}$, and so $hg(x,h)$ goes to 0 as h goes to 0. Consequently,

$$\lim_{h\to 0}\frac{(a+h)^m - a^m}{h} \;=\; ma^{m-1},$$

as asserted.

$\square$

Here is the next step.

Proof when $t = 1/n$ with n a positive integer. For any $a \in \mathbb{R}$, note that

$$h \;=\; z^n - w^n \quad \text{if} \quad z = (a+h)^{1/n},\; w = a^{1/n}.$$

Thus

$$\frac{(a+h)^{1/n} - a^{1/n}}{h} \;=\; \frac{z-w}{z^n - w^n}.$$

But we have the factorization

$$z^n - w^n \;=\; (z-w)(z^{n-1} + z^{n-2}w + z^{n-3}w^2 + \ldots + zw^{n-2} + w^{n-1}),$$

which can be verified by direct computation. Thus

$$\lim_{h\to 0}\frac{(a+h)^{1/n}-a^{1/n}}{h} \;=\; \lim_{h\to 0}\frac{z-w}{z^n-w^n} \;=\; \lim_{h\to 0}\frac{1}{z^{n-1}+z^{n-2}w+\ldots+zw^{n-2}+w^{n-1}}.$$

Since the function $x\mapsto x^{1/n}$ is continuous, the limit of z^iw^j as h goes to 0 is simply w^{i+j}. We then obtain

$$\lim_{h\to 0}\frac{(a+h)^{1/n}-a^{1/n}}{h} \;=\; \frac{1}{nw^{n-1}} \;=\; \frac{1}{n}a^{\frac{1}{n}-1}.$$

which proves the assertion of Proposition 6.1.1 for $t=1/n$.

□

We will come back to this proposition in the next section after deriving the *product rule*.

Proposition 6.1.2 *The functions $f(x)=\sin x$ and $g(x)=\cos x$ are differentiable at any point a in $\mathbb{R}$, and*

$$f'(a) \;=\; \cos a \quad \textit{and} \quad g'(a) \;=\; -\sin a.$$

Proof. Recall the addition formula

$$\sin(x+y) \;=\; \sin x\cos y+\cos x\sin y.$$

Therefore

$$\frac{\sin(a+h)-\sin a}{h} \;=\; \sin a\frac{\cos h-1}{h}+\cos a\frac{\sin h}{h}.$$

We have already seen that

$$\lim_{h\to 0}\frac{\sin h}{h} \;=\; 1 \quad \text{and} \quad \lim_{h\to 0}\frac{\cos h-1}{h} \;=\; 0.$$

Consequently,

$$\lim_{h\to 0}\frac{\sin(a+h)-\sin a}{h} \;=\; \cos a,$$

as asserted.

Similarly, the addition formula for the cosine function, namely

$$\cos(x+y) \;=\; \cos x\cos y-\sin x\sin y,$$

implies

$$\frac{\cos(a+h)-\cos a}{h} \;=\; \cos a\frac{\cos h-1}{h}-\sin a\frac{\sin h}{h}.$$

The expression on the right has the limit $-\sin a$ as h goes to 0. Done.

□

One can check that if f is differentiable at a, f must necessarily be continuous at a, but it is not sufficient. Indeed, the function $f(x)=|x|$ is continuous everywhere, but it is not integrable at $x=0$. This is because for $h>0$,

$$\frac{f(h)-f(0)}{h} \;=\; \frac{h-0}{h} \;=\; 1,$$

while for $h < 0$,

$$\frac{f(h) - f(0)}{h} \;=\; \frac{-h - 0}{h} \;=\; -1.$$

So $\frac{f(h)-f(0)}{h}$ is 1, resp. -1, as h approaches 0 from the right, resp. left. So there is no unique limit and f is not differentiable at 0. Note, however, that f is differentiable everywhere else.

