NOTES ON ANALYSIS.

RALPH HOWARD

Contents

1. More on limits of sequences.

To make a definition we used more or less implicitly last semester:

Definition 1.1. Let $\langle a_k\rangle_{k=1}^\infty$ be a sequence of real numbers. Then

$$\lim_{n\to\infty} a_n = \infty$$

means that for any real number M there is a $N > 0$ so that

$$n \geq N \quad \text{implies} \quad a_n \geq M.$$

Likewise

$$\lim_{n\to\infty} a_n = -\infty$$

means that for any real number M there is a $N > 0$ so that

$$n \geq N \quad \text{implies} \quad a_n \leq M.$$

(In this case think of M as be a negative number very far to the left on the number axis.)

Example 1.2. As a more or less obvious example let p be a positive integer and $a > 0$ then $\lim_{n\to\infty} an^p = \infty$.

To prove this let $M > 0$ and let $N = (M/a)\frac{1}{p}$. Then if $n \geq N$ we have $an^p \geq a\left((M/a)^{\frac{1}{p}}\right)^p = M$. □

Since it is easier to work in inequalities using positive numbers rather than negative numbers the following is obvious, but nice.

Proposition 1.3. *Let $\langle a_n\rangle_{n=1}^{\infty}$ be a sequence or real numbers. Then*

$$\lim_{n\to\infty} a_n = -\infty$$

if and only if

$$\lim_{n\to\infty} -a_n = \infty.$$

Proposition 1.4. *Let $\langle a_n\rangle_{n=1}^{\infty}$ and $\langle b_n\rangle_{n=1}^{\infty}$ be sequences of real numbers with*

$$\lim_{n\to\infty} a_n = L$$

(That is a usual finite limit) and

$$\lim_{n\to\infty} b_n = \infty.$$

Then

$$\lim_{n\to\infty} (a_n + b_n) = \infty.$$

Problem 1.1. Prove this. □

Along the same lines we have

Proposition 1.5. *Let $\langle a_n\rangle_{n=1}^{\infty}$ and $\langle b_n\rangle_{n=1}^{\infty}$ be sequences of real numbers with*

$$\lim_{n\to\infty} a_n = L$$

where $L > 0$

$$\lim_{n\to\infty} b_n = \infty.$$

Then

$$\lim_{n\to\infty} a_n b_n = \infty.$$

Problem 1.2. Prove this. □

Problem 1.3. Here is an example of Proposition 1.5 in action. Let

$$p(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0$$

be a polynomial with $a_n > 0$ with $m \geq 1$. Show

$$\lim_{n\to\infty} p(n) = \infty.$$

Hint: To start rewrite $p(n)$ as

$$p(n) = n^m \left(a_m + \frac{a_{m-1}}{n} + \frac{a_{n-2}}{n^2} + \cdots + \frac{a_1}{n^{m-1}} + \frac{a_0}{n^m} \right).$$

Then

$$\lim_{n\to\infty} n^m = \infty$$

$$\lim_{n\to\infty} \left(a_m + \frac{a_{m-1}}{n} + \frac{a_{n-2}}{n^2} + \cdots + \frac{a_1}{n^{m-1}} + \frac{a_0}{n^m} \right) = a_m > 0.$$

□

Problem 1.4. If $\lim_{n\to\infty} a_n = \infty$, show $\lim_{n\to\infty} \dfrac{1}{a_n} = 0$. □.

Proposition 1.6. *Let $\langle a_n \rangle_{n=1}^{\infty}$ be a monotone sequence. Then $\lim_{n\to\infty} a_n$ exists, but might have the value ∞ or $-\infty$.*

Problem 1.5. Prove this. *Hint:* It is enough to prove this in the case the case $\langle a_n \rangle_{n=1}^{\infty}$, otherwise replace the sequence with $\langle -a_n \rangle_{n=1}^{\infty}$, It the sequence is bounded above, then his is some thing we have seen before. So you just have to show that a monotone increasing sequence that is not bounded below has ∞ as its limit. □

Problem 1.6. If $a_n > 0$ and $\lim_{n\to\infty} = 0$, show $\lim_{n\to\infty} \dfrac{1}{a_n} = \infty$. □

Let us extend our notation of supremum and infinitum a bit. If S is a nonempty subset of $\mathbb{R}$ which is bounded above, then $\sup(S)$ is the least upper bound of S. If $S \subseteq \mathbb{R}$ then set

$$\sup(S) = \begin{cases} -\infty, & S = \varnothing; \\ \sup(S), & S \text{ is bounded above;} \\ \infty, & S \text{ is not bounded above.} \end{cases}$$

That $\sup(S)$ should be defined as ∞ when S is not bounded above makes sense. That $\sup(\varnothing) = -\infty$ takes a bit more logic to seem reasonable. Let $b \in \mathbb{R}$ then the implication

$$s \in \varnothing \quad \text{implies} \quad s \leq b$$

holds. This is because an implication $P \implies Q$ is always true when the hypothesis, P, is false. As $s \in \varnothing$ is always false the implication $s \in \varnothing \implies s \leq b$ true. Thus b is an upper bound for $\varnothing$. This holds for all $b \in \mathbb{R}$ and therefore every real number is an upper bound for $\varnothing$. Viewed this way the only reasonable definition of $\sup(\varnothing)$ is ∞.

Likewise we have

$$\inf(S) = \begin{cases} \infty, & S = \varnothing; \\ \inf(S), & S \text{ is bounded below;} \\ -\infty, & S \text{ is not bounded below.} \end{cases}$$

Problem 1.7. Let $\langle a_n \rangle_{n=1}^{\infty}$ is a sequence in $\mathbb{R}$. For each n set

$$A_n = \sup\{a_n, a_{n+1}, a_{n+2}, \ldots\}.$$

(Note that A_n may be infinite, for example for the sequence where $a_n = n^2$ we have $A_n = \infty$ for all n.)

(a) Show that the sequence $\langle A_n \rangle_{n=1}^{\infty}$ is monotone decreasing.
(b) Show that $\lim_{n\to\infty} A_n$ exists (but might be either ∞ or $-\infty$.

Definition 1.7. Let $\langle a_n \rangle_{n=1}^{\infty}$ be a sequence in $\mathbb{R}$. Then define

$$\limsup_{n\to\infty} a_n = \sup\{a_n hj, a_{n+1}, a_{n+2}, \ldots\}.$$

(In the notation of the previous problem this is just $\lim_{n\to\infty} A_n$.

The number $\limsup_{n\to\infty} a_n$ is the ***limit superior*** of the sequence, but is generally just called the ***lim sup***. Likewise we have

Definition 1.8. Let $\langle a_n \rangle_{n=1}^{\infty}$ be a sequence in $\mathbb{R}$. Then define

$$\liminf_{n\to\infty} a_n = \inf\{a_, a_{n+1}, a_{n+2}, \ldots\}.$$

(In the notation of the previous problem this is just $\lim_{n\to\infty} A_n$.

Problem 1.8. Find the lim sup and lim inf for the following sequences. $\langle a_n \rangle_{n=1}^{\infty}$

(a) $a_n = (-1)^n$.
(b) $a_n = \sin(n)$ (This requires the following fact which you can assume: the set $\{\sin(n) : n \in \mathbb{N}\}$ is dense in $[-1, 1]$. That is between any two numbers $\alpha < \beta$ in $[-1, 1]$ there is a n (in fact infinitely many) with $\alpha < \sin(n) < \beta$.)

Problem 1.9. Let $\langle a_n \rangle_{n=1}^{\infty}$ be a bounded sequence. Show $\lim_{n\to\infty} a_n$ exists if and only if $\liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n$. □

Problem 1.10. Let $\langle a_n \rangle_{n=1}^{\infty}$ and $\langle b_n \rangle_{n=1}^{\infty}$ are sequences of real numbers prove the following

(a) $\lim_{n\to\infty} = L$ and $\limsup_{n\to\infty} = M$ then $\limsup_{n\to\infty}(a_n + b_n) = L + M$.
(b) $\lim_{n\to\infty} = L > 0$ and $\limsup_{n\to\infty} = M$ then $\limsup_{n\to\infty} a_n b_n = LM$.

Problem 1.11. Let $\langle a_n \rangle_{n=1}^{\infty}$ and $\langle b_n \rangle_{n=1}^{\infty}$ be bounded sequences. Prove

$$\lim_{n\to\infty}(a_n + b_n) \le \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n$$

and given an examples both where equality holds and where the inequality is strict. □

2. The derivative.

2.1. The derivative at a point.

Definition 2.1. Let (α, β) be an open interval, $f\colon (\alpha,\beta) \to \mathbb{R}$ a function, and $a \in (\alpha, \beta)$. Then f is ***differentiable*** at if and only if the limit

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

exists. When this limit exists it is the ***derivative*** of f at a and is denoted by $f'(a)$. □

The limit defining $f'(a)$ can be rewritten in several ways. For example if we do the change of variable $x = a + h$ in the limit it becomes

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

which is the way it is often presented in calculus books. And sometimes, especially in older books, h is replaced by Δx and the limit is written as

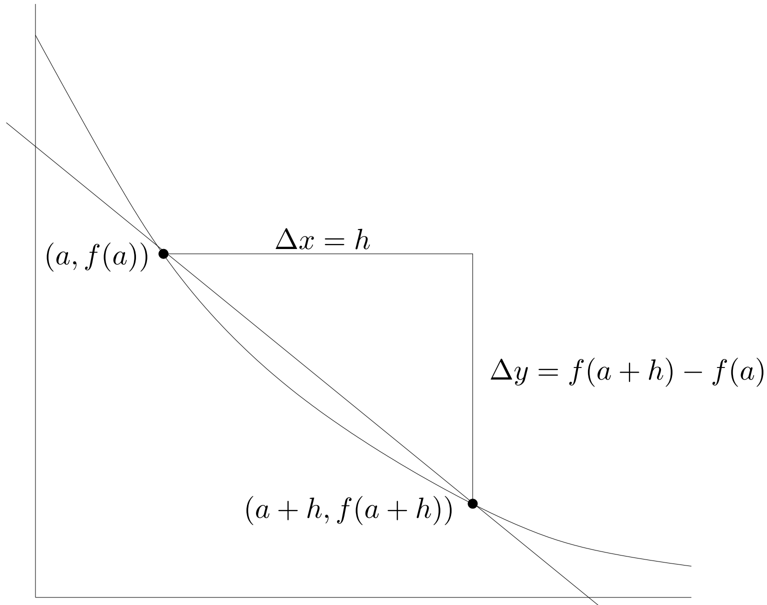
$$f'(a) = \lim_{\Delta x\to 0} \frac{f(a+\Delta x) - f(a)}{\Delta x}.$$

And finally $f(a+\Delta x) - f(a)$ can be abbreviated as $\Delta y = f(a+\Delta x) - f(a)$ and then the limit becoves

$$f'(a) = \lim_{\Delta x\to 0} \frac{\Delta y}{\Delta x}$$

a notation meshes well with the Leibniz notation $\dfrac{dy}{dx}$ for the derivative.

I am now obligated to draw the standard picture that shows that the ***difference quotient*** $(f(a+h) - f(a))/h$ is the slope through the points $(a, f(a))$ and $(a+h, f(a+h))$ and therefore taking the limit as $h \to 0$ of this difference quotient is a reasonable definition of the slope of the tangent line to the graph of $y = f(x)$ at the point $(a, f(a))$.



$$\frac{\Delta y}{\Delta x} = \frac{f(a+h) - f(a)}{h} = \text{slope of line}$$
through $(a, f(a))$ and $(a+h, f(a+h))$.

We now do some examples of derivatives that you no doubt already know from calculus.

Let $f(x) = mx + b$ where m and b are constants. Then for any $a \in \mathbb{R}$

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{x \to a} \frac{m(x-a)}{x-a} = m.$$

For a slightly more complicated example consider $f(x) = x^2$ we have , using that $x^2 - a^2 = (x-a)(x+a)$:

$$\begin{aligned} f'(a) &= \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \\ &= \lim_{x \to a} \frac{x^2 - a^2}{x - a} \\ &= \lim_{x \to a} (x + a) \\ &= 2a. \end{aligned}$$

Problem 2.1. Use the identities $x^3 - a^3 = (x-a)(x^2+ax+a^2$ and $x^4 - a^4 = (x-a)(x^3+ax^2+a^2x+a^3$ to prove that the functions $f(x) = x^3$ and $g(x) = x^4$ have the derivatives

$$\begin{aligned} f'(a) &= 3a^2 \\ g'(a) &= 4a^3. \end{aligned}$$

□

The classic example of a function that does not have a derivative at a point is the absolute value function $f(x) = |x|$ which does not have a derivative at $x = 0$. Here are some other examples

Problem 2.2. Let $f: \mathbb{R} \to \mathbb{R}$ be the function

$$f(x) = \max\{x, 2 - x\}.$$

Graph $y = f(x)$ and show that it is differentiable at every point other than $x = 1$. What is $f'(a)$ when $a \neq 1$? □

Problem 2.3. Let $f: \mathbb{R} \to \mathbb{R}$ be the function

$$f(x) = \min\{x^2, 1\}$$

Find the points where f is not differentiable and prove your result.

We now start proving the basic rules for derivatives you know from calculus.

Proposition 2.2 (Sum rule for derivatives). *Let f_1 and f_2 be defined on an interval containing the point a and assume that f_1 and f_2 are both differentiable at a. Let c_1 and c_2 be constants. Then the function $g = c_1 f_1 + c_2 f_2$ is differentiable at a and*

$$g'(a) = c_1 f_1'(a) + c_2 f_2'(a).$$

Problem 2.4. Prove this. □

We extend this to sums with more terms:

Proposition 2.3. *Let $f_1, f_2, \ldots, f_n$ be functions defined on an interval containing a and with each f_j differentiable at a. Let $c_1, c_2, \ldots, c_n$ be constants. Then $g = c_1 f_1 + c_2 f_2 + \cdots + c_n f_n$ is differentiable at a and*

$$g'(a) = c_1 f_1'(a) + c_2 f_2'(a) + \cdots + c_n f_n'(a).$$

Proof. This is an easy induction proof. □

Proposition 2.4. *Let f be defined on an interval containing a and assume that f is differentiable at a. Then f is continuous at a.*

Problem 2.5. Prove this. *Hint:* To show that f is continuous at a we need to show $\lim_{x \to a} f(x) = f(a)$. As f is differentiable we know that

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

exists. To use this write $f(x)$ as

$$f(x) = f(a) + \frac{f(x) - f(a)}{x - a}(x - a).$$

Now you can use standard results about limits (no ε, δ needed). □

Proposition 2.5 (Product rule)**.** *Let f and g be defined in an interval containing a and assume they are both differentiable at a. Then the product $p(x) = f(x)g(x)$ is differentiable at a and*

$$p'(a) = f'(a)g(a) + f(a)g'(a).$$

Problem 2.6. Prove this. *Hint:* One way is to do some adding and subtracting in the difference quotient for p:

$$\begin{aligned}\frac{p(x)-p(a)}{x-a} &= \frac{f(x)g(x)-f(a)g(a)}{x-a} \\ &= \frac{f(x)g(x)-f(x)g(a)+f(x)g(a)-f(a)g(a)}{x-a} \\ &= f(x)\frac{g(x)-g(a)}{x-a} + \frac{f(x)-f(a)}{x-a}g(a)\end{aligned}$$

As f is continuous at a (why?) we have $\lim_{x\to a} f(x) = f(a)$. Finishing now should be easy. □

Proposition 2.6. *Let g be defined in an interval containing a and assume g is differentiable at a and $g(a) \neq 0$. Then $h(x) = 1/g$ is differentiable at a and*

$$h'(a) = \frac{-g'(a)}{g(a)^2}.$$

Problem 2.7. Prove this. *Hint:* Write the difference quotient for h as

$$\begin{aligned}\frac{h(x)-h(a)}{x-a} &= \frac{1}{x-a}\left(\frac{1}{g(x)} - \frac{1}{g(a)}\right) \\ &= \frac{1}{x-a}\frac{g(a)-g(x)}{g(x)g(a)} \\ &= \frac{-1}{g(x)g(a)}\frac{g(x)-g(a)}{x-a}\end{aligned}$$

and now it should be easy to take the limit defining $h'(a)$. □

Proposition 2.7 (Quotient rule)**.** *Let f and g be defined on an interval containing a and with f and g differentiable at a. Also assume $g(a) \neq 0$. Then the quotient $q(x) = f(x)/g(x)$ is differentiable at a and*

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

Problem 2.8. Prove this. *Hint:* We have already done most of the work for this. Write q as a product

$$q(x) = f(x)\left(\frac{1}{g(x)}\right)$$

and use Proposition 2.6 and the product rule. □

Proposition 2.8 (The power rule for positive powers)**.** *Let f be defined on an interval containing a and assume that f is differentiable at a. Then for any positive integer n the function $p(x) = f(x)^n$ is differentiable at a and*

$$p'(x) = nf(a)^{n-1}f'(a).$$

In particular letting $f(x) = x$ yields that the when $p(x) = x^n$, then $p'(a) = na^{n-1}$.

