31. L'HÔSPITAL'S RULE AND TAYLOR'S THEOREM

We proceed with further applications of the Mean-Value Theorem to l'Hôspital's Rule and, after introducing higher derivatives, Taylor's Theorem.

31.1 l'Hôspital's Rule.

All of our previous uses of the Mean-Value Theorem referred to the part attributed to J.L. Lagrange. For an application of A.L. Cauchy's version of MVT, we recall the well-known (and terribly overrated) Calculus technique called l'Hôspital's Rule. As happens for many of these ancient results, the attribution is incorrect. Indeed, quoting from the corresponding wiki page: "The rule is named after the 17th-century French mathematician Guillaume de l'Hôpital, who published it in his 1696 textbook after learning it from his tutor, the Swiss mathematician Johann Bernoulli."

Theorem 31.1 (l'Hôspital's Rule, proved by J. Bernoulli in 1694) *Given $a \in \mathbb{R}$ and $\delta > 0$, let $f, g\colon \mathbb{R} \to \mathbb{R}$ be defined and continuous $(a-\delta, a+\delta)$ and differentiable on $(a-\delta, a+\delta) \smallsetminus \{a\}$. Assume, in addition, that*

$$f(a) = 0 = g(a) \ \wedge\ \forall x \in (a-\delta, a+\delta) \smallsetminus \{a\}\colon\ g(x) \neq 0 \ \wedge\ g'(x) \neq 0 \tag{31.1}$$

Then

$$\lim_{x\to a} \frac{f'(x)}{g'(x)} \text{ exists} \quad\Rightarrow\quad \lim_{x\to a} \frac{f(x)}{g(x)} \text{ exists} \ \wedge\ \lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)} \tag{31.2}$$

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

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(z_x)}{g'(z_x)} \tag{31.3}$$

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

$$\lim_{x\to a^+} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)} \tag{31.4}$$

A similar statement holds for the limit from the left thus proving (31.2). □

The first condition in (36.1) shows that (as is practiced to a point of exhaustion in Calculus) l'Hôspital's Rule is a tool to compute limits of expressions of the *indeterminate form* $\frac{0}{0}$. To demonstrate this on an example, we note

$$\lim_{x\to 1} \frac{x^\alpha - 1}{x^\beta - 1} = \lim_{x\to 1} \frac{\alpha x^{\alpha-1}}{\beta x^{\beta-1}} = \frac{\alpha}{\beta} \tag{31.5}$$

Similar statements exist for other indeterminate expressions such as $\frac{\infty}{\infty}$ or limits of such ratios as $x \to \pm\infty$. (One way to deal with the case $\frac{\infty}{\infty}$ is to write f/g as $(g/f)^{-1}$ which effectively converts the indeterminate form $\frac{\infty}{\infty}$ to the form $\frac{0}{0}$.)

That being said, there are examples where l'Hôspital's Rule does not yield any conclusion; e.g., for $\lim_{x\to 0} \frac{1}{x} e^{-1/x^2}$ where a formal application of l'Hôspital's Rule asks us to compute $\lim_{x\to 0} \frac{2}{x^3} e^{-1/x^2}$ which seems harder than the limit we started with. (Iterating further only makes this worse.) The fact is that, for this and other reasons, most working mathematicians pretty much never use l'Hôspital's Rule as it stands but rather proceed more sophisticated methods such as Taylor's theorem.

31.2 Higher derivatives.

The Lagrange notation for the derivative is suggestive of a map that assigns the function f' to the function f (with $\text{Dom}(f')$ being a subset, even proper or possibly empty, of $\text{Dom}(f)$). Iterating this map further leads to the concept of higher derivatives which we need to introduce next.

