30. Mean-Value and Taylor's Theorems

Equipped with the concept of the derivative and, in particular, the characterization of local extrema by the first derivative test, we will now draw a couple of interesting conclusions. The key words are the Mean-Value Theorem and Taylor's theorem.

30.1 Mean-Value Theorems.

We start with with a theorem that goes back to M. Rolle in 1691 albeit with a rigorous proof first given by A.L. Cauchy in 1823. The name Mean-Value Theorem is usually use to refer to the version attributed to J.-L. Lagrange.

Theorem 30.1 (Mean-Value Theorems of Rolle, Lagrange and Cauchy) *Let $a < b$ be reals and $f\colon [a,b] \to \mathbb{R}$ a function (with $\operatorname{Dom}(f) = [a,b]$) that is continuous on $[a,b]$ and differentiable on (a,b). Then*

(1) (Rolle's Theorem)

$$f(a) = f(b) \;\Rightarrow\; \exists x \in (a,b)\colon\; f'(x) = 0 \tag{30.1}$$

(2) (Lagrange's Theorem)

$$\exists x \in (a,b)\colon\; f'(x) = \frac{f(b) - f(a)}{b-a} \tag{30.2}$$

(3) (Cauchy's Theorem) *If also $g\colon [a,b] \to \mathbb{R}$ (with $\operatorname{Dom}(f) = [a,b]$) is continuous on $[a,b]$ and differentiable on (a,b), then*

$$\forall x \in (a,b)\colon g'(x) \neq 0 \tag{30.3}$$

implies

$$g(b) \neq g(a) \;\wedge\; \exists x \in (a,b)\colon\; \frac{f'(x)}{g'(x)} = \frac{f(b) - f(a)}{g(b) - g(a)} \tag{30.4}$$

Proof. We start with (1). Suppose f is as given with $f(a) = f(b)$. Then one of the three alternatives occur

$$\sup_{x \in [a,b]} f(x) > f(a) \;\vee\; \inf_{x \in [a,b]} f(x) < f(a) \;\vee\; \forall x \in [a,b]\colon\; f(x) = f(a) \tag{30.5}$$

Since (by Corollary 24.17) a continuous real-valued function on a compact set achieves its minimum and maximum, in all three cases f has a global extremum (maximum or minimum) at some $x \in (a,b)$. By Theorem 29.8, we then have $f'(x) = 0$ as desired.

Moving to the proof of (2), let $h\colon [a,b] \to \mathbb{R}$ be defined by

$$h(x) := f(x) - \frac{f(b) - f(a)}{b-a}(x-a) \tag{30.6}$$

Then, as is readily checked, h is continuous on $[a,b]$ and differentiable on (a,b) (being the difference of two functions with these properties). Moreover, $h(a) = f(a) = h(b)$ and so, by (1), there exists $x \in (a,b)$ with

$$0 = h'(x) = f'(x) - \frac{f(b) - f(a)}{b-a}. \tag{30.7}$$

This is the statement in (30.2).

Part (3), which subsumes part (2), is proved using a similar trick. First note that (30.2) and (30.3) imply $g(a) \neq g(b)$ for otherwise there would a point $x \in (a,b)$ where $g'(x)$ vanishes. This proves the first half of (30.4) and allows us to define $h\colon [a,b] \to \mathbb{R}$ by

$$h(x) := f(x) - \frac{f(b)-f(a)}{g(b)-g(a)}\big(g(x)-g(x)\big) \tag{30.8}$$

which is then again continuous on $[a,b]$ and differentiable on (a,b). By (1) we thus get the existence of $x \in (a,b)$ such that

$$0 = h'(x) = f(x) - \frac{f(b)-f(a)}{g(b)-g(a)}g'(x) \tag{30.9}$$

Dividing by $g'(x)$, which is non-zero by (30.3), we get the second half of (30.4). □

Lagrange's Mean-Value Theorem can be interpreted by saying that, for each continuous differentiable function on interval $[a,b]$ there is a point where the tangent line has the same slope as the secant line between the endpoints of the interval. Another "practical" consequence of the theorem is that, for a body that moves distance L in time T, there must be a time $t \in (0,T)$ where the instantaneous speed of motion equals L/T.

30.2 Applications.

Moving back to mathematics, we will now go over a couple of standard applications of Mean-Value Theorems. The first one concerns a well-known characterization of monotone differentiable functions:

Lemma 30.2 *Let $a < b$ be reals and $f\colon [a,b] \to \mathbb{R}$ a function (with $\mathrm{Dom}(f) = [a,b]$) that is continuous on $[a,b]$ and differentiable on (a,b). Then*

$$f \text{ non-decreasing on } [a,b] \;\Leftrightarrow\; \forall x \in (a,b)\colon\; f'(x) \geqslant 0 \tag{30.10}$$

Proof. We start with the easy direction ⇒. Indeed, assume that f is as above and non-decreasing. Let $x,y \in (a,b)$ be such that $y \neq x$. Then

$$\frac{f(y)-f(x)}{y-x} \geqslant 0 \tag{30.11}$$

and so $f'(x) \geqslant 0$ by the definition of the derivative (29.1).