Problem 2.9. Prove this. *Hint:* Induction. □

Proposition 2.9.

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

is a polynomial (thus $a_0, a_1, \ldots, a_n$ are constants) then f is differentiable at all points a and

$$f'(a) = na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + 2a_2 x + a_1.$$

Proof. This follows by combining Propositions 2.2 and 2.8. □

Proposition 2.10. *The function $f(x) = \sqrt{x}$ is differentiable on $(0, \infty)$ and*

$$f'(x) = \frac{1}{2\sqrt{x}}.$$

Problem 2.10. Prove this. *Hint:* The calculation

$$\begin{aligned} f(x+h) - f(x) &= \sqrt{x+h} - \sqrt{x} \\ &= \frac{(\sqrt{x+h} - \sqrt{x})(\sqrt{x+h} + \sqrt{x})}{\sqrt{x+h} + \sqrt{x}} \\ &= \frac{(x+h) - x}{\sqrt{x+h} + \sqrt{x}} \\ &= \frac{h}{\sqrt{x+h} + \sqrt{x}} \end{aligned}$$

might be useful. Also useful is that we have shown the square root function is continuous on $(0, \infty)$. □

The next result makes precise that the graph $y = f(x)$ of a function has a tangent line at the point $(a, f(a))$ if and only if the derivative $f'(a)$ exists.

Theorem 2.11. *Let f be a real valued function defined in a neighborhood of a. Then the following are equivalent.*

(a) $f'(a)$ exists.

(b) There is a constant m such that

$$f(x) = f(a) + m(x-a) + \rho(x; a)$$

where $\rho(x; a)$ satisfies

(1) $$\lim_{x \to a} \frac{\rho(x; a)}{x - a} = 0.$$

Before going on with the proof, let us think a bit about what condition (b) of the theorem says. Rewrite (1) as

$$f(x) = f(a) + \left(m + \frac{\rho(x;a)}{x-a}\right)(x-a).$$

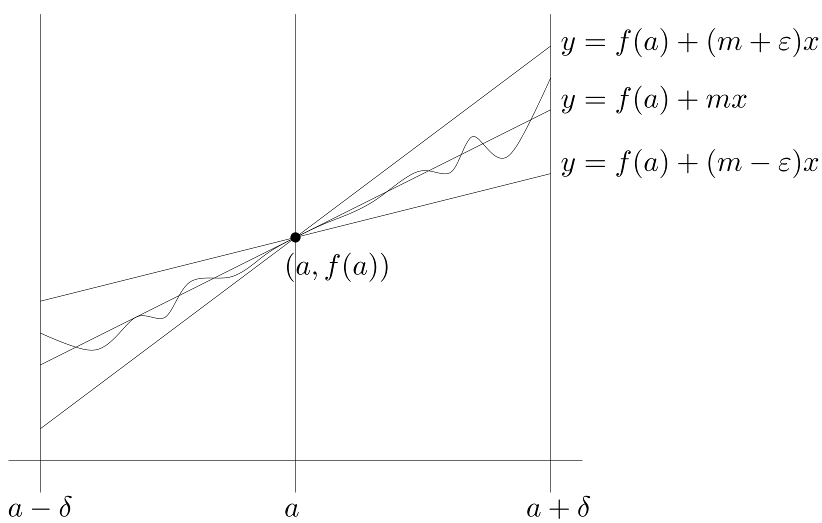
As $\lim_{x\to a} \rho(x;a)/(x-a) = 0$ for any $\varepsilon > 0$ there is a $\delta > 0$ such that

$$0 < |x-a| < \delta \quad \text{implies} \quad \frac{|\rho(x;a)|}{|x-a|} < \varepsilon.$$

Therefore

$$0 < |x-a| < \delta \quad \text{implies} \quad m - \varepsilon < m + \frac{\rho(x:a)}{x-a} < m + \varepsilon.$$

Problem 2.11. Show that these inequalities imply that on the interval $(a-\delta, a+\delta)$ that the graph of $y = f(x)$ stays between the graphs of the lines $y = f(a)+(m+\varepsilon)(x-a)$ and $y = f(a)+(m-\varepsilon)(x-a)$ as in the figure below. Therefore by making ε small we have that near $x = a$ the graph $y = f(x)$ is sandwiched between two line that have slope very close to m. □



Problem 2.12. Prove Theorem 2.11. □

Lemma 2.12. *Let g be differentiable at a. Then there is a $\delta > 0$ such that*

$$|x-a| < \delta \quad \textit{implies} \quad |g(x) - g(a)| \le (|g'(a)| + 1)|x-a|.$$

Problem 2.13. Prove this. □

Theorem 2.13 (The chain rule)**.** *Let g be defined on an interval containing a and f defined and on an interval containing $g(a)$. Assume that g is differentiable at a and f is differentiable at $g(a)$. Then the composition $h = f \circ g$ is differentiable at a and*

$$h'(a) = (f \circ g)'(a) = f'(g(a))g'(a).$$

Proof. As g is differentiable at a, by Theorem 2.11, we have

$$g(x) - g(a) = g'(a)(x-a) + \rho_1(x;a)$$

where

$$\lim_{x\to a} \frac{\rho_1(x;a)}{x-a} = 0.$$

Likewise as f is differentiable at $g(a)$ we have

$$f(y) - f(g(a)) = f'(g(a))(y - g(a))(x-a) + \rho_2(y; g(a))$$

where

$$\lim_{y\to g(a)} \frac{\rho_2(y;g(a))}{y-g(a)} = 0.$$

Therefore we have

$$\begin{aligned} f(g(x)) - f(g(a)) &= f'(g(a))(g(x) - g(a)) + \rho_2(g(x); g(a)) \\ &= f'(g(a))\big(g'(a)(x-a) + \rho_1(x;a)\big) + \rho_2(g(x); g(a)) \\ &= f'(g(a))g'(a)(x-a) + \rho_2(g(x); g(a)) \end{aligned}$$

where

$$\rho(x;a) = f'(g(a))g'(a)\rho_1(x;a) + \rho_2(g(x); g(a)).$$

Therefore, by Theorem 2.11, to finish the proof it is enough to show

$$\lim_{x\to a} \frac{\rho(x:a)}{x-a} = 0.$$

The first term in the definition of is easy to deal with:

$$\begin{aligned} \lim_{x\to a} \frac{f'(g(a))g'(a)\rho_1(x;a)}{x-a} &= f'(g(a))g'(a) \lim_{x\to a} \frac{\rho_1(x;a)}{x-a} \\ &= f'(g(a))g'(a)0 \\ &= 0 \end{aligned}$$

The second term takes a bit more work. Let $\varepsilon > 0$

$$\lim_{y\to g(a)} \frac{\rho_2(y;g(a))}{y-g(a)} = 0$$

there is a $\delta_1 > 0$ such that

$$0 < |y - g(a)| < \delta \quad \text{implies} \quad \left|\frac{\rho_2(y;g(a))}{y-g(a)}\right| < \frac{\varepsilon}{1 + |g'(a)|}$$

and therefore

$$|y - g(a)| < \delta \quad \text{implies} \quad |\rho_2(y; g(a))| \le \frac{\varepsilon |y - g(a)|}{1 + |g'(a)|}.$$

Problem 2.14. If g is differentiable at x show that there is a $\delta_2 > 0$ such that

$$|x - a| < \delta_2 \quad \text{implies} \quad |g(x) - g(a)| \le (|g'(a)| + 1)|x - a|$$

□

Getting back to the proof of the chain rule, there is a if $|g(x) - g(a)| < \delta_1$ and $|x - a| < \delta_2$ then

$$|\rho_2(y; g(a))| \leq \frac{\varepsilon|y - g(a)|}{1 + |g'(a)|} \leq \frac{\varepsilon(1 + |g'(a)|)|x - a|}{1 + |g'(a)|} = \varepsilon|x - a|.$$

and therefore

$$\left|\frac{\rho_2(g(x); g(a))}{x - a}\right| \leq \varepsilon$$

Finally, as g is continuous at a, there is $\delta_3 > 0$ such that

$$|x - a| < \delta_1 \quad \text{implies} \quad |g(x) - g(a)| < \delta_1.$$

Whence if $\delta = \min\{\delta_2, \delta_2\}$,

$$0 < |x - a| < \delta \quad \text{implies} \quad \left|\frac{\rho_2(g(x); g(a))}{x - a}\right| \leq \varepsilon$$

and therefore

$$\lim_{x \to a} \frac{\rho_2(g(x); g(a))}{x - a} = 0$$

which completes the proof. □

2.2. **The derivative of the inverse of a function.**

Theorem 2.14. *Let $f\colon [a, b] \to [c, d]$ be onto, continuous and strictly increasing (or strictly decreasing). Then the inverse $f^{-1}\colon [c, d] \to [a, b]$ is also continuous.*

Problem 2.15. Prove this. *Hint:* (This is mostly going to be a review of results from last semester. At each place where I ask (why?) you should quote a result form last term.) We are given that f is onto. As f is strictly increasing it is also one-to-one. This implies that the inverse f^{-1} exists. One of the ways to shot that function g is continuous is to show that g^{-1}[closed set] is a closed. In our case $g = f^{-1}$ and thus $g^{-1} = (f^{-1})^{-1} = f$. So we only need show that for any closed subset C of $[a, b]$ that $f[C]$ is a closed subset of $[c, d]$. As $[a, b]$ is compact (why?) we have that C is compact (why?). Therefore $f[C]$ is the continuous image of a compact set and thus $f[C]$ is a compact subset of $[c, d]$. Therefore $f[C]$ is compact (why?). Finally this implies that $f[C]$ is closed (why?) which finishes the proof. □

There is a generalization of this that has almost exactly the same proof.

Problem 2.16 (Extra Credit). If $f\colon E \to E'$ is a continuous one-to-one onto (that is f is bijective) between metric spaces with E compact. Then the inverse $f^{-1}\colon E' \to E$ is continuous. *Hint:* To start note that since f is onto we have $f[E] = E'$ and thus E' is the continuous image of a compact set and thus compact. □

Theorem 2.15. *Let $f\colon [\alpha,\beta] \to [c,d]$ be onto, continuous and strictly increasing (or strictly decreasing). Let $x_0 \in (\alpha,\beta)$ and assume that f is differentiable at x_0 with $f'(x_0) \neq 0$. Then f^{-1} is differentiable at $f(x_0)$ and*

$$(f^{-1})'(f(x_0)) = \frac{1}{f'(x_0)}.$$

Proof. Let $\varepsilon > 0$. Then by the definition of the derivative that is a $\delta_1 > 0$ such that

$$0 < |x - x_0| < \delta \quad \text{implies} \quad \left| \frac{f(x) - f(x_0)}{x - a} - f'(x_0) \right| < \varepsilon. \tag{2}$$

By Theorem 2.14 the function f^{-1} is continuous and therefore there is a $\delta > 0$ such that

$$|y - f(x_0)| < \delta \quad \text{implies} \quad |f^{-1}(y) - f^{-1}(f(x_0)| < \delta_1.$$

Let $y_0 := f(x_0)$ and for y with $|y - y_0| < \delta$ let $x = f^{-1}(y)$ in the inequality (2) to get

$$0 < |y - y_0| < \delta \quad \text{implies} \quad \left| \frac{y - y_0}{f^{-1}(y) - f^{-1}(y_0)} - f'(x_0) \right| < \varepsilon.$$

Therefore

$$\lim_{y\to y_0} \frac{y - y_0}{f^{-1}(y) - f^{-1}(y_0)} = f'(x_0) \neq 0.$$

Thus

$$\lim_{y\to\infty} \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \frac{1}{\displaystyle\lim_{y\to y_0} \frac{y - y_0}{f^{-1}(y) - f^{-1}(y_0)}} = \frac{1}{f'(x_0)}.$$

This shows $(f^{-1})'(y_0)$ exists and equals $1/f'(x_0)$, which is just what we were to prove. □

Proposition 2.16. *Let n be a nonzero integer. Then the function $f(x) = x^{\frac{1}{n}}$ is differentiable on $(0,\infty)$ (and if n is odd also on the interval $(-\infty,0)$) and*

$$f'(x) = \frac{1}{n} x^{\frac{1}{n}-1}.$$

Proof. The function f is the inverse of the differentiable function $g(x) = x^n$. Also $g'(x) = nx^{n-1} \neq 0$ for $x \neq 0$. Therefore by Theorem 2.15 we know f is differentiable on $(0,\infty)$. Then

$$f(x)^n = x.$$

Taking the derivative of this gives

$$nf(x)^{n-1} f'(x) = 1$$

so that

$$f'(x) = \frac{1}{n}\frac{1}{f(x)^{n-1})} = \frac{1}{n}\frac{1}{(x^{\frac{1}{n}}}^{\,n-1} = \frac{1}{n} x^{\frac{1}{n}-1}.$$

□

Let $r = p/q \neq 0$ be a rational number with p and q integers. Then for $x > 0$ we can define

$$x^r = x^{\frac{p}{q}} = (x^{\frac{1}{q}})^p.$$

This is the composition of two differentiable functions and therefore is differentiable. Let $f(x) = x^r = (x^{\frac{1}{q}})^p$ with this definition. Then by the chain rule

$$f'(x) = p(x^{\frac{1}{q}})^{p-1}\left(\frac{1}{q}x^{\frac{1}{q}-1}\right) = \frac{p}{q}x^{\frac{p}{q}-1} = rx^{r-1}.$$

We have just proven

Theorem 2.17. *Let r be a rational number, say $r = p/q$ with p and q integers and with $f(x) = x^r$ as defined above. Then f is differentiable on $(0, \infty)$ and*

$$f'(x) = rx^{r-1}.$$

□

2.3. **Functions differentiable on an interval, the first derivative test, and the mean value theorem.**

Definition 2.18. Let f be defined on an open set U containing x_0. Then f has a ***local maximum*** (respectively a ***local minimum***) at x_0 if and only if there is a $\delta > 0$ such that

$$f(x) \leq f(x_0) \quad (\text{respectively } f(x) \geq f(x_0)) \quad \text{for } x \text{ with } |x - x_0| < \delta$$

In this case x_0 is a ***local maximizer*** (respectively a ***local minimizer***) of f. The point x_0 is a ***local extrema*** if it is either a local maximizer or a local minimizer. □

Theorem 2.19 (First Derivative Test)**.** *If f is defined on an open U set containing the point x_0 and*

- *f is differentiable at x_0*
- *f has a local extrema at x_0.*

then

$$f'(x_0) = 0.$$

Lemma 2.20. *Let f be differentiable at x_0 and let $\langle x_n \rangle_{n=1}^{\infty}$ be a sequence with*

$$\lim_{n\to\infty} x_n = x_0 \quad \textit{and for all } n \quad x_n \neq x_0.$$

Then

$$\lim_{n\to\infty} \frac{f(x_n) - f(x_0)}{x_n - x_0} = f'(x_0)$$

Problem 2.17. Prove this. □

Problem 2.18. Prove Theorem 2.19. *Hint:* You do not have to follow this hint, but here is one way to start. Without loss of generality we can assume

f has a local maximum at x_0. (If it has a local minimum, then replace f by $-f$.) Let

$$x_n = x_0 - \frac{1}{n} \quad \text{and} \quad y_n = x_0 + \frac{1}{n}.$$

Then show

$$\lim_{n\to\infty} \frac{f(x_n) - f(x_0)}{x_n - x_0} \geq 0 \quad \text{and} \quad \lim_{n\to\infty} \frac{f(x_n) - f(x_0)}{x_n - x_0} \leq 0$$

and use the lemma. □

Theorem 2.21 (Rôlle's Theorem)**.** *Let f be a function that is continuous on $[a,b]$ and differentable at all points of (a,b). Assume*

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

Then there exists a point $\xi \in (a,b)$ such that

$$f'(\xi) = 0.$$

Problem 2.19. Prove this. *Hint:* Start by showing that either (or both) of the maximum or minimum of f occur in the open interval (a,b). □

Theorem 2.22 (Mean Value Theorem)**.** *Let f be a function that is continuous on $[a,b]$ and differentiable at all points of (a,b). There there exists a point $\xi \in (a,b)$ such that*

$$f(b) - f(a) = f'(\xi)(b-a)$$

Problem 2.20. Prove this. *Hint:* One way to start is to show

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a)$$

satisfies the hypothesis of Rôlle's Theorem. □

Definition 2.23. Let x_1, x_2 and ξ be three real numbers. Then ξ is ***between*** x_1 and x_2 if and only if one of the following three cases holds:

$$\begin{aligned} x_1 <& \xi < x_2 \\ x_2 <& \xi < x_1 \\ x_1 =& \xi = x_2. \end{aligned}$$

□

Often we will use the Mean Value Theorem in the following slightly less general form:

Theorem 2.24 (Mean Value Theorem)**.** *Let f be differentiable on the open interval (a,b) and let $x_1, x_2 \in (a,b)$. There there is ξ between x_1 and x_2 such that*

$$f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1).$$

Proof. If $x_1 = x_2$, then let $\xi = x_1$ and we have $f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1) = 0$. If $x_1 \neq x_2$, then by possibly changing the names of x_1 and x_2 we can assume that $x_1 < x_2$. Then f is continuous on $[x_1, x_2]$ and differentiable on $I(x_1, x_2)$. Therefore we can use our first form of the Mean Value Theorem to conclude there is a $\xi \in (x_1, x_2)$ with $f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1)$. □

Before using the Mean Value Theorem to prove theorems let us note that it can be use to prove interesting results about concrete functions. Here are a couple of examples.