Let $f\colon \mathbb{R} \to \mathbb{R}$ be defined (and f' thus defined) in an open set containing x. If f' is itself differentiable at x, we write

$$f''(x) := (f')'(x) \tag{31.6}$$

for the *second derivative* of f at x. Similarly, we write

$$f'''(x) := (f'')'(x) \tag{31.7}$$

to denote the *third derivative* of f at x whenever f'' is defined in an open set containing x and is differentiable at x, etc. Formalizing this further, this leads to:

Definition 31.2 (Derivatives of higher order) *For all $f\colon \mathbb{R} \to \mathbb{R}$, let $\{f^{(n)}\}_{n\in\mathbb{N}}$ be functions such that*

$$\text{Dom}(f^{(0)}) = \text{Dom}(f) \,\wedge\, \forall x \in \text{Dom}(f)\colon\; f^{(0)}(x) = f(x) \tag{31.8}$$

and that

$$\forall n \in \mathbb{N}\colon\; \text{Dom}(f^{(n+1)}) = \big\{x \in \text{int}(\text{Dom}(f^{(n)}))\colon f^{(n)} \text{ differentiable at } x\big\} \tag{31.9}$$

and

$$\forall n \in \mathbb{N}\, \forall x \in \text{Dom}(f^{(n+1)})\colon\; f^{(n+1)}(x) = (f^{(n)})'(x) \tag{31.10}$$

We then call $f^{(n)}$ the n-th derivative *of f. In the Leibnitz notation, we write $f^{(n)}$ as $\frac{\mathrm{d}^n f}{\mathrm{d}x^n}$.*

If $[x, x+\delta) \subseteq \text{Dom}(f^{(n)})$ then we can of course talk about the right $(n+1)$-th derivative at x, and similarly for the left derivative. The reader has probably encountered the second derivative in the following context:

Theorem 31.3 (Second derivative test) *Let $f\colon \mathbb{R} \to \mathbb{R}$ and $x \in \text{int}(\text{Dom}(f))$ be such that f is differentiable on $(x-\delta, x+\delta)$, for some $\delta > 0$. Assume also that $f''(x)$ exists. Then*

$$f'(x) = 0 \,\wedge\, f''(x) > 0 \;\Rightarrow\; f \text{ has a strict local minimum at } x \tag{31.11}$$

If instead $f''(x) < 0$, then f has a strict local maximum at x.

Proof. The existence of $f''(x)$ implies that f' is continuous at x. The inequality $f''(x) > 0$ than forces existence of $\delta' \in (0, \delta)$ such that

$$\forall z \in (x-\delta', x)\colon\; f'(z) < f'(x) \tag{31.12}$$

and

$$\forall z \in (x, x+\delta')\colon\; f'(z) > f'(x) \tag{31.13}$$

Since $f'(x) = 0$, we get $f'(z) < 0$ for $x \in (x-\delta', x)$ while $f'(z) > 0$ for $x \in (x, x+\delta')$. The Mean-Value Theorem then gives $f(z) > f(x)$ for all $z \in (x-\delta', x+\delta') \smallsetminus \{x\}$, proving that f has a strict local minimum at x. □

As the example $f(x) := x^3$ shows that that no information is obtained in the case where $f''(x) = 0$. However, the conclusion still holds when $f''(x) = 0$ but $f''(z) \geqslant 0$ for z in an open interval containing x such that $z \neq x$. The points where f'' changes sign are called *inflection points*. This is due to the relation to the following concept:

Definition 31.4 (Convex/concave functions) *Let V be a vector space. We say that a function $f\colon V \to \mathbb{R}$ is convex if for all $x, y \in \operatorname{Dom}(f)$ and all $\alpha \in [0,1]$,*

$$\alpha x + (1-\alpha)y \in \operatorname{Dom}(f) \;\wedge\; f\big(\alpha x + (1-\alpha)y\big) \leqslant \alpha f(x) + (1-\alpha)f(y) \tag{31.14}$$

We say that f is concave if $-f$ is convex.

The relation to the derivative now comes in:

Lemma 31.5 *Let $a < b$ be reals and $f\colon (a,b) \to \mathbb{R}$ with $\operatorname{Dom}(f) = (a,b)$ be convex. Then*

1. *f is continuous on (a,b)*
2. *f is right and left differentiable on (a,b)*
3. *the one-sided derivatives $f^{\prime,+}$ and $f^{\prime,-}$ are non-decreasing with*
$$\forall x \in (a,b)\colon\; f^{\prime,-}(x) \leqslant f^{\prime,+}(x) \tag{31.15}$$
and
$$\forall x, y \in (a,b)\colon\; x < y \;\Rightarrow\; f^{\prime,+}(x) \leqslant \frac{f(y)-f(x)}{y-x} \leqslant f^{\prime,-}(y) \tag{31.16}$$
4. *$f^{\prime,+}$ is right-continuous and $f^{\prime,-}$ is left-continuous*
5. *if $x \in (a,b)$ is such that $f''(x)$ exists, then $f''(x) \geqslant 0$*