The proof of ⇔ is done by contrapositive. Indeed, assume that $x,y \in [a,b]$ are such that $x < y$ and $f(y) < f(x)$. Then (30.3) implies existence of $z \in (x,y) \subseteq (a,b)$ such that

$$f'(z) = \frac{f(y)-f(x)}{y-x} < 0. \tag{30.12}$$

Hence, non-negativity of f' on (a,b) forces upward monotonicity of f on $[a,b]$. □

The use of non-strict monotonicity and non-negativity of the derivative is necessary. This is because a strictly increasing differentiable functions may not always have a strictly positive derivative. (The function $f(x) = x^3$ is an example.) The above theorem has a version that is useful in applications:

Corollary 30.3 *Let $f,g\colon [a,b]\to\mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b) and such that*

$$f(a)\leqslant g(a)\ \wedge\ \forall x\in(a,b)\colon\ f'(x)\leqslant g'(x) \tag{30.13}$$

Then

$$\forall x\in[a,b]\colon\ f(x)\leqslant g(x) \tag{30.14}$$

Proof. Let $h(x) := g(x)-f(x)$. By Lemma 30.2, h is non-decreasing. Since $h(a)\geqslant 0$ we have $h(x)\geqslant 0$ for all $x\in[a,b]$. □

One application arises when we need to prove a bound on a function. An example of this is $f(x) := \sin(x)$ (which we can think of as the unique solution to the second-order ODE $f''=-f$ with $f(0)=0$ and $f'(0)=1$). Then $f'(x)=\cos(x)$ and so one we know that $\cos(x)\leqslant 1$, the above gives $f(x)\leqslant x$ for all $x\geqslant 0$.

Another application is to the solutions of ordinary differential equations. Here is a statement in this vain:

Lemma 30.4 (Comparison of ODEs) *Let $F,G\colon\mathbb{R}\to\mathbb{R}$ (with domain all of the reals) be continuous functions with*

$$\forall u,v\in\mathbb{R}\colon\ u\leqslant v\ \Rightarrow\ F(u)\leqslant G(v). \tag{30.15}$$

Let $y,z\colon[a,b]\to\mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b) functions that solve the ordinary differential equations

$$\forall x\in(a,b)\colon\ y'(x)=F\big(y(x)\big)\ \wedge\ z'(x)=G\big(z(x)\big) \tag{30.16}$$

with the "initial" values such that $y(a)<z(a)$. Then

$$\forall x\in[a,b]\colon\ y(x)\leqslant z(x) \tag{30.17}$$

We leave the easy proof of this lemma to the reader. To demonstrate this on an example, consider the ODE

$$y'=y+\sqrt{y} \tag{30.18}$$

with initial value $y(0)=1$. Then $y'\geqslant y$ and so the above shows that y is bounded from below by the solution to the ODE

$$z'=z \tag{30.19}$$

with initial value $z(0)=a$ for any $a<1$. (This ODE happens to be solved by $z(x)=a\mathrm{e}^x$ so, taking $a\to 1$ from below we get $y(x)\geqslant \mathrm{e}^x$ for all $x\geqslant 0$.) These ideas drive the technique for solving *differential inequalities* which sometimes arise in applications.

Our next application of the Mean-Value Theorems is of more abstract nature. We start with a definition:

Definition 30.5 *Let $I\subseteq\mathbb{R}$ be a non-degenerate interval. A function $f\colon\mathbb{R}\to\mathbb{R}$ with $I\subseteq\mathrm{Dom}(f)$ is said to have the* intermediate value property *(IVP) on I if*

$$\forall x,y\in I\colon\ f(x)\leqslant f(y)\ \Rightarrow\ \big[f(x),f(y)\big]\subseteq f\big([\min\{x,y\},\max\{x,y\}]\big) \tag{30.20}$$

We remark that, at some point in early 19th century, the IVP was considered for a definition of continuity. As it turns out, the IVP is actually weaker than continuity as we defined it above. Indeed, the Intermediate Value Theorem shows that any continuous function on (a,b) will have an IVP. However, there are functions that have an IVP yet are not continuous. (Examples will be given after the next lemma.)