Example 2.25. Assume that we know that the derivative of $\sin(x)$ is $\cos(x)$. Then for all $a, b \in \mathbb{R}$ we have

$$|\sin(b) - \sin(a)| \leq |b - a|.$$

To see this let $f(x) = \sin(x)$. Then the Mean Value Theorem tells us there is a ξ between b and a such that

$$|\sin(b) - \sin(a)| = |f(b) - f(a)| = |f'(\xi)(b - a)| = |\cos(\xi)(b - a)| \leq |b - a|$$

where at the last step we used that $|\cos(\xi)| \leq 1$. □

Example 2.26. If $a, b \geq 2$, then

$$\left| \frac{a-1}{a+1} - \frac{b-1}{b+1} \right| \leq \frac{2}{9}|b - a|.$$

To see this let

$$f(x) = \frac{x-1}{x+1}.$$

Then if $\xi \geq 2$ we have

$$f'(\xi) = \frac{2}{(\xi+1)^2} \leq \frac{2}{(2+1)^2} = \frac{2}{9}.$$

Thus if $a, b \geq 2$ the Mean Value Theorem gives us a ξ between a and b (and therefore $\xi \geq 2$ such that

$$\left| \frac{a-1}{a+1} - \frac{b-1}{b+1} \right| = |f(b) - f(a)| = |f'(\xi)(b - a)| = \frac{2}{(\xi+1)^2}|b - a| \leq \frac{2}{9}|b - a|.$$ □

Problem 2.21. Use the ideas above to show the following

(a) For all $x, y \in \mathbb{R}$ the inequality

$$|\cos(4y) - \cos(4x)| \leq 4|y - x|.$$

(b) If $a, b > 1$ then

$$|\sqrt{b^2 - 1} - \sqrt{a^2 - 1}| \geq |b - a|.$$

(c) If $x > 0$ then

$$e^x - 1 > x.$$

Hint: $e^x - 1 = e^x - e^0$. □

Theorem 2.27. *Let f be differentiable on the open interval (a,b) and assume*

$$f'(x) = 0 \quad \text{for all} \quad x \in (a,b).$$

Then f is constant.

Problem 2.22. Use the Mean Value Theorem to prove this. □

Definition 2.28. If f is a function defined on an interval I, then f is ***increasing*** if and only if for all $x_1, x_2 \in I$

$$x_1 < x_2 \qquad \Longrightarrow \qquad f(x_1) < f(x_2).$$

Theorem 2.29. *Let f be a function on the open interval and assume that f' exists on all of (a,b) and that $f'(x) > 0$ for all $x \in (a,b)$. Then f is increasing on (a,b).*

Problem 2.23. Use the Mean Value Theorem to prove this. □

Problem 2.24. Show that $f'(x_0)$ exists if and only if the limit

$$\lim_{h \to 0} \frac{f(x_0) - f(x_0 - h)}{h}$$

exists. When this limit exists what is its value? □

Problem 2.25. Show that if $f'(x_0)$ exists, then so does the limit

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0 - h)}{2h}$$

and its value is $f'(x_0)$. □

Problem 2.26. Let α be a positive real number and set

$$f(x) = \begin{cases} |x|^\alpha \cos(1/x), & x \neq 0; \\ 0, & x = 0. \end{cases}$$

For what values of α does $f'(0)$ exist. When it doses exist what is its value? □

Theorem 2.30 (Cauchy Mean Value Theorem)**.** *Let f and g be functions that are differentiable on the open interval (a,b) and continuous on the closed interval $[a,b]$. Then there is a $\xi \in (a,b)$ such that*

$$g'(\xi)\big(f(b) - f(a)\big) = f'(\xi)\big(g(b) - g(a)\big).$$

(Note when g is the function $g(x) = x$ this reduces to the usual mean value theorem.

Problem 2.27. Prove this. *Hint:* Let

$$h(x) = (g(b) - g(a))(f(x) - f(a)) - (f(b) - f(a))(g(x) - g(a))$$

and show $h(a) = h(b) = 0$. □

We now wish to look at one of the other standard topics in differential calculus, l'hôpital's rule. Recall this involves evaluating limits of the type

$$\lim_{x\to x_0} \frac{f(x)}{g(x)}$$

where $f(x_0) = g(x_0) = 0$ which leads to

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

which, at least formally, does not make sense. Here is the basic result.

Theorem 2.31 (L'hôpital's rule)**.** *Let f and g be differentiable in a neighborhood of x_0 with $g'(x)] \neq 0$ for $x \neq x_0$. Assume that $f(x_0) = g(x_0) = 0$ and that*

$$\lim_{x\to x_0} \frac{f'(x)}{g'(x)} = L$$

exists. Then $\lim_{x\to x_0} f(x)/g(x)$ exists and is given by

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = L$$

This is usually stated informally as that if $f(x_0) = g(x_0) = 0$ then

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

The important part is that the existence of the limit $\lim_{x\to x_0} \frac{f'(x)}{g'(x)}$ implies the existence of the limit $\lim_{x\to x_0} \frac{f(x)}{g(x)}$.

Problem 2.28. Prove Theorem 2.31 as follows. Let $\varepsilon > 0$ then as $\lim_{x\to x_0} \frac{f'(x)}{g'(x)} = L$ there is a $\delta > 0$ such that

$$0 < |x - x_0| < \delta \quad \implies \quad \left|\frac{f'(x)}{g'(x)} - L\right| < \varepsilon.$$

Let x be so that $0 < |x - x_0| < \delta$. Then, by the Cauchy Mean Value Theorem, there is a ξ between x and x_0 such that

$$\frac{f(x)}{g(x)} = \frac{f(x) - 0}{g(x) - 0} = \frac{f(x) - f(x_0)}{g(x) - g(x_0)} = \frac{f'(\xi)}{g'(\xi)}.$$

Use this to show

$$0 < |x - x_0| < \delta \quad \implies \quad \left|\frac{f(x)}{g(x)} - L\right| < \varepsilon$$

and thus $\lim_{x\to x_0} \frac{f(x)}{g(x)} = L$. (A main point is that $0 < |\xi - x_0| < \delta$, so be sure to explain why this holds.) □

Here is a standard application of l'hôpital's rule:

$$\lim_{x\to 0}\frac{\sin(2x)}{3x} = \lim_{x\to 0}\frac{\sin(2x)'}{(3x)'} = \lim_{x\to 0}\frac{2\cos(2x)}{3} = \frac{2\cos(0)}{3} = \frac{2}{3}.$$

It can also be applied several times in a row:

$$\begin{aligned}
\lim_{x\to 0}\frac{1-\cos(x)}{x^2} &= \lim_{x\to 0}\frac{\sin(x)}{2x} && \left(\text{take } \frac{d}{dx} \text{ of top and bottom}\right)\\
&= \lim_{x\to 0}\frac{\cos(x)}{2} && \left(\text{take } \frac{d}{dx} \text{ of top and bottom}\right)\\
&= \frac{\cos(0)}{2}\\
&= \frac{1}{2}.
\end{aligned}$$

So we have shown $\lim_{x\to 0}\frac{1-\cos(x)}{x^2} = \frac{1}{2}$. Note that in terms of showing this limit exists, this should be read from the bottom up. That is l'hôpital's rule shows that $\lim_{x\to 0}\frac{\sin(x)}{2x}$ exists as $\lim_{x\to 0} = \frac{\sin(x)'}{(2x)'} = \lim_{x\to 0}\frac{\cos(0)}{2} = \frac{1}{2}$ exists. Then anther application of l'hôpital's rule shows that $\lim_{x\to 0}\frac{1-\cos(x)}{x^2} = \lim_{x\to 0}\frac{(1-\cos(x))'}{(x^2)'} = \lim_{x\to 0}\frac{\sin(x)}{2x}$ exists.

Problem 2.29. Here are some problems to practice the use of l'hôpital's rule. Compute the following

(a) $\lim_{x\to 0}\frac{\sin(x)-x}{x^3}$
(b) $\lim_{x\to 0}\frac{e^x-e^{-x}}{x}$
(c) $\lim_{\theta\to \pi}\frac{\sin^3(x)}{x(\cos(x)+1)}$

□

2.4. **Generalized Rôlle's Theorem and Taylor's Theorem.** Now back to Rôlle's theorem. First a definition.

Definition 2.32. Let f be defined on an open integral I. Then f is ***twice differentiable*** on I if f' exists at all points of I and the function f' is differentiable on I. We denote the derivative of f' as f'' or $f^{(2)}$ and it is called the ***second derivative*** of f. If f'' exists at all points of I and f'' is differentiable on I its derivative is denoted by f''' or $f^{(3)}$ and is called the ***third derivative*** of f and f is said to be ***three times differentiable***. Continuing recursively, if we have defined what it means for f to be n ***times differentiable*** on I and the n***-th derivative***, $f^{(n)}$, is differentiable on I then the derivative of $f^{(n)}$ is denoted by $f^{(n+1)}$ and f is $(n+1)$ ***times differentiable*** on I. □

Remark 2.33. For consistency's sake we set $f^{(0)} = f$ and $f^{(1)} = f'$ □

Problem 2.30. Show that the function f on $\mathbb{R}$ defined by

$$f(x) = \begin{cases} x^2, & x \geq 0; \\ -x^2, & x < 0. \end{cases}$$

is differentiable on $\mathbb{R}$ but not twice differentiable. *Hint:* Show $f'(x) = 2|x|$. You may have to use the limit definition to compute $f'(0)$. □

Problem 2.31. Find a function that is twice differentiable on $\mathbb{R}$ but not three times differentiable. More generally can you give an example of a function that is n times differentiable, but not $n+1$ times differentiable. □

Proposition 2.34. *Let I be an open interval and assume f is twice differentiable on I. Let $x_0, x_1 \in I$ with $x_0 \neq x_1$. Assume $f(x_0) = f'(x_0) = 0$ and $f(x_1) = 0$. Then there is a point ξ between x_0 and x_1 with $f''(\xi) = 0$.*

Proof. As $f(x_0) = f(x_1) = 0$ by Rôlle's Theorem there is a ξ_1 between x_0 and x_1 with $f'(\xi_1) = 0$. But the function f' is differentiable on I and $f'(x_0) = f'(\xi_1) = 0$ and thus anther application of Rôlle's Theorem gives us a ξ between x_0 and ξ_1 with $f''(\xi) = (f')'(\xi) = 0$. As ξ_1 is between x_0 and x_1 and ξ is between x_0 and ξ_1 we have that ξ is between x_0 and x_1. □

This generalizes

Theorem 2.35 (Generalized Rôlle's Theorem). *Let f be $n+1$ times differentiable on the open interval I. Let $x_0, x_1 \in I$ with $x_0 \neq x_1$. Assume that*

- $f(x_0) = f'(x_0) = \cdots = f^{(n)}(x_0) = 0$,
- $f(x_1) = 0$.

Then there is a point ξ between x_0 and x_1 with

$$f^{(n+1)}(\xi) = 0.$$

□

Problem 2.32. Prove this. *Hint:* There are several ways to do this. One is to look at the proof of Proposition 2.34 and meditate upon induction. □

Proposition 2.36. *Let f be twice differentiable on the open interval I and let $a, b \in I$ with $a \neq b$. Then there is a ξ between a and b with*

$$f(b) = f(a) + f'(a)(b-a) + \frac{f''(\xi)}{2}(b-a)^2.$$

Proof. Let h be defined on I by

$$h(x) = f(x) - f(a) - f'(a)(x-a) - c(x-a)^2$$

where c is a constant to be chosen shortly. Note

$$h(a) = 0$$

and

$$h'(x) = f'(x) - f'(a) - 2c(x-a),$$

and thus

$$h'(a) = 0.$$

With applying Theorem 2.35 in mind, we choose c so that $h(b) = 0$. That is

$$h(b) = f(b) - f(a) - f'(a)(b-a) - c(b-a)^2 = 0$$

which leads to

$$c = \frac{f(b) - f(a) - f'(a)(b-a)}{(b-a)^2}.$$

With this choice of c we have $h(a) = h'(a) = h(b) = 0$ and thus by Theorem 2.35 there is a ξ between a and b with

$$h''(\xi) = 0.$$

By direct calculation

$$h''(x) = f''(x) - 2c.$$

Then $h''(\xi) = 0$ yields

$$f''(\xi) - 2c = 0.$$

But using the formula for c above we find

$$f''(\xi) - 2\left(\frac{f(b) - f(a) - f'(a)(b-a)}{(b-a)^2}\right) = 0$$

which can be rearranged to give

$$f(b) = f(a) + f'(a)(b-a) + \frac{f''(\xi)}{2}(b-a)^2$$

as required. □

As this was a more or less direct consequence of Proposition 2.34 it makes sense to look for a generalization that depends on Theorem 2.35. To make life a little easier on ourselves we first do the case of $n = 4$.

Lemma 2.37. *Let f be a function that is four times differentiable on an open interval I and let $a \in I$. Let $T(x)$ be the polynomial*
(3)
$$T(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \frac{f^{(4)}(a)}{4!}(x-a)^4,$$

and set

$$g(x) = f(x) - T(x).$$

Then

$$g(a) = g'(a) = g''(a) = g^{(3)}(a) = g^{(4)}(a) = 0.$$

Problem 2.33. Prove this. □

Theorem 2.38. *Let f be five times differentiable on the open interval I and $a, b \in I$ with $a \neq b$. Then there is a ξ between a and b such that*

$$f(b) = f(a) + f'(a)(b-a) + \frac{f''(a)}{2}(b-a)^2 + \frac{f^{(3)}(a)}{3!}(b-a)^3 + \frac{f^{(4)}(a)}{4!}(b-a)^4 + \frac{f^{(5)}(\xi)}{5!}(b-a)^5.$$

Or in different notation let $T(x)$ be the polynomial (3)*, then this is*

$$f(b) = T(b) + \frac{f^{(5)}(\xi)}{5!}(b-a)^5.$$

Problem 2.34. Prove this. *Hint:* Let

$$h(x) = f(x) - T(x) - c(x-a)^5$$

where we choose c so that

$$h(b) = 0.$$

Show $h(a) = h'(a) = h''(a) = h^{(3)}(a) = h^{(4)}(a) = 0$. Now use Theorem 2.35 and now proceed as in the proof of Proposition 2.36. □

Definition 2.39. Let f be n times differentiable on a neighborhood of a. Then the ***degree*** n ***Taylor polynomial*** of f at x is

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

□

Problem 2.35. Show that if f is n times differentiable on an open interval I and T_n is its degree n Taylor polynomial at a, then for $0 \leq k \leq n$

$$T_n^{(k)}(a) = f^{(k)}(a).$$

That is the k-th derivatives of T_n and f agree at a for $0 \leq k \leq n$. □

Theorem 2.40 (Taylor's Theorem with Lagrange's form of the remainder)**.** *Let f be $(n+1)$ times differentiable on the open interval I and let $a, b \in I$ with $a \neq b$. Let T_n be the degree n Taylor polynomial of f at a. Then there is a ξ between a and b such that*

$$f(b) = T_n(b) + f^{(n+1)}(\xi)\frac{(b-a)^{n+1}}{(n+1)!}.$$

*(The term $E_n(b) = f^{(n+1)}(\xi)\frac{(b-a)^{n+1}}{(n+1)!} = f(b) - T_n(b)$ is the **error term** or **remainder term** when approximating f by its Taylor polynomial T_n.)*

Problem 2.36. Prove this. □

We restate this with slightly different notation (just replacing a and b with x_0 and x.)