Let us end this section by looking at a couple of more examples. The first one below is **not differentiable at** $x = 0$:

$$f(x) \;=\; \begin{cases} x \sin\left(\frac{1}{x}\right), & \text{if} \quad x \neq 0 \\ 0, \quad \text{if} \quad x = 0 \end{cases}$$

Indeed, for any $h \neq 0$,

$$\frac{f(h) - f(0)}{h} \;=\; \frac{h \sin\left(\frac{1}{h}\right) - 0}{h} \;=\; \sin\left(\frac{1}{h}\right),$$

and $\sin\left(\frac{1}{h}\right)$ has no limit as h goes to 0; it fluctuates wildly between -1 and 1.

The second example is

$$g(x) \;=\; \begin{cases} x^2 \sin\left(\frac{1}{x}\right), & \text{if} \quad x \neq 0 \\ 0, \quad \text{if} \quad x = 0 \end{cases}$$

In this case

$$\frac{g(h) - g(0)}{h} \;=\; h \sin\left(\frac{1}{h}\right),$$

which approaches 0 as h goes to 0. So $g(x)$ is differentiable at $x = 0$ with

$$g'(0) \;=\; 0.$$

6.2 Rules of differentiation, consequences

The two basic results here are the following:

Theorem 6.2.1 *Let f, g be differentiable functions at some $a \in \mathbb{R}$. Then we have:*

*(i) (**Linearity**) For all α, β in $\mathbb{R}$, the function $\alpha f + \beta g$ is differentiable at a, with*

$$(\alpha f + \beta g)'(a) \;=\; \alpha f'(a) + \beta g'(a).$$

*(ii) (**Product rule**) The product function fg is differentiable at a, with*

$$(fg)'(a) \;=\; f'(a)g(a) + f(a)g'(a).$$

(iii) (**Quotient rule**) *If $g(x)$ is non-zero in an interval around a, then the ratio f/g is differentiable at a, with*

$$\left(\frac{f}{g}\right)(a) \;=\; \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}.$$

Theorem 6.2.2 (**Chain rule**) *Let f be a differentiable function at some a in $\mathbb{R}$, g a differentiable function at $f(a)$. Then the composite function $g \circ f$ is differentiable at a, with*

$$(g \circ f)'(a) \;=\; g'(f(a)) \cdot f'(a).$$

Before giving proofs of these assertions, which will be done in the next section, let us note that as a **consequence**, **rational functions**, **trigonometric functions**, and **various combinations** of them are differentiable wherever they are defined.For example, the function

$$f(x) \;=\; \frac{x^{72} - 21x^3 + \sin^9(x^4 + 43x) - \cos^3(x^3 - 5)}{12 \sin\sqrt{x} - 29x^4 - 4}$$

is differentiable at any x where the denominator is non-zero. A simpler, and more **commonly occurring example** is

$$g(x) \;=\; \tan x.$$

Since $\tan x = \sin x / \cos x$, we get by applying the *quotient rule* and Proposition 6.1.2,

$$\text{(6.2.3}a\text{)} \qquad \frac{d}{dx}\tan x \;=\; \frac{(\cos x)(\cos x) - (\sin x)(-\sin x)}{\cos^2 x} \;=\; \frac{1}{\cos^2 x} \;=\; \sec^2 x.$$

Here we have used the fact that $\sin^2 x + \cos^2 x = 1$, and the formula is valid at any x where $\cos x$ is non-zero, i.e., at any real number x not equal to an odd integer multiple of $\pi/2$. Similarly we have

$$\text{(6.2.3}b\text{)} \qquad \frac{d}{dx}\cot x \;=\; -{}^2 x,$$

$$\text{(6.2.3}c\text{)} \qquad \frac{d}{dx}\sec x \;=\; \sec x \tan x,$$

and

$$\text{(6.2.3}d\text{)} \qquad \frac{d}{dx} x \;=\; -x \cot x,$$

for all x where the function is defined.

Now we will complete the **proof of Proposition 6.1.1**.