Lemma 30.6 *Let $f\colon [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b). Then f' has the intermediate value property on (a,b).*

Proof. Let $x,y \in (a,b)$ be such that $x < y$ and, without loss of generality, $f'(x) < f'(y)$ (otherwise, swap $-f$ for f). Pick $t \in (f'(x), f'(y))$ and define $h(u) := f(u) - tu$. Then $h'(x) = f'(x) - t < 0$ and so $h(x)$ is not a local minimum of h on $[x,y]$. Similarly, $h'(y) = f'(y) - t > 0$ and so $h(y)$ is not a local minimum of h on $[x,y]$ either. As h is continuous, Corollary 24.17 implies that it achieves its minimum in (x,y) and so, by Theorem 29.8, there exists $u \in (x,y)$ such that $h'(u) = 0$. This translates into $f'(u) = t$. As t was arbitrary in $(f'(x), f'(y))$, the function f' has the IVP on (a,b). □

Hereby we get:

Corollary 30.7 *Let $f\colon [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b). Then f' has no discontinuities of the first kind on (a,b).*

Proof. As is readily checked, if a function h has a discontinuity of the first kind at $x \in \operatorname{int}(\operatorname{Dom}(h))$, then h fails to have IVP in any open interval containing x. □

To give some examples, note that the function in (33.10) cannot be a derivative because it has a discontinuity of the first kind at every rational (and already suffices). For a positive example, consider the function $f\colon \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) := \begin{cases} x^2 \sin(1/x), & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases} \tag{30.21}$$

Then f is continuous and differentiable at each $x \neq 0$ with

$$f'(x) = 2x\sin(1/x) - \cos(1/x) \tag{30.22}$$

Note that $\lim_{x\to 0} f'(x)$ does NOT exist yet

$$\frac{f(x) - f(0)}{x - 0} = x\sin(1/x) \tag{30.23}$$

shows that $f'(0) = 0$. So f' exists on all of $\mathbb{R}$, has the IVP but is NOT continuous at 0. We will see that this example can be boosted to have a function which is differentiable on $\mathbb{R}$ but whose derivative is not continuous at any rational.

As our last application of the Mean-Value Theorem, we recall the well-known (and terribly overrated) result from Calculus:

Theorem 30.8 (l'Hospital's Rule, proved by J. Bernoulli in 1694) *Let $f\colon \mathbb{R} \to \mathbb{R}$ be defined, continuous and differentiable on $(a-\delta, a+\delta)$ for some $a \in \mathbb{R}$ and $\delta > 0$. Assume*

$$f(a) = 0 = f(b) \,\wedge\, \forall x \in (a-\delta, a) \cup (a, a+\delta)\colon\ g(x) \neq 0 \,\wedge\, g'(x) \neq 0. \tag{30.24}$$

Then

$$\lim_{x\to a} \frac{f'(x)}{g'(x)} \text{ exists } \Rightarrow \lim_{x\to a} \frac{f(x)}{g(x)} \text{ exists } \wedge \lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)} \tag{30.25}$$

Proof. The claim actually holds even for one-sided limits. Indeed, using that $f(a) = g(a) = 0$, for $x \in (a-\delta, a+\delta)$ with $x > a$, Cauchy's Mean-Value Theorem implies the existence of a $z_x \in (a, x)$ such that

$$\frac{f(x)}{g(x)} = \frac{f(x)-f(0)}{g(x)-g(0)} = \frac{f'(z_x)}{g'(z_x)} \tag{30.26}$$

As $x \to a^+$, we have $z_x \to a^+$ and, assuming the existence of the right limit of ratio of derivatives, we get

$$\lim_{x\to a^+} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)} \tag{30.27}$$

A similar statement holds for the limit from the left thus proving (30.25). □

As is well known from Calculus, l'Hospital's Rule is a tool to compute limit values of expressions of the indeterminate form $\frac{0}{0}$. Similar statements exist for other indeterminate expressions such as $\frac{\infty}{\infty}$ or limits of such ratios as $x \to \pm\infty$. That being said, there are examples where l'Hospital's Rule does not yield any conclusion; e.g., for $\lim_{x\to 0} \frac{1}{x}\mathrm{e}^{-1/x^2}$ where a formal application of l'Hospital's Rule asks us to compute $\lim_{x\to 0} \frac{2}{x^3}\mathrm{e}^{-1/x^2}$ instead. The fact is that, for this and other reasons, most working mathematicians pretty much never use l'Hospital's Rule as it stands but rather proceed more sophisticated methods such as the theorem that we will treat next.