Theorem 2.41 (Taylor's Theorem with Lagrange's form of the remainder, form 2)**.** *Let f be $(n+1)$ times differentiable on the open interval I and let $x, x_0 \in I$ with $x \neq x_0$. There there is a ξ betweem x and x_0 such that*

$$f(x) = \sum_{k=0}^{n} f^{(k)}(x_0)\frac{(x-x_0)^k}{k!} + f^{(n+1)}(\xi)\frac{(x-x_0)^{n+1}}{(n+1)!}.$$

□

Remark 2.42. In the case that $n = 0$ this becomes

$$f(x) = f(x_0) + f'(\xi)(x - x_0),$$

which can be rewritten as $f(x) - f(x_0) = f'(\xi)(x - x_0)$. That is for $n = 0$ we just get the mean value theorem. □

One last restatement of Taylor's theorem. If we let $x = x_0 + h$ we get

$$f(x_0 + h) = \sum_{k=0}^{n} f^{(k)}(x_0)\frac{h^k}{k!} + f^{(n+1)}(\xi)\frac{h^{n+1}}{(n+1)!}$$

where ξ is between x_0 and $x_0 + h$.

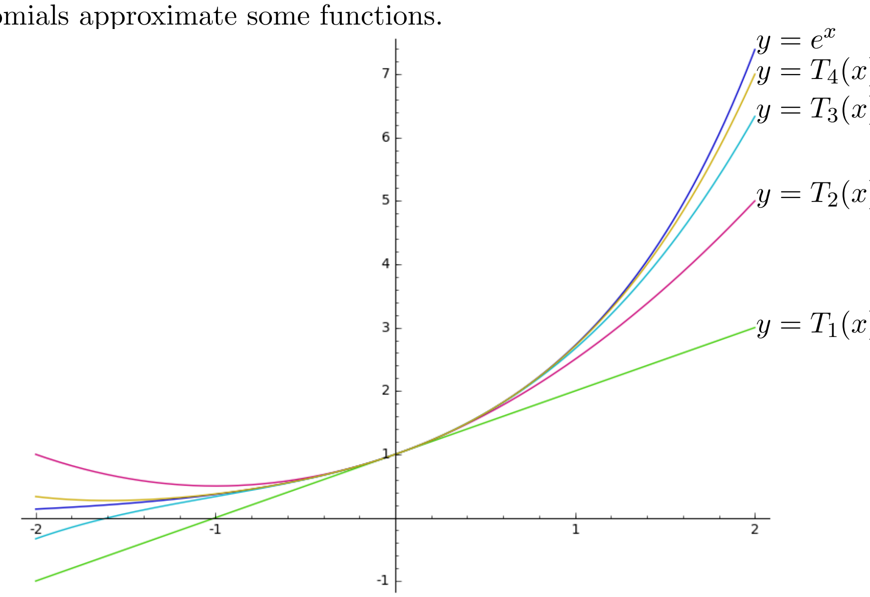
As an examples of Taylor's theorem we have

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{e^\xi x^4}{4!} \qquad \text{(Used } n = 3.)$$

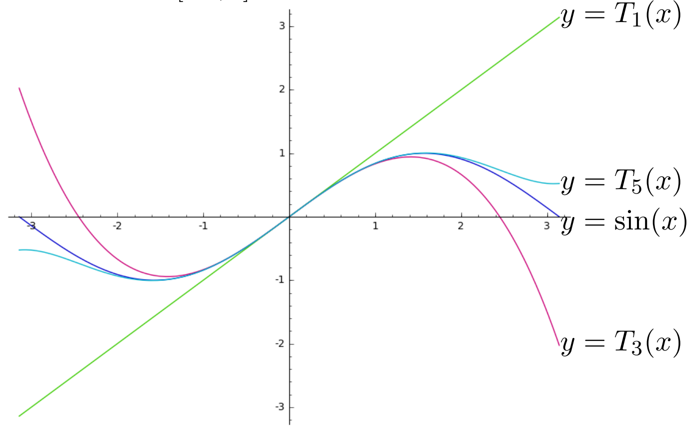
$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{\cos(\xi)x^6}{6!} \qquad \text{(Used } n = 5.)$$

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{\cos(\xi)x^7}{7!} \qquad \text{(Used } n = 6.)$$

where ξ is between x and 0 (and of course the value of ξ is different in each of the three equations). Here are some graphs that show how closely Taylor polynomials approximate some functions.



Graphs of $y = e^x$ and the Taylor polynomials $y = T_n(x)$ for $k = 1, 2, 3, 4$ on the interval $[-2, 2]$



Graphs of $y = \sin(x)$ and the Taylor polynomials $y = T_1(x)$, $y = T_2(x) = T_3(x)$, and $y = T_4(x) = T_5(x)$. on the interval $[-\pi, \pi]$. (If $f(x) = \sin(x)$, then for odd positive integers n we have $f^n(0) = 0$. Therefore $T_{2k}(x) = T_{2k+1}(x)$.)

2.5. **Some applications of Taylor's Theorem.** The next result is just the second derivative test from calculus.

Theorem 2.43. *Let I be an open interval and $f\colon I \to \mathbb{R}$ twice differentiable with f'' continuous. Assume that $f'(x_0) = 0$ then,*

- *if $f''(x_0) < 0$ then x_0 is a local maximizer of f.*
- *if $f''(x_0) > 0$ then x_0 is a local minimizer of f.*

Problem 2.37. Prove this. *Hint:* It is enough to prove in the case $f''(x_0) > 0$.

(a) Use that f'' is continuous to show that there is $\delta > 0$ such that $f'' > 0$ on the interval $(x_0 - \delta, x_0 + \delta)$.
(b) Now use Taylor's Theorem to show for $x \in (x_0 - \delta, x_0 + \delta)$ that

$$f(x) \geq f(x_0)$$

and that equality holds if and only if $x = x_0$. □

If we know that the second derivative has the same sign on an entire interval we can get a global maximum or minimum.

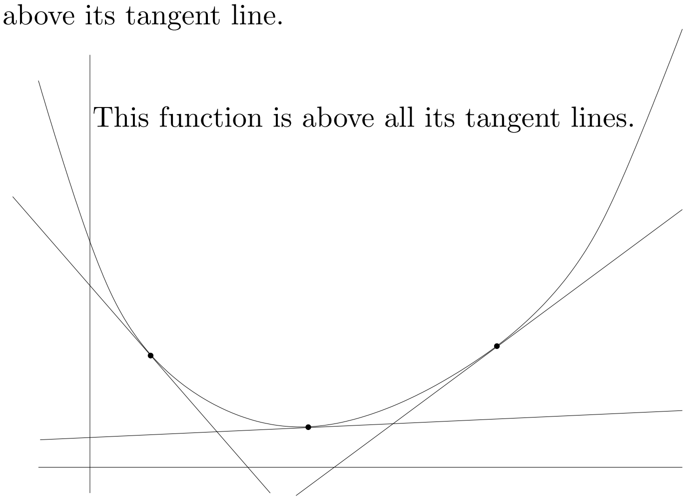
Theorem 2.44. *Let I be an interval and $x_0 \in I$ in the interior of I. Assume that f'' exists and $f'' \geq 0$ at interior points of I. Then $f'(x_0) = 0$ implies that*

$$f(x) \geq f(x_0)$$

for all $x \in I$. (And if $f'' < 0$ on the interior of I, then $f(x) \leq f(x_0)$ on I.)

Problem 2.38. Prove this in the case of $f'' > 0$. □

The last result can be generalized to giving a condition for a function to always be above its tangent line.



Theorem 2.45. *Let $f'' \geq 0$ on an open interval I. Then the graph of f is above all its tangent lines. More precisely if $a \in I$, then*

$$f(a) + f'(a)(x-a) \leq f(x)$$

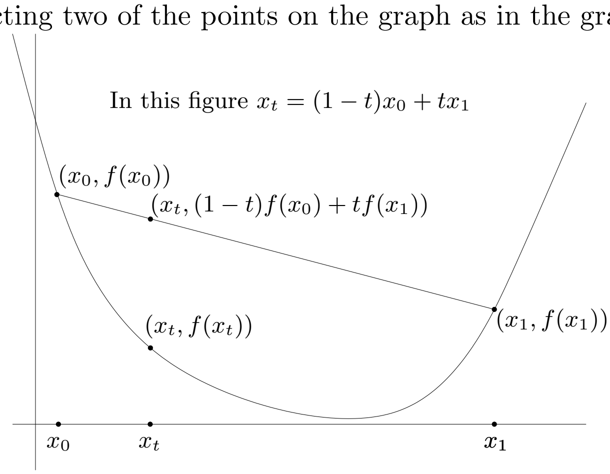
for all $x \in I$.

Problem 2.39. Prove this. □

Let $f\colon I \to \mathbb{R}$ where I is an interval. Then f is ***convex*** if and only if for all $x_0, x_1 \in \mathbb{R}$ and $t \in [0,1]$ the inequality

$$f((1-t)x_0 + tx_1) \leq (1-t)f(x_0) + tf(x_1)$$

holds. Geometrically this means that the graph of f lies below any of the chords connecting two of the points on the graph as in the graph below.



Theorem 2.46. *Let I be an open interval and $f\colon I \to \mathbb{R}$ be a function that is twice differentiable and with $f'' \geq 0$. Then f is convex on I.*

Problem 2.40. Prove this. *Hint:* To simplify notation let

$$x_t = (1-t)x_0 + tx_1.$$

By Theorem 2.45 we have that the graph of f is above its tangent line at x_t, which implies

$$\begin{aligned} f(x_t) + f'(x_t)(x_0 - x_t) &\le f(x_0) \\ f(x_t) + f'(x_t)(x_1 - x_t) &\le f(x_1). \end{aligned}$$

Show

$$\begin{aligned} x_0 - x_t &= -t(x_1 - x_0) \\ x_1 - x_t &= (1-t)(x_1 - x_0). \end{aligned}$$

And therefore

$$\begin{aligned} f(x_t) - tf'(x_t)(x_1 - x_0) &\le f(x_0) \\ f(x_t) + (1-t)f'(x_t)(x_1 - x_0) &\le f(x_1). \end{aligned}$$

Now manipulate these inequalities to complete the proof. □

Proposition 2.47. *Let $x < y$ be real numbers and $\alpha, \beta > 0$ with $\alpha + \beta = 1$. Then the linear combination $\alpha x + \beta y$ is between x and y. That is $x < \alpha x + \beta y < y$.*

Proof. Write $\alpha x + \beta y = x - x + \alpha x + \beta y = x - (1-\alpha)x + \beta y = x - \beta x + \beta y = x + \beta(y - x)$. But $x < y$ so $(y - x) > 0$ and $0 < \beta < 1$ and thus $0 < \beta(y - x) < (y - x)$. There

$$x < \alpha x + \beta y = x + \beta(y - x) < x + (y - x) = y$$

as required. □

Remark 2.48. If we do not make the assumption that $x < y$ we can just say that $\alpha x + \beta y$ is between x and y. That is, when $x \neq y$, we have $\min\{x, y\} < \alpha x + \beta y < \max\{x, y\}$. □

Definition 2.49. Let x, y be real numbers. Then a ***convex combination***, also called a ***weighted average***, of x and y is a linear combination of the form $\alpha x + \beta y$ where $\alpha, \beta \ge 0$ and $\alpha + \beta = 1$.

Thus Proposition 2.47 tells us that the convex combination of two real numbers x and y is between x and y. We can make a more general definition

Definition 2.50. Let $x_1, \ldots, x_n$ be real numbers. Then a ***convex combination*** (and again this is often called a ***weighted average***) of these numbers is a linear combination of the form

$$\alpha_1 x_1 + \cdots + \alpha_n x_n = \sum_{k=1}^{n} \alpha_k x_k$$

where

$$\alpha_1, \ldots, \alpha_n > 0 \quad \text{and} \quad \alpha_1 + \cdots + \alpha_n = \sum_{k=1}^{n} \alpha_k = 1.$$

The following is useful in the induction step of a couple of the proofs below.

Lemma 2.51. *Let $\alpha_1, \ldots, \alpha_{n+1} > 0$ with $\alpha_1 + \cdots + \alpha_{n+1} = 1$. Then for any real numbers $x_1, \ldots, x_{n+1}$ we have*

$$\sum_{k=1}^{n+1} \alpha_k x_k = (1 - \alpha_{n+1}) \sum_{k=1}^{n} \Big(\frac{\alpha_k}{1 - \alpha_{n+1}} \Big) x_k + \alpha_{n+1} x_{n+1}.$$

and

$$\sum_{k=1}^{n} \Big(\frac{\alpha_k}{1 - \alpha_{n+1}} \Big) = 1.$$

Problem 2.41. Prove this. □

Remark 2.52. One way to think about the last lemma is that if x is a convex combination of $x_1, \ldots, x_{n+1}$, then x can be written as

$$x = \alpha y + \beta x_{n+1}$$

where $\alpha = 1 - \alpha_{n+1} > 0$, $\beta = \alpha_{n+1} > 0$ (so that $\alpha + \beta = 1$) and y is a convex combination of $x_1, \ldots, x_n$. This is exactly the set up needed for induction proofs. □

Proposition 2.53. *Let x be a convex combination of $x_1, \ldots, x_n$. Then*

$$\min\{x_1, \ldots, x_n\} \le x \le \max\{x_1, \ldots, x_n\}.$$

(The reason that we have "$\le$" rather than "$<$" is to cover the case when $x_1 = x_2 = \cdots = x_n$. In all other cases the inequalities are strict.)

Problem 2.42. Prove this. *Hint:* See Remark 2.48 (for the base case) and Remark 2.52 (for the induction stop).

Definition 2.54. A function f defined on an interval I is ***convex*** iff for all $\alpha, \beta > 0$ with $\alpha + \beta = 1$ and all $x, y \in I$ the inequality

$$f(\alpha x + \beta y) \le \alpha f(x) + \beta f(y) \tag{4}$$

holds. □

Definition 2.55. A function f defined on an interval I is ***strictly convex*** iff for all $\alpha, \beta > 0$ with $\alpha + \beta = 1$ and all $x, y \in I$ with $x \neq y$ the inequality

$$f(\alpha x + \beta y) < \alpha f(x) + \beta f(y)$$

holds. □

Remark 2.56. Anther way to say that f is strictly convex is that equality holds in the inequality (4) if and only if $x = y$. □

Problem 2.43. Show that $f(x) = x$ and $g(x) = |x|$ are convex on $\mathbb{R}$. *Hint:* For the absolute value, use the triangle inequality. □

Next is a basic result about convex functions.

Theorem 2.57 (Jensen's inequality)**.** *If f is convex on the interval I, $x_1, \ldots, x_n \in I$ and $\alpha_1, \ldots, \alpha_n > 0$ with $\alpha_1 + \cdots + \alpha_n = 1$, then*

$$f(\alpha_1 x_1 + \cdots + \alpha_n x_n) \le \alpha_1 f(x_1) + \cdots + \alpha_n f(x_n).$$

If f is strictly convex, then equality holds if and only if $x_1 = x_2 = \cdots = x_n$.