We leave the proof of these claims to homework. A key point of our interest here is part (5); particularly, if combined with:

Lemma 31.6 *Let $a < b$ be reals and $f\colon (a,b) \to \mathbb{R}$ with $\operatorname{Dom}(f) = (a,b)$ be twice differentiable on (a,b). Then*

$$f'' \geqslant 0 \text{ on } (a,b) \;\Rightarrow\; f \text{ is convex on } (a,b) \tag{31.17}$$

Proof. Let $x, y \in (a,b)$ with $x < y$ and let $h\colon [0,1] \to \mathbb{R}$ be defined by

$$h(\alpha) := f\big(\alpha x + (1-\alpha)y\big) - \big[\alpha f(x) + (1-\alpha)f(y)\big] \tag{31.18}$$

The Chain Rule then tells us

$$h''(\alpha) = f''\big(\alpha x + (1-\alpha)y\big)(x-y)^2 \tag{31.19}$$

and so $h'' \geqslant 0$ on (a,b). The Second-derivative test implies that h cannot have a local maximum inside (a,b) and, noting that h is continuous, the maximum of h on $[0,1]$ is achieved at the endpoints. But $h(0) = 0 = h(1)$ and so $h \leqslant 0$ on $[0,1]$. This now readily translates into the convexity of f on $[x,y]$. □

Using this lemma we conclude that the regions of convexity of a twice-differentiable function are exactly those where the second derivative is non-negative, while those of concavity are those where the second derivative is non-positive. The inflection points are those where type of convexity changes.

For readers wondering why the above lemmas are stated over an open interval, observe that if $f\colon [a,b] \to \mathbb{R}$ is convex, it is continuous on (a,b) but it may not be continuous at the endpoints of $[a,b]$. Similarly, even if it is continuous $[a,b]$, the right/left derivative may not exist in proper sense at the endpoints as the associate limit diverges to negative infinity at a or positive infinity at b.

Convexity is an extremely useful property in all sorts of analytic arguments. To give an example, we note the following consequence of Lemma 31.5:

Lemma 31.7 (Derivative of exponentials) *Given $a > 0$, let $f_a\colon \mathbb{R} \to \mathbb{R}$ with $\operatorname{Dom}(f_a) = \mathbb{R}$ be defined by $f_a(x) := a^x$. Then f_a is convex and, in fact, continuously differentiable on $\mathbb{R}$ with $f_a'(x) > 0$ for all $x \in \mathbb{R}$. Moreover, we have*

$$\forall x \in \mathbb{R}\colon\;\; f_a'(x) = f_a'(0)a^x \tag{31.20}$$

with

$$\forall a,b > 0\colon\;\; b \neq 1 \;\Rightarrow\; f_a'(0) = f_b'(0)\log_b(a) \tag{31.21}$$

where $\log_b$ is the inverse to f_b.

In order to prove convexity of f_a, we need to show that

$$a^{(1-\alpha)x+\alpha y} \leqslant (1-\alpha)a^x + \alpha a^y \tag{31.22}$$

holds for all $x,y \in \mathbb{R}$ and all $\alpha \in [0,1]$. Here we note that, given any $u,v \geqslant 0$, the inequality $0 \leqslant (\sqrt{u}-\sqrt{v})^2 = u+v-2\sqrt{uv}$ implies the *AMGM inequality*

$$\forall u,v \geqslant 0\colon\;\; \sqrt{uv} \leqslant \frac{u+v}{2} \tag{31.23}$$

Setting $u := a^x$ and $v := a^y$ then proves (31.22) for $\alpha = 1/2$. This suffices thanks to:

Lemma 31.8 *Let $f\colon \mathbb{R} \to \mathbb{R}$ with $\operatorname{Dom}(f)$ an open interval be such that*

$$\forall x,y \in \operatorname{Dom}(f)\colon\;\; f\Bigl(\frac{x+y}{2}\Bigr) \leqslant \frac{f(x)+f(y)}{2} \tag{31.24}$$

If f is continuous, then it is convex.