In the previous section we proved the formula for the functions $f(x) = x^m$ and $g(x) = x^{1/n}$ for integers m, n with $m \geq 0$, $n > 0$. The next step is to look at the function

$$h(x) \;=\; x^{m/n},$$

for a rational number m/n, with $m, n > 0$. Then f is the composite of the functions f, g above, which are differentiable at every point in their respective domains. Thus for any $a > 0$, we have by the *chain rule*,

$$h'(a) \;=\; g'(f(a)) \cdot f'(a) \;=\; \left(\frac{1}{n}(f(a))^{1/n-1}\right) \cdot \left(ma^{m-1}\right),$$

which simplifies to

$$\frac{m}{n} a^{\frac{m}{n}-m} a^{+m-1} \;=\; \frac{m}{n} a^{\frac{m}{n}-1},$$

as desired. Now suppose

$$\phi(x) \;=\; x^{-m/n},$$

for some positive integers m, n. Then $\phi(x)$ is the reciprocal of the function $h(x) = x^{m/n}$ which we just looked at. By applying the *quotient rule*, we see that h is differentiable at any positive real number a and moreover,

$$\phi'(a) \;=\; \frac{-h'(a)}{h(a)^2} \;=\; -\frac{m}{n} \frac{a^{\frac{m}{n}-1}}{a^{\frac{2m}{n}}} \;=\; -\frac{m}{n} a^{-\frac{m}{n}-1},$$

as asserted.

The final step is to consider the function

$$f(x) \;=\; x^t,$$

for any *real* number t. But t is, by the construction of $\mathbb{R}$, the limit of a sequence $\{t_n\}$ of rational numbers, and for all $x \geq 0$ we have

$$x^t \;=\; \lim_{n\to\infty} x^{t_n}. \tag{6.2.4}$$

(This could be taken as the definition of x^t for real t.) We then have, for all $a > 0$,

$$\lim_{h\to 0} \frac{(a+h)^t - a^t}{h} \;=\; \lim_{h\to 0} \lim_{n\to\infty} \frac{(a+h)^{t_n} - a^{t_n}}{h}. \tag{6.2.5}$$

It can be shown that the order of taking the two limits can be reversed (see the later chapter of *infinite series of functions*). Thus the right hand side of (6.2.5) becomes

$$\lim_{n\to\infty} \lim_{h\to 0} \frac{(a+h)^{t_n} - a^{t_n}}{h} \;=\; \lim_{n\to\infty} \frac{d\{x^{t_n}\}}{dx}(a),$$

which simplifies to

$$\lim_{n\to\infty} t_n a^{t_n - 1} \;=\; ta^{t-1}, \tag{6.2.6}$$

which is what we wanted to show.

$\square$

6.3 Proofs of the rules

Now we will prove Proposition 6.2.1 and Theorem 6.2.2.

Proof of Proposition 6.2.1 By the linearity of limits (see Proposition 4.1.2) we have

$$\lim_{h\to 0} \frac{(\alpha f+\beta g)(a+h)-(\alpha f+\beta g)(a)}{h} = \alpha \lim_{h\to 0} \frac{f(a+h)-f(a)}{h} + \beta \lim_{h\to 0} \frac{g(a+h)-g(a)}{h},$$

which is

$$\alpha f'(a)+\beta g'(a),$$

proving (i).

Next,

$$\lim_{h\to 0} \frac{(fg)(a+h)-(fg)(a)}{h} = \lim_{h\to 0} \frac{(f(a+h)-f(a))g(a+h)+f(a)(g(a+h)-g(a))}{h},$$

which equals, by Proposition 4.1.2,

$$\lim_{h\to 0} \frac{(f(a+h)-f(a))g(a+h)}{h} + \lim_{h\to 0} \frac{f(a)(g(a+h)-g(a))}{h} = f'(a)g(a)+f(a)g'(a),$$

yielding (ii).

Now on to the *quotient rule*. We have

$$\lim_{h\to 0} \frac{\frac{f}{g}(a+h)-\frac{f}{g}(a))}{h} = \lim_{h\to 0} \frac{f(a+h)-f(a)}{hg(a+h)} - \frac{f(a)}{g(a)} \lim_{h\to 0} g(a)-g(a+h)hg(a+h),$$

which equals

$$\frac{f'(a)g(a)-f(a)g'(a)}{g(a)^2},$$

as asserted in (iii).

□

Before beginning the proof of the *chain rule*, we need the following

Lemma 6.3.1 *Let f be a function defined in an interval around a real number a. Then the following are equivalent:*

(i) f is differentiable at a

(ii) There is a real number λ such that

$$\lim_{x\to a} \frac{f(x)-f(a)-\lambda(x-a)}{x-a} = 0.$$

When (ii) holds, the number λ is $f'(a)$.