Problem 2.44. Prove this. *Hint:* See the hint to Problem 2.42. □

If would be nice to have an easily checked criterion that implies that f is convex. You most likely recall from calculus that a function is concave up, that is convex, if its second derivative is positive. As a first step in toward proving this we have

Proposition 2.58. *Let f be twice differentiable on the open interval I with $f''(x) \ge 0$ for all $x \in I$. Then for any $a \in I$*

$$f(x) \ge f(a) + f'(a)(x - a) \tag{5}$$

for all $x \in I$. If the stronger condition $f''(x) > 0$ holds for all $x \in I$ then equality holds in (5) *if and only if $x = a$.*

Proof. This is a straightforward application of Taylor's theorem. From Taylor's theorem with Lagrange's form of the remainder we have

$$f(x) = f(a) + f'(a)(x-a) + f''(\xi)\frac{(x-a)^2}{2} \ge f(a) + f'(a)(x-a)$$

as $f''(\xi)\dfrac{(x-a)^2}{2} \ge 0$ because $(x-a)^2 \ge 0$ and we are assuming $f'' \ge 0$. If $f'' > 0$ then equality can only hold if $x = a$. □

Recall that $y = f(a) + f'(a)(x - a)$ is the equation of the tangent line to the graph of $y = f(x)$ at the point $(a, f(a))$. Therefore Proposition 2.58 tells us that if $f'' \ge 0$, then the graph of $y = f(x)$ lies above all its tangent lines. See Figure 1.

Theorem 2.59. *Let f be twice differentiable on the open interval I and with $f'' \ge 0$ on I. Then f is convex on I. If $f''(x) > 0$ for all $x \in I$, then f is strictly convex.*

Problem 2.45. Prove this. *Hint:* Let $x, y \in I$. If $x = y$ there is nothing to prove (as the inequality (4) reduces to $f(x) = f(x)$). So assume $x \ne y$. Let $\alpha, \beta > 0$ with $\alpha + \beta = 1$ and set

$$a = \alpha x + \beta y.$$

Then we wish to show

$$f(a) \le \alpha f(x) + \beta f(y). \tag{6}$$

From Proposition 2.58 we know

$$f(x) \ge f(a) + f'(a)(x-a), \qquad f(y) \ge f(a) + f'(a)(y-a).$$

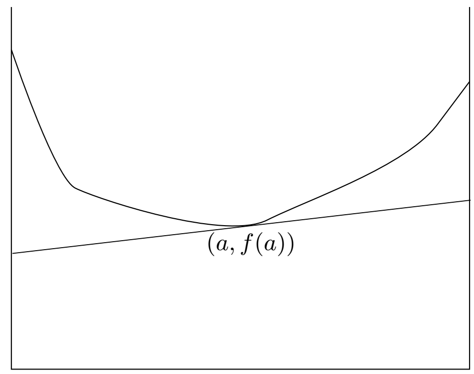


FIGURE 1. If $f'' \geq 0$, then the second order Taylor's theorem tells us

$$\begin{aligned} f(x) &= f(a) + f'(a)(x-a) + f''(\xi)\frac{(x-a)^2}{2} \\ &\geq f(a) + f'(a)(x-a) \end{aligned}$$

As $y = f(a) + f'(a)(x-a)$ is the equation of the tangent line to the graph of $y = f(x)$ at $(a, f(a))$ the graph of f is lies above all of its tangent lines. If $f''(\xi) > 0$ then equality can only if $x = a$, that is the graph $y = f(x)$ is strictly about the tangent line except at the point of tangency.

Multiply the first of these by α and the second by β and add to get an inequality for $\alpha f(x) + \beta f(y)$ and show that this simplifies to (6). Then show if $f'' > 0$ that this inequality is strict. □

It is now easy to check (just by computing the second derivative and noting it is positive) the following

Proposition 2.60. *The following are strictly convex on the indicted intervals.*

(a) $f(x) = x^n$ *where* n *is an integer with* $n \geq 2$ *and* $I = (0, \infty)$.
(b) $f(x) = e^x$ *on* $I = \mathbb{R}$.
(c) $f(x) = -\ln(x)$ *on* $I = (0, \infty)$.
(d) $f(x) = x^{2n}$ *where* $n \geq 1$ *is an integer on* $I = \mathbb{R}$. *(Showing this is strictly convex takes a bit of work.)* □

We recall the ***Arithmetic-Geometric mean inequality***. This is that if a, b are positive real numbers, then

$$\sqrt{ab} \leq \frac{a+b}{2} \tag{7}$$

and equality holds if and only if $a = b$. The proof is simple

$$\frac{a+b}{2} - \sqrt{ab} = \frac{a - 2\sqrt{a}\sqrt{b} + b}{2} = \frac{\left(\sqrt{a} - \sqrt{b}\right)^2}{2} \geq 0$$

and equality can only hold if $\sqrt{a} = \sqrt{b}$. That is if only if $a = b$. The number $\sqrt{ab}$ is the ***geometric mean*** of a and b, while $\frac{a+b}{2}$ is the ***arithmetic mean*** of a and b, which is where the inequality gets its name. It can be greatly generalized.

Theorem 2.61 (Generalized Arithmetic-Geometric Mean Inequality). *Let $\alpha_1, \ldots, \alpha_n > 0$ with $\alpha_1 + \cdots + \alpha_n = 1$. Then for any positive real numbers $a_1, \ldots, a_n$ the inequality*

$$a_1^{\alpha_1} a_2^{\alpha_2} \cdots a_n^{\alpha_n} \le \alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_n a_n$$

holds. Equality holds if and only if all the a_j's are equality.

Problem 2.46. Prove this. *Hint:* We know that the function $f(x) = e^x$ is strictly convex on $\mathbb{R}$. That is for any real numbers $x_1, \ldots, x_n$ we have

$$f(\alpha_1 x_1 + \cdots \alpha_n x_n) \le \alpha_1 f(x_1) + \cdots + \alpha_n f(x_n)$$

and equality holds if and only if all the x_j's are equal. Show this can be rewritten as

$$(e^{x_2})^{\alpha_1} (e^{x_2})^{\alpha_2} \cdots (e^{x_n})^{\alpha_n} \le \alpha_1 e^{x_1} + \alpha_2 e^{x_2} + \cdots + \alpha_n e^{x_n}$$

and equality holds if and only if all the x_j's are equal.

Now given positive numbers $a_1, \ldots, a_n$ there are unique real numbers $x_1, \ldots, x_n$ with $a_j = e^{x_j}$ for all $j = 1, 2, \ldots, n$. (You can assume these x_j's exist.) And you take it from here. □

Remark 2.62. In different notation the generalized Arithmetic-Geometric inequality is

$$\prod_{k=1}^{n} a_k^{\alpha_k} \le \sum_{k=1}^{n} \alpha_k a_x$$

with equality holding if and only if all the a_k's are equal. □

The can you may have seen before is

$$\sqrt[n]{a_1 a_2 \cdots a_n} \le \frac{a_1 + \cdots + a_n}{n}$$

coming form $\alpha_1 = \alpha_2 = \cdots = a_n = 1/n$ and equality holds if and only if all the a_j's are equal. The can of $n = 2$ is often useful. Then letting $\alpha = \alpha_1$ and $\beta = \alpha_2$ we have

$$a^\alpha b^\beta \le \alpha a + \beta b$$

with equality holding if and only if $a = b$. (And as usual $\alpha, \beta > 0$ with $\alpha + \beta = 1$.)

Here is an example of the use of the generalized arithmetic geometric mean inequality

Example 2.63. For $x, y, z \ge 0$ maximize the product xyz subject to the constraint $x + y + z = c$, where c is a constant. We have

$$xyz = \left((xyz)^{1/3} \right)^3 \le \left(\frac{x+y+z}{3} \right)^3 = \left(\frac{c}{3} \right)^3$$

and equality holds if and only of $x = y = z$. Thus the maximum is $(c/3)^3$ with equality if any only if $x = y = z = c/3$. □

Here is a bit more challenging problem.

Problem 2.47. Show that if $f\colon I \to \mathbb{R}$ is a continuous function on an interval I, such that for all $x, y \in I$ the inequality

$$f\left(\frac{x+y}{2}\right) \le \frac{f(x)+f(y)}{2} \tag{8}$$

holds, then f is convex on I. □

Remark 2.64. A function that satisfies the inequality (8) is called ***midpoint convex***. So the problem is asking you to show that a continuous midpoint convex function is convex. Without the assumption of continuity this is false. That is there are midpoint convex functions that are not convex. However they are not continuous. Examples of such functions are hard to produce and require using some version of the Axiom of Choice. □

3. Riemann Integration

We start with some definitions.

Definition 3.1. Let $[a, b]$ be a closed bounded interval. Then a ***partition*** of $[a, b]$ is a list of points $a = x_0 < x_1 < x_2 < \cdots < x_n = b$. We denote it by $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$. We also use the notation

$$\Delta x_j = x_j - x_{j-1}.$$

(See Figure 2.) □

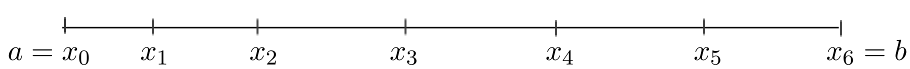


FIGURE 2. A partition of the interval $[a, b]$ into $n = 6$ pieces. The j-th interval $[x_{j-1}, x_j]$ has length $\Delta x_j = x_j - x_{j-1}$.

Definition 3.2. The function φ is a ***step function*** on $[a, b]$ so that there is a partition $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$ of $[a, b]$ and φ is constant on each open interval (x_{j-1}, x_j) for $j = 1, 2, \dots, n$. We denote the set of all step functions on $[a, b]$ by $\mathcal{S}[a, b]$. □

Proposition 3.3. *The set $\mathcal{S}[a, b]$ is a vector space. That if if $\varphi_1, \varphi_2 \in \mathcal{S}[a, b]$ and $c_1, c_2 \in \mathcal{S}[a, b]$, then $c_1\varphi_1 + c_2\varphi_2 \in \mathcal{S}[a, b]$.*

Problem 3.1. Prove this. If φ_1 is a step function with respect to the partition $\mathcal{P}_1$ and φ_2 is a step function with respect to the partition $\mathcal{P}_2$, then show $c_1\varphi_1 + c_2\varphi_2$ is a step function with respect to the partition $\mathcal{P} = \mathcal{P}_1 \cup \mathcal{P}_2$. □

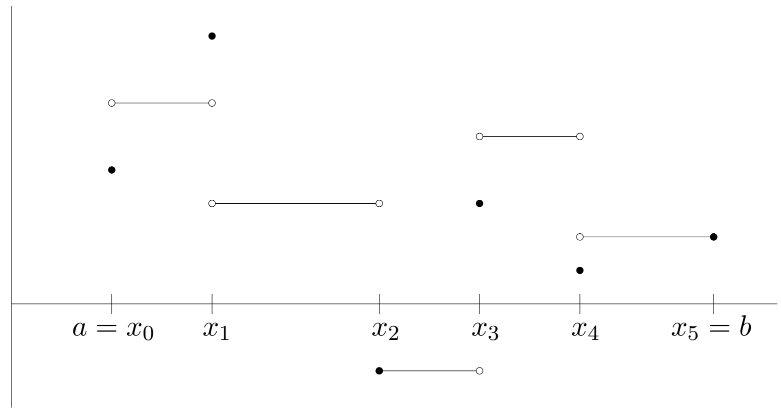


FIGURE 3. A step function, φ, for the interval $[a, b]$ partitioned into five subintervals. By definition φ is constant on each of the open intervals (x_{j-1}, x_j) for $j = 1, 2, 3, 4, 5$. No assumption is made about the values at the points x_j.

Definition 3.4. Let $\varphi \in \mathcal{S}[a, b]$ and let $\mathcal{P} = \{a = x_0, x_1, \ldots, x_n = b\}$ be a partition such that φ has the constant value c_j on the open interval (x_{j-1}, x_j). Then the ***integral*** of φ on the interval $[a, b]$

$$\int_a^b \varphi(x)\,dx = \sum_{j=1}^{n} c_j(x_j - x_{j-1}).$$

□

Setting $\Delta x_j = x_j - x_{j-1}$ (the length of the interval (x_{j-1}, x_j)) this can also be written as

$$\int_a^b \varphi(x)\,dx = \sum_{j=1}^{n} c_j \Delta x.$$

When a all the c_j's are positive, this is just the area under the graph of φ computed by adding up the area rectangles. In the general case (where the c_j's can be negative) this is the area where the area below the graph x-axis counted as negative.

Proposition 3.5. *The integral is linear on $\mathcal{S}[a, b]$. That is if $\varphi_1, \varphi_2 \in \mathcal{S}[a, b]$, and $c_1, c_2 \in \mathbb{R}$, then*

$$\int_a^b (c_1\varphi_1 + c_2\varphi_2)\,dx = c_1 \int_a^b \varphi_1(x)\,dx + c_2 \int_a^b \varphi_2(x)\,dx.$$

□

Problem 3.2. Prove this. *Hint:* If φ_1 is defined by the partition $\mathcal{P}_1$ and φ_2 is defined by the partition $\mathcal{P}_2$ then let $\mathcal{P} = \mathcal{P}_1 \cup \mathcal{P}_2$. If $\mathcal{P}$ is the list $x_0 < x_1 < \cdots < x_n$ then both φ_1 and φ_2 will be constant on each of the open intervals (x_{j-1}, x_j) for $j = 1, 2, \ldots, n$. Thus are constants $a_1, a_2, \ldots, a_n$

and $b_1, b_2, \ldots, b_n$ such that $\varphi_1 = a_j$ and $\varphi_2 = b_j$ on (x_{j-1}, x_j) and

$$\int_a^b \varphi_1(x)\,dx = \sum_{j=1}^n a_j \Delta x_j \qquad \int_a^b \varphi_2(x)\,dx = \sum_{j=1}^n b_j \Delta x_j$$

and the rest should be easy. □

Proposition 3.6. *If $\varphi \in step[a,b]$ and $\varphi(x) \geq 0$ for $x \in [a,b]$, then*

$$\int_a^b \varphi(x)\,dx \geq 0.$$

Problem 3.3. Prove this. □

Corollary 3.7. *If $\varphi, \psi \in \mathcal{S}[a,b]$ and $\varphi \leq \psi$ on $[a,b]$ then*

$$\int_a^b \varphi(x)\,dx \leq \int_a^b \psi(x)\,dx.$$

Problem 3.4. Prove this. □

Definition 3.8. Let $f\colon [a,b] \to \mathbb{R}$ be a bounded function. Then define the ***upper integral*** of f to be

$$\overline{\int_a^b} f(x)\,dx = \inf\left\{ \int_a^b \psi(x)\,dx : \psi \in \mathcal{S}[a,b] \text{ and } f(x) \leq \psi(x) \text{ all } x \in [a,b] \right\}.$$

Likewise the ***lower integral*** of f is

$$\underline{\int_a^b} f(x)\,dx = \sup\left\{ \int_a^b \varphi(x)\,dx : \varphi \in \mathcal{S}[a,b] \text{ and } f(x) \geq \varphi(x) \text{ all } x \in [a,b] \right\}.$$

□

Definition 3.9. The bounded function $f\colon [a,b] \to \mathbb{R}$ is ***Riemann integrable*** if and only if

$$\underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x)\,dx.$$

In this case the ***integral*** of f is the common value of the upper and lower integral

$$\int_a^b f(x)\,dx = \underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x)\,dx.$$

We denote the set of all Riemann integrable functions on $[a,b]$ by $\mathcal{R}[a,b]$. □

Proposition 3.10. *For any bounded $f\colon [a,b] \to \mathbb{R}$ we have*

$$\underline{\int_a^b} f(x)\,dx \leq \overline{\int_a^b} f(x)\,dx$$

Proof. If $\varphi, \psi \in \mathcal{S}[a,b]$ with $\varphi \le f \le \psi$ on $[a,b]$, then

$$\int_a^b \varphi(x)\,dx \le \int_a^b \psi(x)\,dx$$

This holds for all $\varphi \in \mathcal{S}[a,b]$ with $\varphi \le f$, and so $\int_a^b \psi(x)\,dx$ is an upper bound for the set $\{\int_a^b \varphi(x)\,dx : \varphi \in \mathcal{S}[a,b] \text{ and } f(x) \ge \varphi(x) \text{ all } x \in [a,b]\}$ and therefore, by the definition of sup as the least upper bound

$$\underline{\int_a^b} f(x)\,dx \le \int_a^b \psi(x)\,dx.$$

Thus $\underline{\int_a^b} f(x)\,ds$ is a lower bound for $\psi \in \{\int_a^b \psi(x)\,dx : \psi \in \mathcal{S}[a,b] \text{ and } f(x) \le \psi(x) \text{ all } x \in [a,b]\}$ and therefore

$$\underline{\int_a^b} f(x)\,dx \le \overline{\int_a^b} f(x)\,dx$$

as required. □

The following gives a method for showing that a function is Riemann integrable without having to compute the upper and lower integrals.