Proof. Let $\mathbb{D}_n := \{k2^{-n}\colon k = 0,\dots,2^n\}$. We will prove by induction that, for all $n \in \mathbb{N}$,

$$\forall \alpha \in \mathbb{D}_n\, \forall x,y \in \operatorname{Dom}(f)\colon\;\; f\bigl((1-\alpha)x+\alpha y\bigr) \leqslant (1-\alpha)f(x) + \alpha f(y) \tag{31.25}$$

This holds trivially for $n = 0$ because $\mathbb{D}_0 = \{0,1\}$, so let us suppose that it holds for some $n \in \mathbb{N}$. Let $\alpha \in \mathbb{D}_{n+1} \smallsetminus \mathbb{D}_n$. Denote $\tilde\alpha := \alpha - 2^{-n-1}$ and note that $\alpha \notin \mathbb{D}_n$ implies $\tilde\alpha, \tilde\alpha + 2^{-n} \in \mathbb{D}_n$. Invoking (31.24) along with (31.25) for n then gives

$$\begin{aligned}
f\bigl((1-\alpha)x+\alpha y\bigr) &= f\Biggl(\frac{(1-\tilde\alpha)x+\tilde\alpha y}{2} + \frac{(1-\tilde\alpha-2^{-n})x+(\tilde\alpha+2^{-n})y}{2}\Biggr)\\
&\leqslant \frac12 f\bigl((1-\tilde\alpha)x+\tilde\alpha y\bigr) + \frac12 f\bigl((1-\tilde\alpha-2^{-n})x+(\tilde\alpha+2^{-n})y\bigr)\\
&\leqslant \frac12\Bigl((1-\tilde\alpha)f(x)+\tilde\alpha f(y)\Bigr) + \frac12\Bigl((1-\tilde\alpha-2^{-n})f(x)+(\tilde\alpha+2^{-n})f(y)\Bigr)
\end{aligned} \tag{31.26}$$

The right-hand side is now readily identified with $(1-\alpha)f(x)+\alpha f(y)$ proving the claim for $n+1$ and, by induction, for all $n \in \mathbb{N}$.

We have thus shown that the inequality in (31.25) holds for all $\alpha \in \mathbb{D} := \bigcup_{n\in\mathbb{N}} \mathbb{D}_n$. But the set $\mathbb{D}$ of dyadic rationals is dense in $[0,1]$ and so (taking suitable limits) the claim holds for all $\alpha \in [0,1]$ by the assumed continuity of f. □

Proof of Lemma 31.7. The function f_a is continuous and, by our reasoning, obeys (31.25) and so it is convex. By Lemma 31.5, the one-sided derivatives of f exist. Moreover,

$$f'^{,\pm}(x) = \lim_{h\to 0^\pm} \frac{a^{x+h}-a^x}{h} = a^x \lim_{h\to 0^\pm} \frac{a^h-1}{h} = a^x f_a'^{,\pm}(0) \tag{31.27}$$

Since $\{x \in \mathbb{R} \colon f'^{,+}(x) \neq f'^{,-}(x)\}$ is at most countable, there exists $x \in \mathbb{R}$ where $f_a'(x)$ exists. But then the derivative exists at zero and thus at all points with (31.20) in force. For (31.21) we just note that $a^x = b^{\log_b(a)x}$ which upon diffentiation gives the result. □

For $b > 1$, the function $a \mapsto \log_b(a)$ is strictly increasing, continuous with range being all of $\mathbb{R}$. Hence there exists a unique $a > 1$ such that $f_a'(0) = f_b'(0)\log_b(a) = 1$. We give this a a special name:

Definition 31.9 *The Euler constant* e *is the unique positive real such that* $f_{\mathrm{e}}{}'(0) = 1$.

The definition of e is such that $x \mapsto \mathrm{e}^x$ differentiates to itself. This is naturally advantageous in calculations. The corresponding inverse, to be denoted $\log(x)$ with $\ln(x)$ sometimes used instead, is the so called *natural logarithm*.