Proof of Lemma. Note that f is differentiable at a iff the limit

$$L = \lim_{x\to a} \frac{f(x)-f(a)}{x-a}$$

exists, and when it exists, L is denoted by $f'(a)$. On the other hand,

$$\lim_{x\to a} \frac{f(x)-f(a)-\lambda(x-a)}{x-a} = \lim_{x\to a} \frac{f(x)-f(a)}{x-a} - \lambda = L-\lambda.$$

So L exists iff $L-\lambda$ exists and equals zero for some λ, namely for $\lambda = L$. It is now clear that (i) and (ii) are equivalent, and that when (ii) is satisfied, $\lambda = f'(a)$.

$\square$

Proof of Theorem 6.2.2. Let us put

$$b = f(a), \quad \phi = g\circ f,$$

$$F(x) = f(x)-f(a)-f'(a)(x-a), \; G(y) = g(y)-g(b)-g'(b)(y-b)$$

and

$$\Phi(x) = h(x)-h(a)-g'(b)f'(a)(x-a).$$

Then we have, by Lemma 6.3.1 above,

$$\lim_{x\to a}\frac{F(x)}{x-a} = 0 = \lim_{y\to b}\frac{G(y)}{y-b}. \tag{6.3.2}$$

In view of Lemma 6.3.1, it suffices to show:

$$\lim_{x\to a}\frac{\Phi(x)}{x-a} = 0. \tag{6.3.3}$$

But

$$\Phi(x) = g(f(x))-g(b)-g'(b)\left(f'(a)(x-a)\right),$$

and since $f'(a)(x-a) = f(x)-f(a)-F(x)$, we get

$$\Phi(x) = [g(f(x))-g(b)-g'(b)(f(x)-f(a))]+g'(b)F(x) = G(f(x))+g'(b)F(x).$$

Thanks to (6.3.2), it then suffices to prove:

$$\lim_{x\to a}\frac{G(f(x))}{x-a} = 0. \tag{6.3.4}$$

We know that $\lim_{y\to b}\dfrac{|G(y)|}{|y-b|}$ is zero. So we can find, for every $\varepsilon > 0$, a $\delta > 0$ such that $|G(f(x))| < \varepsilon|f(x)-b|$ if $|f(x)-b| < \delta$. But since f is integrable at a, it is also continuous there (see section 6.1). Hence $|f(x) - b| < \delta$ whenever $|x - a| < \delta_1$, for any small enough $\delta_1 > 0$. Hence

$$\begin{aligned} (f(x))| &< \varepsilon|f(x) - b| = \varepsilon|F(x) + f'(a)(x-a)| \\ &\leq \varepsilon(x)| + \varepsilon|f'(a)(x-a)|, \end{aligned}$$

by the triangle inequality. Since $\lim\limits_{x\to a} \frac{F(x)}{x-a}$ is zero, we get

$$\lim_{x\to a} \frac{|G(f(x))|}{|x-a|} \leq \varepsilon \lim_{x\to a} \frac{|f'(a)(x-a)|}{|x-a|} = \varepsilon|f'(a)|.$$

Now (6.3.4) follows easily, and we are done.

6.4 Tangents

Recall from High school Geometry that the **equation of any non-vertical line L in the plane** is given by

(6.4.1) $$y = mx + c,$$

where m is the **slope** of the line and $(0, c)$ is a point on L. The equation of a vertical line is given by $x = c_1$ for some constant c_1, and this will not concern us here because the graph of any function $f(x)$ will never be a vertical line (check it!).