Theorem 3.11. *The bounded function $f\colon [a,b] \to \mathbb{R}$ is Riemann integrable if and only if for all $\varepsilon > 0$ there is are step functions $\varphi, \psi \in \mathcal{S}[a,b]$ with*

$$\varphi \le f \le \psi$$

on $[a,b]$ and

$$\int_a^b (\psi(x) - \varphi(x))\,dx < \varepsilon.$$

Proof. Let $\varepsilon > 0$. Then there are $\varphi, \psi \in \mathcal{S}[a,b]$ with

$$\varphi(x) \le f(x) \le \psi(x)$$

for $x \in [a,b]$.

$$\int_a^b (\psi(x) - \varphi(x))\,dx < \varepsilon.$$

Then

$$\int_a^b \varphi(x)\,dx \le \underline{\int_a^b} f(x) \le \overline{\int_a^b} f(x)\,dx \le \int_a^b \psi(x)\,dx$$

which implies

$$0 \le \overline{\int} f(x)\,dx - \underline{\int_a^b} f(x)\,dx \le \int_a^b (\psi(x) - \varphi(x))\,dx < \varepsilon.$$

This holds for all $\varepsilon > 0$ and so $\underline{\int_a^b} fx)\,dx = \overline{\int_a^b} f(x)\,dx$, that is $f \in \mathcal{R}R[a,b]$. □

Lemma 3.12. *Let $f, g\colon [a,b] \to \mathbb{R}$ be bounded functions. Then*

$$\overline{\int_a^b} (f(x)+g(x))\,dx \le \overline{\int_a^b} f(x)\,dx + \overline{\int_a^b} g(x)\,dx$$
$$\underline{\int_a^b} (f(x)+g(x))\,dx \ge \underline{\int_a^b} f(x)\,dx + \underline{\int_a^b} g(x)\,dx$$

Problem 3.5. Prove this. *Hint:* It is enough to prove the first of these, as the proof of the second is similar. Let $\varepsilon > 0$ and choose $\varphi_1, \psi_2 \in \mathcal{S}[a,b]$ with $f \le \psi_1$, $g \le \psi_2$ and

$$\int_a^b \psi_1(x)\,dx \le \overline{\int_a^b} f(x)\,dx + \varepsilon/2, \quad \int_a^b \psi_2(x)\,dx \le \overline{\int_a^b} g(x)\,dx + \varepsilon/2,$$

Use this and $f + g \le \psi_1 + \psi_2$ to show

$$\overline{\int_a^b} f(x)\,dx + \overline{\int_a^b} g(x)\,dx \le \overline{\int_a^b} f(x)\,dx + \overline{\int_a^b} g(x) + \varepsilon.$$

Since this holds for all $\varepsilon > 0$ the result follows.

Problem 3.6. This problem shows that equality need not hold in Lemma **??**. Let

$$f(x) = \begin{cases} 1, & x \in \mathbb{Q}; \\ 0, & x \notin \mathbb{Q}. \end{cases} \qquad g(x) = \begin{cases} 0, & x \in \mathbb{Q}; \\ 1, & x \notin \mathbb{Q}. \end{cases}$$

Then $f + q = 1$. Show

$$\overline{\int_0^1} f(x)\,dx = \overline{\int_0^1} g(x)\,dx = \overline{\int_a^b} (f(x)+g(x))\,dx = 1.$$

Thus $\overline{\int_a^b}(f(x)+g(x))\,dx = 1 < 2 = \overline{\int_a^b} f(x)\,dx + \overline{\int_a^b} g(x)\,dx$. □

Lemma 3.13. *If $f, g \in \mathcal{R}[a,b]$ then $f+g \in \mathcal{R}[a,b]$ and*

$$\int_a^b (f(x)+g(x))\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx.$$

Proof. As $f, g \in \mathcal{R}[a,b]$ we have $\overline{\int_a^b} f(x)\,dx = \underline{\int_a^b} f(x)\,dx = \int_a^b f(x)\,dx$ and $\overline{\int_a^b} g(x)\,dx = \underline{\int_a^b} g(x)\,dx = \int_a^b g(x)\,dx$ We now use Lemma 3.12

$$\begin{aligned}
\overline{\int_a^b}(f(x)+g(x))\,dx &\le \overline{\int_a^b} f(x)\,dx + \overline{\int_a^b} g(x)\,dx \\
&= \int_a^b f(x)\,dx + \int_a^b g(x)\,dx \\
&= \underline{\int_a^b} f(x)\,dx + \underline{\int_a^b} g(x)\,dx \\
&\le \underline{\int_a^b}(f(x)+g(x))\,dx \\
&\le \overline{\int_a^b}(f(x)+g(x))\,dx.
\end{aligned}$$

So all the inequalities are equalities and we have

$$\overline{\int_a^b}(f(x)+g(x))\,dx = \underline{\int_a^b}(f(x)+g(x))\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx,$$

as required. □

Lemma 3.14. *If $f \in \mathcal{R}[a,b]$, then $-f \in \mathcal{R}[a,b]$ and*

$$\int_a^b (-f(x))\,dx = -\int_a^b f(x)\,dx$$

Problem 3.7. Prove this. □

Lemma 3.15. *If $f \in \mathcal{R}[a,b]$ and $c \in \mathbb{R}$ then $cf \in \mathcal{R}[a,b]$ and*

$$\int_a^b cf(x)\,dx = c\int_a^b f(x)\,dx.$$

Problem 3.8. Prove this. *Hint:* You may have to consider the cases $c \ge 0$ and $c \le 0$ separately. □

Theorem 3.16. *The set $\mathcal{R}[a,b]$ is a vector space and the integral is linear on $\mathcal{R}[a,b]$.*

Problem 3.9. Use the lemmas and propositions above to prove this. □

If f is a monotone increasing function on $[a,b]$ and $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$ is a partition of $[a,b]$ define two step functions by $\varphi_{f,\mathcal{P}}(b) = f(b)$,

$$\varphi_{f,\mathcal{P}}(x) = f(x_{j-1}) \quad \text{for} \quad x_{j-1} \le x < x_j$$

and $\psi_{f,\mathcal{P}}(b) = f(b)$

$$\psi_{f,\mathcal{P}} = f(x_j) \quad \text{for} \quad x \in x_{j-1} \le x < x_j.$$

See Figure 4

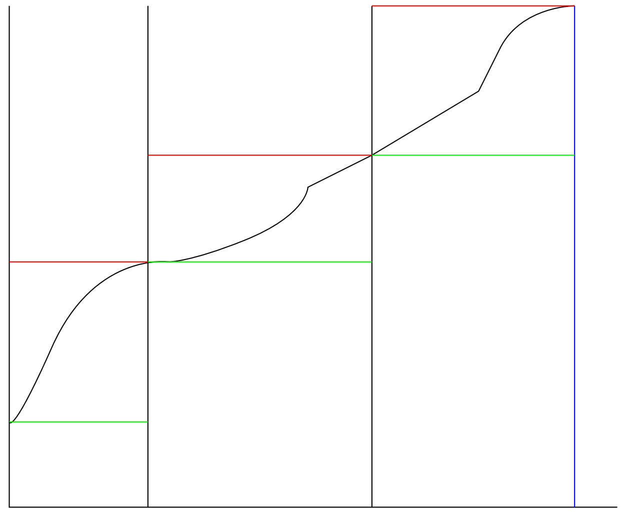


FIGURE 4. A monotone increasing function on $[a,b]$ and a partition, $\mathcal{P}$, with $n = 3$ showing the lower step function $\varphi_{f,\mathcal{P}}$ (in green) and the upper step function $\psi_{f,\mathcal{P}}$ (in red).

Proposition 3.17. *If f is monotone increasing on $[a,b]$ then for any partition, $\mathcal{P}$, of $[a,b]$, with the notation above,*

$$\varphi_{f,\mathcal{P}} \leq f \leq \psi_{f,\mathcal{P}}$$

on $[a,b]$.

Problem 3.10. Prove this. □

Definition 3.18. Given a positive integer n and a closed bounded interval $[a,b]$ the ***uniform partition*** of $[a,b]$ into n sub-intervals is the partition $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$ with

$$x_j = a + j\left(\frac{b-a}{n}\right)$$

for $j = 0, 1, \dots, n$. Note in this case all the lengths, Δx_j of the sub-intervals $[x_{j-1}, x_j]$ have the same value $\Delta x = \Delta x_j = (b-a)/n$. □

Now let us consider the monotone increasing function f on the interval $[a,b]$ with the uniform partition, $\mathcal{P}$, of $[a,b]$ with $n = 4$. Then $\Delta x = \Delta x_j = (b-a)/4$ and $\varphi_{f,\mathcal{P}} \leq f \leq \psi_{f,\mathcal{P}}$. Also

$$\int_a^b \varphi_{f,\mathcal{P}}(x)\,dx = \big(f(x_0) + f(x_1) + f(x_2) + f(x_3)\big)\Delta x$$

and

$$\int_a^b \psi_{f,\mathcal{P}}(x)\,dx = \big(f(x_1) + f(x_2) + f(x_3) + f(x_4)\big)\Delta x.$$

Thus

$$\int_a^b (\psi_{f,\mathcal{P}}(x) - \psi_{f,\mathcal{P}}(x))\, dx = (f(x_4) - f(x_0))\,\Delta x = (f(b) - f(a))\,\Delta x$$

There is nothing special about $n = 4$ in this:

Problem 3.11. Show that if f is monotone increasing on $[a, b]$, n is a positive integer and $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ is the uniform partition of $[a, b]$ into n subintervals, then, with the notation above,

$$\int_a^b (\psi_{f,\mathcal{P}}(x) - \varphi_{f,\mathcal{P}}(x))\, dx = (f(b) - f(a))\,\Delta x = \frac{(f(b) - f(a))(b - a)}{n}. \qquad \square$$

Theorem 3.19. *If f is a monotone function on the closed bounded interval $[a, b]$, then f is integrable on $[a, b]$.*

Problem 3.12. Prove this. *Hint:* With out loss of generality assume f is monotone increasing (if f is monotone decreasing replace f by $-f$). Let $\varepsilon > 0$ and let n be a positive integer such that

$$\frac{(f(b) - f(a))(b - a)}{n} < \varepsilon$$

and use Proposition 3.17 and the last problem. $\square$

Theorem 3.20. *Let f be a continuous function on $[a, b]$. Then f is integrable on $[a, b]$.*

Proof. Let $\varepsilon > 0$. As f is continuous on the closed bounded set $[a, b]$ it is uniformly continuous on $[a, b]$. Thus there is an $\delta > 0$ such that for $x, y \in [a, b]$.

$$|x - y| < \delta \quad \Longrightarrow \quad |f(x) - f(y)| < \frac{\varepsilon}{b - a}.$$

Let n be a positive integer such that

$$\frac{b - a}{n} = \Delta x < \delta$$

and let $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ be the uniform partition of $[a, b]$ into n subintervals. Set

$$m_j = \inf\{f(x) : x \in [x_{j-1}, x_j]\} = \min\{f(x) : x \in [x_{j-1}, x_j]\},$$
$$M_j = \sup\{f(x) : x \in [x_{j-1}, x_j]\} = \max\{f(x) : x \in [x_{j-1}, x_j]\}$$

where the infimum is achieved as a minimum and the supremum is achieved as a maximum because continuous functions on closed bounded sets achieve their maximums and minimums. Define step functions φ and ψ on $[a, b]$ $\varphi(b) = \psi(b) = f(b)$ and

$$\varphi(x) = m_j \quad \text{for} \quad x_{j-1} \le x < x_j$$
$$\psi(x) = M_j \quad \text{for} \quad x_{j-1} \le x < x_j.$$

Then

$$\varphi \le f \le \psi$$

and

$$\int_a^b (\varphi(x) - \psi(x))\,dx = \sum_{j=1}^{n}(M_j - m_j)\left(\frac{b-a}{n}\right).$$

As f is continuous on the closed bounded interval $[x_{j-1}, x_j]$, f achieves its maximum and minimum on this interval. Thus there are $\alpha_j, \beta_j \in [x_{j-1}, x_j]$ with $f(\alpha_j) = m_j$ and $f(\beta_j) = M_j$. But then $|\alpha_j - \beta_j| \le \Delta x < \delta$ and therefore

$$M_j - m_j = |f(\beta_j) - f(\alpha_j)| < \frac{\varepsilon}{b-a}.$$

Thus

$$\int_a^b (\varphi(x) - \psi(x))\,dx = \sum_{j=1}^{n}(M_j - m_j)\left(\frac{b-a}{n}\right) < \sum_{j=1}^{n}\frac{\varepsilon}{b-a}\left(\frac{b-a}{n}\right) = \varepsilon$$

and the result now follows from Theorem 3.11. □

Lemma 3.21. *Let* $\alpha, \beta \in \mathbb{R}$*, then*

$$|\max\{\alpha, 0\} - \max\{\beta, 0\}| \le |\alpha - \beta|.$$

Problem 3.13. Prove this by splitting it into the four cases *(i)* $\alpha, \beta \ge 0$, *(ii)* $\alpha \ge 0, \beta < 0$, *(iii)* $\alpha < 0, \beta \ge 0$, and *(iv)* $\alpha, \beta < 0$. This is not to be handed in. □

Proposition 3.22. *If* $f \in \mathcal{R}[a,b]$ *then so is* $g = \max\{f, 0\}$.

Proof. Let $\varepsilon > 0$ Let φ and ψ be step functions on $[a,b]$ such that $\varphi \le f \le \psi$ and $\int_a^b (\psi - \varphi)\,dx < \varepsilon$. Then

$$\varphi_0 = \max\{0, \varphi\}, \qquad \psi_0 = \max\{0, \psi\}$$

are step functions, $\varphi_0 \le \max\{f, 0\} \le \psi_0$ and $0 \le \psi_0 - \varphi_0 \le \psi - \varphi$. Thus, using Lemma 3.21,

$$\int_a^b (\psi_0 - \varphi_0)\,dx \le \int_a^b (\psi - \varphi)\,dx < \varepsilon$$

and so $\max\{f, 0\}$ is integrable by Theorem 3.11. □

This implies a good deal more because of the following elementary result.