A corollary of the above proof is the well-known inequality:

Corollary 31.10 (General AMGM inequality)

$$\forall u,v \geqslant 0\, \forall \alpha \in [0,1]\colon\;\; u^{1-\alpha}v^\alpha \leqslant (1-\alpha)u + \alpha v \tag{31.28}$$

Proof. Apply Lemma 31.8 with (31.23) used as input. Alternatively, with the exponentials differentiable, compute the second derivative of the left-hand side with respect to α is non-negative and apply Lemma 36.7. □

31.3 Taylor's Theorem.

Moving to another application of the MVT, recall that, as we learned in the proof of Lemma 29.5 (and again in Lemma 29.6), f having the derivative at a entails a linear approximation near a of the form

$$f(x) = f(a) + f'(a)(x-a) + u_a(x)(x-a), \tag{31.29}$$

where the "error term" $u_a(x)(x-a)$ is a quantity of smaller order than the previous terms. In short, f is approximated by a linear function up to errors that vanish at linear order. This idea can be iterated to yield:

Theorem 31.11 (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. Then for all*

$a, x \in I$ there exists $z \in (\min\{a,x\}, \max\{a,x\})$ such that

$$f(x) = \left[\sum_{k=0}^{n} \frac{f^k(a)}{k!}(x-a)^k\right] + \frac{f^{(n+1)}(z)}{(n+1)!}(x-a)^{n+1} \tag{31.30}$$

Let $P_n(x)$ be the quantity in the square brackets in (31.30); i.e.,

$$P_n(x) := \sum_{k=0}^{n} \frac{f^k(a)}{k!}(x-a)^k \tag{31.31}$$

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:

$$\forall k = 0, \dots, n\colon\ P_n^{(k)}(a) = f^k(a) \tag{31.32}$$

As we will see, this is an important ingredient for:

Proof of Theorem 31.11. Fix $a, x \in I$ and for simplicity assume $a < x$. Abbreviate

$$A := \frac{f(x) - P_n(x)}{(x-a)^{n+1}} \tag{31.33}$$

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

$$h(t) := f(t) - P_n(t) - A(t-a)^{n+1} \tag{31.34}$$

Observe that (31.32) implies

$$\forall k = 0, \dots, n\colon\ \ h^{(k)}(a) = 0 \tag{31.35}$$

We now claim

$$\forall k = 0, \dots, n+1\ \exists z_k \in (a, x]\colon\ \ h^{(k)}(z_k) = 0 \tag{31.36}$$

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 $z_0 := x$. Assume now that $h^{(k)}(z_k) = 0$ for some $k \in \mathbb{N}$ with $k \leqslant n$. Note that $h^{(k)}$ is continuous on $[a, z_k]$ and differentiable on (a, z_k). In light of (31.35) we have $h^{(k)}(a) = h^{(k)}(z_k)$ and so Rolle's Mean-Value Theorem gives existence of a point $z_{k+1} \in (a, z_k) \subseteq (a, x]$ such that $h^{(k+1)}(z_{k+1}) = 0$. This proves (31.36) as stated.

With (31.36) established, observe that $P_n^{(n+1)}$ vanishes. It thus follows that

$$0 = h^{(n+1)}(z_{n+1}) = f^{(n+1)}(z_{n+1}) - (n+1)!A \tag{31.37}$$

Using the definition of A, this rewrites into (31.30) with $z := z_{n+1}$. □

As an applications of Taylor's theorem, we show how to control Newton's method for numerically finding a root of a function:

Lemma 31.12 (Newton's method) *Let $a < b$ be reals and let $f\colon [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and twice differentiable on (a,b). Assume $f(a) < 0 < f(b)$ and the following*

$$A := \inf_{x \in (a,b)} f'(x) > 0\ \wedge\ B := \sup_{x \in (a,b)} \left|f''(z)\right| < \infty \tag{31.38}$$

Then there exists unique $z \in (a,b)$ such that $f(z) = 0$. Moreover, for any $\delta > 0$ such that $(z-\delta, z+\delta) \subseteq (a,b)$ and $B\delta \leqslant 2A$, the function $\phi\colon (a,b) \to \mathbb{R}$ defined by

$$\phi(x) := x - \frac{f(x)}{f'(x)} \tag{31.39}$$

maps $(z-\delta, z+\delta)$ into itself. In particular, for any $x_0 \in (a,b)$ with $|x_0 - z| < \delta$, the sequence $\{x_n\}_{n\in\mathbb{N}}$ defined recursively so that

$$\forall n \in \mathbb{N}\colon\ \ x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \tag{31.40}$$

obeys $\forall n \in \mathbb{N}\colon x_n \in (z-\delta, z+\delta)$ and

$$\forall n \in \mathbb{N}\colon\ \ |x_n - z| \leqslant \left(\frac{B}{2A}\right)^{2^n-1} |x_0 - z|^{2^n} \tag{31.41}$$