30.3 Taylor's Theorem.

As we learned in the proof of Lemma 29.2 (and again in Lemma 29.3), for f having the derivative at a entails linear approximation near a of the form

$$f(x) = f(a) + f'(a)(x-a) + u_a(x)(x-a), \tag{30.28}$$

where the "error term" $u_a(x)(x-a)$ is a quantity of smaller order than the previous terms. This idea can be iterated provided we introduce:

Definition 30.9 (Derivatives of higher order) *Let $f\colon \mathbb{R} \to \mathbb{R}$ be differentiable (and f' thus defined) in an open set containing x. If f' is itself differentiable at x, we write $f''(x) := (f')'(x)$ for the* second derivative *of f at x. Similarly, we write $f'''(x) := (f'')'(x)$ to denote the* third derivative *of f at x provided f'' is differentiable at x, etc.*

Proceeding recursively, for each $n \in \mathbb{N}$ we denote the n-th derivative of f at x by $f^{(n)}(x)$. These are defined so that

$$f^{(0)}(x) = f(x) \ \wedge\ \forall n \in \mathbb{N}\colon\ f^{(n+1)}(x) = (f^{(n)})'(x) \tag{30.29}$$

whenever the derivative on the right exists. In the Leibnitz notation, we write $f^{(n)}$ as $\frac{\mathrm{d}^n f}{\mathrm{d}x^n}$.

We now generalize (30.28) as follows:

Theorem 30.10 (Taylor 1715, Gregory 1691) *Let $I \subseteq \mathbb{R}$ be an open interval, $n \in \mathbb{N}$ and assume that $f\colon I \to \mathbb{R}$ (with $\mathrm{Dom}(f) = I$) is $(n+1)$-times differentiable in I. For each $a, x \in I$ there exists $\xi \in (\min\{a,x\}, \max\{a,x\})$ such that*

$$f(x) = \left[\sum_{k=0}^{n} \frac{f^k(a)}{k!}(x-a)^k\right] + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1} \tag{30.30}$$

Let $P_n(x)$ be the quantity in the square brackets in (30.30); i.e.,

$$P_n(x) := \sum_{k=0}^{n} \frac{f^k(a)}{k!}(x-a)^k. \tag{30.31}$$

This is a polynomial of degree n in x conveniently written in powers of $x - a$ (which is a "small" quantity near a). We call P_n the *n-th order Taylor's polynomial* associated with f near a. The theorem then gives a quantitative bound on the difference $f(x) - P_n(x)$ by a quantity that is higher order than $P_n(x)$ itself and it leads to a *polynomial approximation* of f near a. Note that P_n shares all relevant derivatives of f at a:

$$\forall k = 0, \dots, n\colon\ P_n^{(k)}(a) = f^k(a) \tag{30.32}$$

As we will see, this is key for the proof:

Proof. Fix $a, x \in I$ and for simplicity assume $a < x$. Denote

$$A := \frac{f(x) - P_n(x)}{(x-a)^{n+1}} \tag{30.33}$$

and let $h\colon I \to \mathbb{R}$ be defined by

$$h(t) := f(t) - P_n(t) - A(t-a)^{n+1} \tag{30.34}$$

We now claim:

$$\forall k = 0, \dots, n+1\ \exists \xi_k \in (a, x]\colon\ h^{(k)}(\xi_k) = 0 \tag{30.35}$$

To prove this we proceed by induction. The base case is simple: The definition of A implies $h(x) = 0$ and so we can set $\xi_0 := x$. Assume now that $h^{(k)}(\xi_k) = 0$ for some $k \in \mathbb{N}$ with $k \leqslant n$. Note that, since $h^{(k)}$ is at least once differentiable, it is continuous on $[a, \xi_k]$ and differentiable on (a, ξ_k). Moreover, the observation (30.32) gives $h^{(k)}(a) = 0$ and so $h^{(k)}(a) = h^{(k)}(\xi_k)$. Rolle's Mean-Value Theorem then gives existence of $\xi_{k+1} \in (a, \xi_k) \subseteq (a, x]$ such that $h^{(k+1)}(\xi_{k+1}) = 0$. This proves (30.35) as stated.

With (30.35) established, observe that $P_n^{(n+1)}$ vanishes. It thus follows that

$$0 = h^{(n+1)}(\xi_{n+1}) = f^{(n+1)}(\xi_{n+1}) - (n+1)!A \tag{30.36}$$

Using the definition of A, this rewrites into (30.30) with $\xi := \xi_{n+1}$. □

We will not go into applications of Taylor's theorem as these have been practiced in Calculus and also because we will return to this theorem one more time once we have introduced the Riemann integral (which allows us to write the "error term" in integral form). We remark that the theorem can be given the following asymptotic form:

Theorem 30.11 (Taylor's theorem in asymptotic form) *Let $I \subseteq \mathbb{R}$ be an open interval, $n \in \mathbb{N}$ obey $n \geqslant 1$ and, given $a \in I$, assume that $f\colon I \to \mathbb{R}$ (with $\mathrm{Dom}(f) = I$) is $(n-1)$-times differentiable on I and n-times differentiable at a. Then*

$$\lim_{x \to a} \frac{f(x) - P_n(x)}{(x-a)^n} = 0 \tag{30.37}$$

We leave the proof of this version to a homework exercise. Note that here we only assume the existence of derivatives up to n; i.e., those needed to define P_n.