Lemma 3.23. *For real numbers* a, b *the following hold*

$$\begin{aligned} \min\{a, 0\} &= -\max\{-a, 0\}, \\ |a| &= \max\{a, 0\} + \max\{-a, 0\}, \\ \max\{a, b\} &= a + \max\{0, b-a\}, \\ \min\{a, b\} &= a + \min\{0, b-a\}. \end{aligned}$$

Proof. Left to reader (and you don't have to turn these in). We did enough of this type of thing last term that I believe you can do it. □

Proposition 3.24. *If f and g are integrable on $[a,b]$ then so are $|f|$, $\min\{f,g\}$ and $\max\{f,g\}$.*

Proof. This follows easily from Proposition 3.22 and Lemma 3.23. □

Lemma 3.25. *If f is integrable on $[a,b]$ then so is f^2.*

Problem 3.14. Prove this. *Hint:* As $f^2 = |f|^2$ and $|f|$ is also integrable by replacing f by $|f|$ we can assume $f \geq 0$. As f is integrable it is bounded, say $0 \leq f \leq B$ on $[a,b]$. Also as f is integrable on $[a,b]$ for $\varepsilon > 0$ there is are step functions φ, ψ such that

$$\varphi \leq f \leq \psi$$

and

$$\int_a^b (\psi - \varphi)\, dx < \frac{\varepsilon}{2B}.$$

By replacing φ by $\max\{0,\varphi\}$ and ψ by $\min\{\psi, B\}$ we can assume $0 \leq \varphi$ and $\psi \leq B$. Then φ^2 and ψ^2 are step functions and

$$\varphi^2 \leq f^2 \leq \psi^2$$

and

$$0 \leq \psi^2 - \varphi^2 = (\psi + \varphi)(\psi - \varphi) \leq (\psi + \psi)(\psi - \varphi) \leq (B + B)(\psi - \varphi).$$

You should now be able to show

$$\int_a^b (\psi^2 - \varphi^2)\, dx < \varepsilon$$

so that Theorem 3.11 applies. □

Proposition 3.26. *If f and g are integrable on $[a,b]$ then so is the product fg.*

Problem 3.15. Prove this. *Hint:* Show

$$fg = \frac{(f+g)^2 - (f-g)^2}{4}$$

and use Lemma 3.25. □

4. THE FUNDAMENTAL THEOREM OF CALCULUS.

Proposition 4.1. *If $a < b < c$ and f is integrable on $[a,c]$ then the restrictions $f\big|_{[a,b]}$ and $f\big|_{[b,c]}$ are integrable on $[a,b]$ and $[b,c]$ respectively and*

$$\int_a^c f(x)\, dx = \int_a^b f(x)\, dx + \int_b^c f(x)\, dx.$$

Proof. We have shown in class that for any bounded function on $[a,c]$ that

$$\overline{\int_a^c} f(x))\,dx = \overline{\int_a^b} f(x)\,dx + \overline{\int_b^c} f(x)\,dx,$$
$$\underline{\int_a^c} f(x))\,dx = \underline{\int_a^b} f(x)\,dx + \underline{\int_b^c} f(x)\,dx.$$

As f is integrable on $[a,c]$

$$\begin{aligned}\int_a^c f(x)\,dx &= \overline{\int_a^c} f(x)\,dx \\ &= \underline{\int_a^c} f(x)\,dx \\ &= \underline{\int_a^b} f(x)\,dx + \underline{\int_b^c} f(x)\,dx \\ &\le \overline{\int_a^b} f(x)\,dx + \overline{\int_b^c} f(x)\,dx \\ &= \overline{\int_a^c} f(x)\,dx \\ &= \int_a^c f(x)\,dx.\end{aligned}$$

Thus equality must hold at all the intermediate inequalities. Therefore

$$\underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x)\,dx \quad \text{and} \quad \underline{\int_b^c} f(x)\,dx = \overline{\int_b^c} f(x)\,dx$$

which implies the restrictions $f\big|_{[a,b]}$ and $f\big|_{[b,c]}$ are integrable. The rest follows from

$$\int_a^b f(x)\,dx = \overline{\int_a^b} f(x)\,dx \quad \text{and} \quad \int_b^c f(x)\,dx = \overline{\int_b^c} f(x)\,dx$$

and that equality holds in the displayed inequality. □

Proposition 4.2. *Let f be integrable on $[a,b]$ and let $[\alpha,\beta]\subseteq[a,b]$. The f is integrable on $[\alpha,\beta]$.*

Problem 4.1. Prove this. *Hint:* $[\alpha,\beta]=[a,\beta]\cap[\alpha,b]$ and Proposition 4.1. □

It is useful to define $\int_a^b f(x)\,dx$ even in the cases where $a=b$ and $b<a$.

Definition 4.3. For any function f define

$$\int_a^a f(x)\,dx = 0.$$

If $b < a$ and f is integrable on $[b, a]$ define

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx.$$

□

Proposition 4.4. *If f is integrable on the interval $[x_1, x_2]$ and $a, b, c \in [x_1, x_2]$ then, with the definitions above,*

$$\int_a^c f(x)\,dx = \int_a^b f(x)\,dx + \int_b^c f(x)\,dx.$$

Proof. This is just checking case by case (i.e. $a \le b \le c$, $a \le c \le b$ etc.) and is left to the reader. And please do not hand it in. □

Proposition 4.5. *Let $f(x)$ be integrable on $[a, b]$ and let $F\colon [a, b] \to \mathbb{R}$ be defined by*

$$F(x) = \int_a^x f(t)\,dt$$

then F is Lipschitz. That is there is a constant M such that for all $x_1, x_2 \in [a, b]$,

$$|F(x_2) - F(x_1)| \le M|x_2 - x_1|$$

and therefore F is continuous on $[a, b]$.

Problem 4.2. Prove this. *Hint:* As f is integrable on $[a, b]$, it is bounded on $[a, b]$, say $|f(x)| \le M$ on $[a, b]$. Without loss of generality we can assume that $x_1 \le x_2$. Then

$$|F(x_2)-F(x_1)| = \left|\int_a^{x_2} f(t)\,dt - \int_a^{x_1} f(t)\,dt\right| = \left|\int_{x_1}^{x_2} f(t)\,dt\right| \le \int_{x_1}^{x_2} |f(t)|\,dt$$

and it should be easy from here. □

Theorem 4.6 (Fundamental Theorem of Calculus Form 1)**.** *Let f be integrable on $[a, b]$. Define new function $F\colon [a, b] \to \mathbb{R}$ by*

$$F(x) = \int_a^x f(t)\,dt.$$

If f is continuous at the point $x \in (a, b)$, then the derivative of F exists at x and

$$F'(x) = f(x).$$

Problem 4.3. Prove this. *Hint:* First note

$$1 = \frac{1}{h}\int_x^{x+h} 1\,dt.$$

Multiply by $f(x)$ to get

$$f(x) = \frac{1}{h}\int_x^{x+h} f(x)\,dt$$

Also note

$$F(x+h) - F(x) = \int_a^{x+h} f(t)\,dt - \int_a^x f(t)\,dt = \int_x^{x+h} f(t)\,dt.$$

Combining some of these formulas we get

$$\begin{aligned}\frac{F(x+h)-F(x)}{h} - f(x) &= \frac{1}{h}\int_x^{x+h} f(t)\,dt - \frac{1}{h}\int_x^{x+h} f(x)\,dt \\ &= \frac{1}{h}\int_x^{x+h} (f(t) - f(x))\,dt.\end{aligned}$$

Let $\varepsilon > 0$. As f is continuous at x there is a $\delta > 0$ such that

$$|t - x| < \delta \quad \implies \quad |f(t) - f(x)| < \varepsilon.$$

Put this all together to show

$$|h| < \delta \quad \implies \quad \left|\frac{F(x+h)-F(x)}{h} - f(x)\right| < \varepsilon$$

and explain why this shows $F'(x) = f(x)$. □

Theorem 4.7 (Fundamental Theorem of Calculus Form 2)**.** *Let f be continuous on $[a,b]$ and let F be continuous on $[a,b]$ and differentiable (a,b) with $F' = f$ on (a,b). Then*

$$\int_a^b f(t)\,dt = F(b) - F(a) = F\Big|_a^b.$$

Problem 4.4. Prove this. *Hint:* Let

$$G(x) = \int_a^x f(t)\,dt - F(x)$$

and show $G'(x) = 0$ for $x \in (a,b)$. □

Corollary 4.8. *If f is continuous on $[a,b]$ and F is any anti-derivative of f on $[a,b]$ (that is $F'(x) = f(x)$ for $x \in [a,b]$), then*

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

Problem 4.5. Prove this. □

Definition 4.9. Let f be integrable on $[a,b]$. Then the ***average value*** of f on $[a,b]$ is

$$\frac{1}{b-a}\int_a^b f(x)\,dx.$$ □

Theorem 4.10 (The First Mean Value Theorem for Integrals)**.** *If f is continuous on $[a,b]$, then it achieves its average value. That is there is a $\xi \in (a,b)$ with*

$$f(\xi) = \frac{1}{b-a}\int_a^b f(x)\,dx.$$

Problem 4.6. Prove this. *Hint:* As f is continuous on the closed bounded set $[a,b]$, it achieves its maximum and minimum on this interval. Let $m = \min\{f(x) : x \in [a,b]\}$ and $M = \max\{f(x) : x \in [a,b]\}$ and let $\alpha, \beta \in [a,b]$ such that $f(\alpha) = m$ and $f(\beta) = M$. Now

$$f(\alpha) = m = \frac{1}{b-a}\int_a^b m\,dx \le \frac{1}{b-a}\int_a^b f(x)\,dx$$

and

$$f(\beta) = M = \frac{1}{b-a}\int_a^b M\,dx \ge \frac{1}{b-a}\int_a^b f(x)\,dx$$

and recall the intermediate value theorem. □

We now prove a somewhat stronger version of the second form of the Fundamental Theorem of Calculus.

Theorem 4.11. *Let F be continuous on $[a,b]$ assume that F is differentiable on (a,b) and let*

$$f(x) = F'(x)$$

on $[a,b]$. Then

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

(This differs from Theorem 4.7 as we are only assuming that f is integrable rather than continuous.)

Proof. Let $\varepsilon > 0$. As f is integrable there are step functions φ and ψ on $[a,b]$ with

$$(9)\quad \varphi \le f \le \psi \quad \text{and} \quad \int_a^b f\,dx - \varepsilon \le \int_a^b \varphi\,dx \le \int_a^b \psi\,dx \le \int_a^b f\,dx + \varepsilon.$$

We can assume there is a partition $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$ such that if $I_j = [x_{j-1}, x_j)$ then

$$\varphi = \sum_{j=1}^n m_j \chi_{I_j}, \qquad \psi = \sum_{j=1}^n M_j \chi_{I_j}.$$

We write $F(b) - F(a)$ as a telescoping sum:

$$F(b) - F(a) = F(x_n) - F(x_0) = \sum_{j=1}^n (F(x_j) - F(x_{j-1}))$$

As F is differentiable on $[x_{j-1}, x_j]$ we can apply the mean value theorem to get that there is a $\xi_j \in (x_{j-1}, x_j)$ with

$$F(x_j) - F(x_{j-1} = F'(\xi_j)(x_j - x_{j-1}) = f(\xi_j)(x_j - x_{j-1}) = f(\xi_j)|I_j|.$$

Combining these equations gives

$$F(b) - F(a) = \sum_{j=1}^n (F(x_j) - F(x_{j-1}) = \sum_{j=1}^n f(\xi_j)|I_j|.$$

But $\varphi \le f \le \psi$ which implies $m_j \le f(\xi_j) \le M_j$ and thus

$$\int_a^b \varphi\,dx = \sum_{j=1}^n m_j|I_j| \le F(b) - F(a) = \sum_{j=1}^n f(\xi_j)|I_j| \le \sum_{j=1}^n M_j|I_j| = \int_a^b \psi\,dx.$$

Combining this with the inequalities (9) gives

$$\int_a^b f\,dx - \varepsilon \le F(b) - F(a) \le \int_a^b f\,dx + \varepsilon.$$

As $\varepsilon > 0$ was arbitrary this gives $F(b) - F(a) = \int_a^b f\,dx$ as required. □

Problem 4.7. To see that Theorem 4.11 really is stronger than Theorem 4.7 we need to show that there is a function F on an interval $[a,b]$ such that $f = F'$ exists and is integrable on (a,b) but with f not continuous on (a,b). Let

$$F(x) = \begin{cases} x^2\cos(1/x), & x \ne 0; \\ 0, & x = 0 \end{cases}$$

Show that F is differentiable at all points of $\mathbb{R}$, and $f = F'$ is bounded on $[-1,1]$, but f is not continuous at $x = 0$. As f is continuous at all points other than 0 it is integrable on $[-1,1]$. □

We can now give the familiar integration by parts formula.

Theorem 4.12 (Integration by Parts)**.** *Let u and v continuous on $[a,b]$, differentiable on (a,b), with u' and v' integrable on $[a,b]$. Then*

$$\int_a^b u(x)v'(x)\,dx = u(x)v(x)\Big|_{x=a}^b - \int_a^b u'(x)v(x)\,dx.$$

Problem 4.8. Prove this. *Hint:* This follows from the product rule and the Fundamental Theorem of Calculus in the form

$$\int_a^b (u(x)v(x))'\,dx = u(x)v(x)\Big|_{x=a}^b.$$

You do have to worry a bit about if the integrals involved exist. Theorem 3.26 should help here. □

We now use integration by parts to give another form of the remainder in Taylor's Theorem.

Lemma 4.13. *Let f be $k+1$ times differentable on an open interval (α,β) and assume that $f^{(k+1)}$ is integrable. Then for $a, x \in (\alpha,\beta)$ we have*

$$\int_a^x \frac{(x-t)^{k-1}}{(k-1)!} f^{(k)}(t)\,dt = \frac{f^{(k)}(a)}{k!}(x-a)^k + \int_a^x \frac{(x-t)^k}{k!} f^{(k+1)}(t)\,dt.$$

Problem 4.9. Prove this. *Hint:* Use integration by parts with $v'(t) = \frac{(x-t)^{k-1}}{(k-1)!}$ and $u = f^{(k)}(t)$. □

Theorem 4.14 (Taylor's Theorem with Integral form of the Remainder)**.** *Let f be $n+1$ times differentable on (α, β) and assume that $f^{(n+1)}$ is integrable. Then for $a, x \in (\alpha, \beta)$*

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_n(x)$$

where the remainder term $R_n(x)$ is given by

$$R_n(x) = \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\, dt.$$

Problem 4.10. Prove this. *Hint:* Note that Lemma 4.13 can be rewritten as

$$R_{k-1}(x) = \frac{f^{(k)}(a)}{k!}(x-a)^k + R_k(x)$$

and by the Fundamental Theorem of Calculus and integration by parts

$$\begin{aligned} f(x) - f(a) &= \int_a^x f'(t)\, dt \\ &= -\int_a^x (-1) f'(t)\, dt \\ &= -\int_a^x \left(\frac{d}{dt}(x-t)\right) f'(t)\, dt \\ &= -\frac{d}{dt}(x-t)f'(t)\Big|_{t=a}^{x} + \int_a^x (x-t) f''(t)\, dt \\ &= f(a)(x-a) + R_1(x). \end{aligned}$$

Now use induction. □

Theorem 4.15 (Change of Variable Formula)**.** *Let the map $x = u(t)$ map the interval $[c, d]$ into the interval $[a, b]$ and assume that $u'(t)$ is integrable on $[c, d]$. Then for any continuous function f on $[a, b]$*

$$\int_{u(c)}^{u(d)} f(x)\, dx = \int_c^d f(u(t))u'(t)\, dt.$$

Problem 4.11. Prove this. *Hint:* Do this in steps

(a) Explain why both the integrals exist.
(b) Define F on $[a, b]$ by

$$F(x) = \int_a^x f(y)\, dy$$

and explain why

$$F'(x) = f(x) \quad \text{and} \quad \int_{u(c)}^{u(d)} f(x)\, dx = F(u(d)) - F(u(c)).$$

(c) On $[c, d]$ define
$$G(t) = F(u(t)).$$
By the chain rule
$$G'(t) = F'(u(t))u'(t) = f(u(t))u'(t)$$
and so by Theorem 4.11
$$\int_c^d f(u(t))u'(t)\,dt = \int_c^d G'(t)\,dt = G(d) - G(c).$$
(d) Put the pieces above together to finish the proof. □

5. Definition of the logarithm and exponential functions.

Define a function $L\colon (0,\infty) \to \mathbb{R}$ by
$$L(x) = \int_1^x \frac{dx}{x}.$$
We know this should be the natural logarithm, but we now verify directly from its definition that it has the correct properties.

Proposition 5.1. *The derivative of L is*
$$L'(x) = \frac{1}{x}$$
and thus L is strictly increasing. Therefore L is one-to-one (that is injective).