Proof. The existence of z with $f(z) = 0$ follows from the Intermediate Value Theorem. Since f is strictly increasing, the solution is necessarily unique. Given $x \in (a,b)$, Taylor's theorem centered at x then tells us

$$0 = f(z) = f(x) + f'(x)(z-x) + \frac{f''(y)}{2}(z-x)^2 \tag{31.42}$$

for some y between x and z. Dividing by $f'(x)$ gives us

$$\frac{f(x)}{f'(x)} = x - z + \frac{f''(y)}{2f'(x)}(x-z)^2 \tag{31.43}$$

which upon inserting into the definition of ϕ shows

$$\phi(x) - z = x - z - \left[(x-z) - \frac{f''(y)}{2f'(x)}(x-z)^2\right] = \frac{f''(y)}{2f'(x)}(x-z)^2 \tag{31.44}$$

Taking absolute values and invoking the definitions of A and B results in the bound

$$\big|\phi(x) - z\big| \leqslant \frac{B}{2A}|x-z|^2 \tag{31.45}$$

In particular, if $\delta > 0$ is such that $(z-\delta, z+\delta) \subseteq (a,b)$ and $B\delta \leqslant 2A$, which is possible thanks to $A > 0$, the function ϕ maps $(z-\delta, z+\delta)$ into itself.

With $\delta > 0$ as above, we may iterate ϕ starting from any $x_0 \in (z-\delta, z+\delta)$ to define a sequence $\{x_n\}_{n\in\mathbb{N}}$ satisfying (31.40). The inequality (31.45) then results in the recursive bound

$$\forall n \in \mathbb{N}\colon\ \ |x_{n+1} - z| \leqslant \frac{B}{2A}|x_n - z|^2 \tag{31.46}$$

Using induction, we now verify the claim (31.41). □

The upshot of (31.41) is that the convergence $x_n \to z$ is *doubly-exponentially* fast. (Remember that contraction maps, of which ϕ is an example, typically result only in an exponentially-fast convergence.) To demonstrate this, assume that we start the iterations from x_0 such that $|x_0 - z| \leqslant \delta/2$. The bound (31.41) along with $B\delta \leqslant 2A$ then yields $|x_n - z| \leqslant \delta 2^{-2^n}$ which decays to zero remarkably fast (note that 2^{10} is about 1000 so $2^{2^{10}}$ is about 10^{100}).

A disadvantage of Newton's method is that one needs start them close enough to the root (which is tantamount to finding δ with above properties). Indeed, starting from a generic point in $[a, b]$ often takes the iterations out of the interval. Another disadvantage is that we need to be able to evaluate f', as opposed to just f. This poses no problem for polynomials; for instance, taking $f(x) := x^2 - \theta$ for $\theta > 0$ gives us a way to compute $\sqrt{\theta}$ numerically as the limit of the sequence $\{x_n\}_{n\in\mathbb{N}}$ where x_0 is chosen "close" to $\sqrt{\theta}$ and further terms are defined recursively via

$$x_{n+1} := x_n - \frac{x_n^2 - \theta}{2x_n} = \frac{1}{2}\left(x_n + \frac{\theta}{x_n}\right) \tag{31.47}$$

(The convexity of f ensures convergence regardless of how far x_0 is from $\sqrt{\theta}$ as long as it is positive.) The resulting *Heron's method* for computing square root dates back to Heron of Alexandria who published this algorithm in the year 60AC.

Theorem 31.11 is not the end of the story. Indeed, we will return to it one more time once we have introduced the Riemann integral which allows us to write the "error term" in integral form. Another version of the result is the following asymptotic form:

Theorem 31.13 (Taylor's theorem in asymptotic form) *Given an open interval $I \subseteq \mathbb{R}$ a natural $n \geqslant 1$ and $a \in I$, 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*

$$\lim_{x\to a} \frac{f(x) - P_n(x)}{(x-a)^n} = 0 \tag{31.48}$$

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. Moreover, existence of the n-th derivative is assumed only at the point a.