To know a (non-vertical) line L as above, it is sufficient to know the slope m and any point $P = (x_0, y_0)$ on L. (P need not be $(0, c)$.) Indeed, since the coordinates of P must satisfy the equation of L, we have

(6.4.2) $$y_0 = mx_0 + c,$$

and combining (6.4.1) and (6.4.2) we get

$$y - y_0 = m(x - x_0),$$

or equivalently,

(6.4.3) $$y = mx + (y_0 - mx_0).$$

So we can recover (6.4.1) by setting

$$c = y_0 - mx_0.$$

Now let f be a function defined in an interval I around a point a. Suppose f is differentiable at a. Then, as noted earlier, the slope of the line T_a which is tangent to the graph C of $y = f(x)$, $x \in I$, is given by

(6.4.4) $$m = f'(a).$$

Moreover, since the point $(a, f(a)$ belongs to T_a, the equation of T_a is, thanks to (6.4.3), given by

(6.4.5) $$y = f'(a)x + (f(a) - f'(a)a).$$

6.5 Extrema of differentiable functions

Recall that any **continuous function** f on a closed interval $[a, b]$ attains its extrema, i.e., there exist $c, d \in [a, b]$ such that

$$f(c) \leq f(x) \leq f(d) \quad \forall\, x \in [a, b]. \tag{6.5.1}$$

One says that c is a **minimum** and d a **maximum** of f on $[a, b]$.

While this is a very beautiful result, we do not see how one can go about finding the points c and d where the extremal values occur.

Things become nicer when f is **differentiable** on (a, b), and even nicer when f is **twice differentiable** on (a, b), i.e., when f is differentiable and also the derivative f' is differentiable on (a, b). Note that it does not make sense to require that f is differentiable at either of the end points, because to have differentiability at a point, the function needs to be defined in an open interval surrounding the point.

Definition 6.5.2 *A function f defined on a set S containing a real number c has a* **local minimum**, *resp.* **local maximum**, *at c iff we can find a $\delta > 0$ such that for all $x \in (c - \delta, c + \delta) \cap S$,*

$$f(x) \geq f(c) \quad (\text{resp. } f(x) \leq f(c).$$

Clearly, a local minimum need not be an **absolute minimum** of f on S, but it provides a candidate to be tested for being the absolute minimum (if one exists, as for a continuous f).

Here is a basic result.

Proposition 6.5.3 *Let f be a function differentiable at a real number c.*

(i) If c is a local minimum or a local maximum, then $f'(c) = 0$.

(ii) Suppose f is twice differentiable at c with $f'(c) = 0$ and $f''(c)0$ (resp. $f''(c) < 0$). Then c is a local minimum (resp. local maximum) for f.

Proof. *(i)* Suppose c is a local minimum. Then by definition, we have for all small enough h,

$$f(a + h) - f(a) \geq 0.$$

Thus the **right derivative** of f at c satisfies

$$L^+ = \lim_{h \to 0,\, h > 0} \frac{f(a + h) - f(a)}{h} \geq 0$$

and the **left derivative** of f at c satisfies

$$L^- = \lim_{h \to 0,\, h < 0} \frac{f(a + h) - f(a)}{h} \leq 0.$$

Since by assumption f is differentiable at c, L^+ and L^- both exist and equal $f'(c)$. This forces $f'(c)$ to vanish as claimed. The argument is similar when c is a local maximum. Hence (i).

For the next part we need the following

Lemma 6.5.4 *Let f be a differentiable function on an open interval I such that $f'(x)$ is > 0 (resp. < 0) for all x in I. Then f is a* **strictly increasing function**, *resp.* **strictly decreasing function**, *on I.*

Proof of Lemma. Assume $f'(x) > 0$ for all $x \in I$. Pick any two points $a, b \in I$ with $a < b$. We have to show that $f(a) < f(b)$. Since $f'(a) > 0$, $f(a+h) > f(a)$ for all small $h > 0$. If $f(a) \leq f(b)$, then $f(a+h) > f(b)$ for small $h > 0$, and this will force the graph to peak at some point between a and b, i.e., f will have a local maximum at some c in (a, b). By part (i) of this Proposition, $f'(c)$ will then be 0, contradicting the hypothesis that the derivative of f is positive on all of I. Hence $f(a) < f(b)$, and the assertion is proved. The argument is similar when $f'(x) < 0$ for all $x \in I$.

$\square$

Proof of part (ii) of Proposition 6.5.3. Suppose $f'(c) = 0$ and $f''(c) > 0$. Then

$$f''(c) = \frac{f'(c+h)}{h} > 0.$$

Then $\frac{f(c+h)}{h}$ will need to be positive for all small h. When h is small and positive (resp. negative), this will mean that $f'(a+h)$ is positive (resp. negative). By the Lemma above, f will, for all small positive h, be strictly increasing in $(c, c+h)$ and strictly decreasing in $(c-h, c)$. So c will have to be a local maximum of f. By a similar argument, if $f'(c) = 0$ and $f''(c) < 0$, c will be a local maximum, as asserted.

$\square$

Suppose f is a differentiable function around a point c. We will say that c is a **critical point** of f iff $f'(c) = 0$.

A critical point can be a local minimum or a local maximum or of neither type. For example, the function

$$f(x) = x^3$$

has $x = 0$ as the only critical point, as $f'(x) = 3x^2$. But 0 is neither a local maximum nor a local minimum.

6.6 The mean value theorem

Here is a very useful result to know:

The Mean Value Theorem *Let f be a continuous function on $[a, b]$. Suppose f is differentiable on (a, b). Then we can find a point c in (a, b) where the following holds:*

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

A **special case** of this result, obtained by setting $f(a) = f(b)$, yields the following:

Rolle's Theorem *Let f be a continuous function on $[a, b]$. Suppose f is differentiable on (a, b). Then we can find a point c in (a, b) where $f'(c) = 0$.*

Prof of Rolle's theorem. Since f is continuous on $[a, b]$, we know that there exist c, d in $[a, b]$ for which (6.5.1) holds. If c or d lies in the open interval (a, b), then part (i) of Proposition 6.5.3 implies that the derivative vanishes there, proving Rolle's theorem. So we may assume that neither c nor d is in (a, b), i.e, $\{c, d\} = \{a, b\}$ as sets. Since by assumption, $f(a) = f(b)$, this forces us to have the maximum and the minimum value of f on $[a, b]$ to be $f(a)$ $(= f(b))$. The only possibility then is for f to be a constant function, forcing $f'(x)$ to be 0 at every $x \in (a, b)$.

□

Proof of the Mean Value Theorem. Put

$$m = \frac{f(b) - f(a)}{b - a}, \tag{6.6.1}$$

which is the slope of the line joining $(a, f(a))$ and $(b, f(b))$. Put

$$g(x) = f(x) - m(x - a). \tag{6.6.2}$$

Then by additivity, g is continuous on $[a, b]$ and differentiable on (a, b). Moreover

$$g(a) = f(a),$$

and

$$g(b) = f(b) - m(b - a). \tag{6.6.3}$$

Combining (6.6.1) and (6.6.3), we get

$$g(b) = f(b) - (f(b) - f(a)) = f(a).$$

Hence

$$g(a) = g(b).$$

So we may apply Rolle's theorem and conclude that for some c in (a, b),

$$g'(c) = 0. \tag{6.6.4}$$

On the other hand, by the definition (6.6.2) of $g(x)$,

$$g'(c) = f'(c) - m, \tag{6.6.5}$$

The assertion of the Theorem now follows by combining (6.6.4), (6.6.5) and (6.6.1).

□

We will end the section by noting certain facts concerning **convexity** and **concavity**.

Recall from the Homework assignment 3 that a point u lies in a closed interval $[c, d]$ iff we can write $u = (1-t)c + td$ for some real number t in $[0, 1]$. Given a function f on $[a, b]$, we will say that f is **convex** on $[a, b]$ iff for every subinterval $[c, d]$ and for every t in $[0, 1]$,, we have

$$f((1-t)c + td) \le (1-t)f(c) + tf(d). \tag{6.6.6}$$

It is said to be **concave** if the inequality is reversed, i.e., iff

$$f((1-t)c + td) \ge (1-t)f(c) + tf(d). \tag{6.6.7}$$

Note that the function f is **linear** iff equality holds everywhere.

A useful result, which we will not prove, is the following:

Proposition 6.6.8 *Let f be a function which is differentiable on an open interval containing $[a, b]$. Then f is convex, resp. concave, on $[a, b]$ if f' is an increasing, resp. decreasing, function on this interval.*

Note that if f is twice differentiable, we can assert that f is convex, resp. concave, iff f'' is everywhere positive, resp. negative.