Proof. By the Fundamental Theorem of Calculus
$$L'(x) = \frac{1}{x} > 0$$
as $x > 0$ which implies L is strictly increasing. □

Proposition 5.2. *Let $a, b > 0$ then*
$$\int_a^b \frac{dx}{x} = L(b/a).$$

Problem 5.1. Prove this. *Hint:* In the integral $\int_a^b \frac{dx}{x}$ do the change of variable $x = at$ to get
$$\int_a^b \frac{dx}{x} = \int_1^{b/a} \frac{dt}{t}.$$
□

Proposition 5.3. *If $a, b > 0$ then*
$$L(ab) = L(a) + L(b).$$

Problem 5.2. Prove this. *Hint:*
$$L(ab) = \int_1^{ab} \frac{dx}{x} = \int_1^a \frac{dx}{x} + \int_a^{ab} \frac{dx}{x}$$
and use Proposition 5.2. □

The last Proposition and induction yield:

Corollary 5.4. *If $a > 0$ and n is a positive integer*

$$L(a^n) = nL(a).$$

□

Proposition 5.5. *The function $L\colon (0,\infty) \to \mathbb{R}$ is a bijection.*

Problem 5.3. Prove this. *Hint:* Recall the saying that L is a bijection is just saying that it is one-to-one and onto. We have already seen that L is injective. To see that it is surjective (that is onto) note that $L(2) > 0$ and $L(1/2) < 0$. Also for a positive integer n

$$L(2^n) = nL(2) \qquad \text{and} \qquad L(1/2^n) = nL(1/2).$$

If y_0 is any real number we can find (by Archimedes' principle) a positive integer n such that

$$nL(1/2) < y_0 < nL(2).$$

Also we know that L is continuous (why?). Now you should be able to show that there is a $x_0 \in (0,\infty)$ with $L(x_0) = y_0$. □

Because the function $L\colon (0,\infty) \to \mathbb{R}$ is bijective, it has an inverse $E\colon \mathbb{R} \to (0,\infty)$. As L is strictly increasing, continuous, and differentiable with $L'(x) \neq 0$ for all x theorems from earlier this term imply that E is strictly increasing, continuous, and differentiable.

Proposition 5.6. *The function E satisfies $E(0) = 1$ and*

$$E'(x) = E(x).$$

Problem 5.4. Prove this. *Hint:* $L(1) = 0$. And as L and E are inverses of each other $L(E(x)) = x$ for all $x \in \mathbb{R}$. Therefore $\frac{d}{dx}L(E(x)) = 1$. Use the chain rule and that we know the derivative of L. □

Proposition 5.7. *For all real numbers x*

$$E(-x) = \frac{1}{E(x)}.$$

Problem 5.5. Prove this. *Hint:* There are several ways to do this. One is to take the derivative of $E(x)E(-x)$ and show it is zero. Anther is to note that $L(a) + L(1/a) = L(1) = 0$ □

Proposition 5.8. *For all real numbers a, b*

$$E(a+b) = E(a)E(b).$$

Problem 5.6. Prove this. *Hint:* One way is to deduce this from the property $L(\alpha\beta) = L(\alpha) + L(\beta)$ of L. Anther is to show that the derivative of the function

$$f(x) = E(x+a)E(-x)$$

is zero and therefore f is constant. □

Proposition 5.9. *If n is any integer, positive or negative, and t is any real number*

$$E(nt) = E(t)^n$$

If m is a positive integer then

$$E\left(\frac{1}{m}t\right)^m = E(t)$$

and thus $E(\frac{1}{m}t)$ is the positive m-th root of $E(t)$.

Problem 5.7. Prove this. □

In light of Proposition 5.9 If r is a rational number, say $r = n/m$ with m, n integers and $m > 0$, then for a positive number a we can define

$$a^r = a^{n/m} = (a^n)^{1/m}$$

where $(a^n)^{1/m}$ is the positive m-th root of a^n. We would also like to define a^r when r is irrational. Note that when $r = m/n$ and $a = E(t)$, then Proposition 5.9 shows us that

$$a^r = E(t)^{n/m} = (E(t)^n)^{1/m} = E(nt)^{1/m} = E\left(\frac{1}{m}nt\right) = E(rt). \tag{10}$$

But $E(rt)$ makes sense for all real numbers r. We now formalize all this.

Definition 5.10. We now officially define ***logarithm*** of a positive number x to be

$$\ln(x) = L(x) = \int_1^x \frac{dt}{t},$$

the number e to be

$$e = E(1)$$

and for any real number x we define the power e^x by

$$e^x = E(x).$$

□

Definition 5.11. Let $a > 0$. Then for any real number r define

$$a^r = e^{r\ln(a)}.$$

(Note if $a = E(t) = e^t$ then $\ln(a) = t$ and this becomes $a^r = e^{r\ln(a)} = e^{rt} = E(rt)$ which agrees with our preliminary definition (10).) □

Proposition 5.12. *If $a > 0$ and $r = n/m$ is a rational number with $m > 0$, then*

$$a^r = (a^n)^{1/m}$$

so that our definition agrees with what it should be on the rational numbers. In particular $a^{1/2}$ is the square root of a, $a^{1/3}$ is the cube root of a etc.

Problem 5.8. Prove this. □

Proposition 5.13. *With these definition the following hold*

(a) *If $a > 0$ then for all $r, s \in \mathbb{R}$*

$$a^r a^s = a^{r+s}, \qquad \frac{a^r}{a^s} = a^{r-s}.$$

and,

$$(a^r)^s = a^{rs}.$$

(b) *If $r \in \mathbb{R}$ and $a, b > 0$ then*

$$a^r b^r = (ab)^r.$$

(c) *If $r, s \in \mathbb{R}$ and $a > 0$, then*

$$(a^r)^s = a^{rs}.$$

Problem 5.9. Prove this. □

Proposition 5.14. *Let r be a real number and on define $f\colon (0, \infty) \to (0, \infty)$ by*

$$f(x) = x^r.$$

Then f is differentiable and

$$f'(x) = rx^{r-1}.$$

Problem 5.10. Prove this. *Hint:* We know that $E(x) = e^x$ is differentiable with derivative $E'(x) = E(x)$ and that $\ln(x)$ is differentiable with $\frac{d}{dx}\ln(x) = 1/x$. Thus $f(x) = e^{r\ln(x)} = E(r\ln(x))$ is a composition of differentiable functions. Use the chain rule to derive the formula for $f'(x)$. □

Proposition 5.15. *Let a be a positive real number and define $g\colon \mathbb{R} \to (0, \infty)$ by*

$$g(x) = a^x.$$

Then g is differentiable and

$$g'(x) = \ln(a)a^x.$$

Problem 5.11. Prove this. □

There is is anther way define e^x based on the following

Proposition 5.16. *For any $x \in \mathbb{R}$*

$$\lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n = e^x.$$

Problem 5.12. Here is one method of proving this.

(a) Use Taylor's theorem with the Lagrange form of the remainder to show that for $|y| \le 1/2$ that

$$\ln(1 + y) = y + R(y)$$

where

$$|R(y)| \le 2y^2.$$

(b) Let $x \in \mathbb{R}$ and note that if $|x/n| \le 1/2$, we have

$$\ln(1 + x/n) = \frac{x}{n} + R(x/n)$$

and

$$|R(x/n)| \le \frac{2x^2}{n^2}$$

(c) Use (b) to show

$$\lim_{n\to\infty} n\ln(1 + x/n) = x.$$

(d) Now use that

$$\left(1 + \frac{x}{n}\right)^n = e^{n\ln(1+x/n)}$$

to show that

$$\lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n = e^x$$

holds. □

There are books that instead of defining the $\ln(x)$ as $\int_1^x dt/t$, and then defining e^x as the inverse of $\ln(x)$, first define e^x as

$$e^x = \left(1 + \frac{x}{n}\right)^n,$$

show that this behaves like e^x should and then define $\ln(x)$ as the inverse of e^x and finally $a^r = e^{r\ln(a)}$. This takes more work than what we have done, but has the advantage that it is possible to define e^x, $\ln(x)$, and a^x before defining the derivative and integral.

6. Series

The material here corresponds to parts of Chapter VII Rosenlicht.

6.1. **Basic definitions and results about series.** We now wish to make sense out of infinite sums

$$\sum_{k=1}^{\infty} = a_1 + a_2 + a_3 + \cdots$$

Definition 6.1. Let $\langle a_k\rangle_{k=n_0}^{\infty}$ be a sequence of real numbers. The corresponding ***infinite series*** is (or just ***series***) is the sum

$$\sum_{k=k_0}^{\infty} a_k = a_{k_0} + a_{k_0+1} + a_{k_0+2} + \cdots .$$

The n-th ***partial sum*** of the series is

$$A_n = a_{n_0} + a_{n_0+1} + a_{n_0+2} + \cdots + a_{n-1} + a_n = \sum_{k=n_0}^{n} a_k.$$

We say the series ***converges*** and has sum A if and only if

$$\lim_{n\to\infty} A_n = A.$$

If $\sum_{k=1}^{\infty} a_k$ does not converge, it ***diverges***. □

To make notation easier, when proving results about series we will usually let $n_0 = 0$ or $n_0 = 1$.

Here is a result that follows at once from the facts about limits of sequences.

Theorem 6.2. *If $\sum_{n=1}^{\infty} a_k$ and $\sum_{k=1}^{\infty} b_k$ both converge, then for any constants c_1 and c_2 the series $\sum_{k=1}^{\infty} (c_1 a_k + c_2 b_k)$ also converges and*

$$\sum_{k=1}^{\infty} (c_1 a_k + c_2 b_k) = c_1 \sum_{n=1}^{\infty} a_k + c_2 \sum_{n=1}^{\infty} b_k$$

Proof. Let

$$\begin{aligned} A_n &= (a_1 + \cdots a_n) \\ B_n &= (b_1 + \cdots b_n) \\ C_n &= \big((c_1 a_1 + c_2 b_1) + \cdots + (c_1 a_n + c_2 a_n)\big) \end{aligned}$$

be the partial sums of the series. We are given that

$$\lim_{n\to\infty} A_n = A, \qquad \lim_{n\to\infty} B_n = B$$

exist and want to show $\lim_{n\to\infty} C_n = c_1 A + c_2 B$. Note

$$\begin{aligned} C_n &= \big((c_1 a_1 + c_2 b_1) + \cdots + (c_1 a_n + c_2 a_n)\big) \\ &= c_1(a_1 + \cdots a_n) + c_2(b_1 + \cdots b_n) \\ &= c_1 A_n + c_2 B_n \end{aligned}$$

and therefore

$$\lim_{n\to\infty} C_n = \lim_{n\to\infty} (c_1 A_n + c_2 B_n) = c_1 A + c_2 B$$

as required. □

Before going on we note that for any series $\sum_{k=1}^{\infty} a_k$ with partial sums $A_n = \sum_{k=1}^{n} a_k$ we have the elementary relation

$$A_n = A_{n-1} + a_n,$$

or equivalently

$$a_n = A_n - A_{n-1}.$$

This will come up several times in what follows starting with the following:

Theorem 6.3. *If the series $\sum_{k=1}^{n} a_k$ converges, then*

$$\lim_{n\to\infty} a_n = 0.$$

Proof. If $A_n = \sum_{k=1}^{n} a_k$ then $\lim_{n\to\infty} A_n = A$ exists as the series converges. But then also $\lim_{n\to\infty} A_{n-1} = A$ and so

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} (A_n - A_{n-1}) = A - A = 0.$$

□

Remark 6.4. Often the previous theorem is used in its contrapositive form: If $\lim_{k\to\infty} a_k \neq 0$, then $\sum_{k=1}^\infty a_k$ diverges. From this it is not hard to give lots of examples of series that do not converge. For example none of the following converge

$$\sum_{k=1}^{\infty}(-1)^k, \qquad \sum_{k=1}^{\infty}\sin(k), \qquad \sum_{n=1}^{\infty}\frac{n^2-2}{2n^2+5}.$$

□

Proposition 6.5. *The series* $\sum_{k=1}^\infty a_k$ *converges if and only if for all* $\varepsilon > 0$ *there is a* N *such that*

$$N \le m < n \qquad \Longrightarrow \qquad |a_{m+1} + a_{m+2} \cdots + a_n| < \varepsilon.$$

Problem 6.1. Prove this. *Hint:* What is the Cauchy condition for the sequence $\langle A_n\rangle_{n=1}^\infty$ of partial sums? □

Proposition 6.6. *Let* $\sum_{k=1}^\infty a_k$ *and* $\sum_{k=1}^\infty b_k$ *be two series such that* $a_k = b_k$ *except for a finite number of values* k*. Then either they both converge or both diverge. (An informal way so state this is that changing a finite number of terms of a series does not effect whether it converges or diverges.)*

Proof. By the hypothesis there is an n_0 such that such that

$$a_k = b_k \qquad \text{for all} \qquad k \ge n_0.$$

If $n \ge n_0$ then

$$\begin{aligned} B_n &= B_{n_0} + \sum_{k=n_0+1}^{n} b_k \\ &= B_{n_0} + \sum_{k=n_0+1}^{n} a_k && \text{(as } a_k = b_k \text{ when } k \ge n_0) \\ &= B_{n_0} - A_{n_0} + A_{n_0} + \sum_{k=n_0+1}^{n} a_k \\ &= (B_{n_0} - A_{n_0}) + A_n. \end{aligned}$$

Letting $c = B_{n_0} - A_{n_0}$, which is a constant, we have that $B_n = A_n + c$ for $n \ge n_0$. Thus the sequences $\langle A_n\rangle_{n=1}^\infty$ and $\langle B_n\rangle_{n=1}^\infty$ either both converge or both diverge. □

Lemma 6.7. *If* $r \neq 1$ *then*

$$a + ar + ar^2 + \cdots ar^n = \sum_{k=0}^{n} ar^k = \frac{a - ar^{n+1}}{1-r}.$$

Proof. Let $S_n = a + ar + ar^2 + \cdots ar^n$. Then

$$\begin{aligned} (1-r)S_n &= a + ar + ar^2 + \cdots + ar^n - r(a + ar + ar^2 + \cdots + ar^n) \\ &= a + ar + ar^2 + \cdots ar^n - ar - ar^2 - \cdots - ar^n - ar^{n+1} \\ &= a - ar^{n+1}. \end{aligned}$$

As $r \neq 1$ we can divide by $(1-r)$ to get the desired result. $\square$

Lemma 6.8. *If* $|r| < 1$ *then*

$$\lim_{n\to\infty} |r|^n = 0.$$

Proof. One way to do this is to note that $\langle |r|^n \rangle_{n=1}^{\infty}$ is a monotone decreasing sequence that is bounded below (as all the terms are positive). Therefore the sequence has a limit, say $L = \lim_{n\to\infty} |r|^n$. But then L is also the limit of $\langle |r|^{n+1} \rangle$ and so $L = \lim_{n\to\infty} |r|^{n+1} = \lim_{n\to\infty} |r||r|^n = |r|L$. Thus $L = |r|L$ and as $|r| \neq 1$ this implies $L = 0$.

Here is anther proof using what we have recently covered. Let $\varepsilon > 0$ and set $N = \ln(\varepsilon)/\ln(|r|)$. If $n > N$ it is not hard to check $\big||r|^n - 0\big| = |r|^n < \varepsilon$, which shows $\lim_{n\to\infty} |r|^n = 0$. $\square$

One of the most basic examples of convergent series is a geometric series with ratio less than one. Many results about series involve comparison to such a series.

Theorem 6.9 (Infinite Geometric Series)**.** *Let* a, r *be real numbers with* $a \neq 0$*. Then the series*

$$a + ar + ar^2 + \cdots = \sum_{k=0}^{\infty} ar^k$$

converges if and only if $|r| < 1$ *in which case its sum is*

$$\sum_{k=0}^{\infty} ar^k = \frac{a}{1-r}.$$

Proof. If $|r| \geq 1$ then the n-th term ar^n satisfies $|ar^n| \geq |a| > 0$ and so $\lim_{n\to\infty} ar^n \neq 0$ and thus the series diverges.

Now assume $|r| < 1$. We have seem in Lemma 6.7 that the nth partial sum is

$$S_n = \frac{a - ar^{n+1}}{1-r}.$$

Now by the last lemma,

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} \frac{a - ar^{n+1}}{1-r} = \frac{a - a\cdot 0}{1-r} = \frac{a}{1-r}$$

as required. $\square$

6.2. **Series with positive terms.**

Theorem 6.10. *Let* $\sum_{k=1}^{\infty} a_k$ *be a series with* $a_k \geq 0$ *for all* k*. Then* $\sum_{k=1}^{\infty} a_k$ *converges if and only if the sequence,* $\langle A_n \rangle_{n=1}^{\infty}$ *(with* $A_n = a_1 + \cdots a_n$*) of partial sums is bounded.*

Proof. If $\sum_{k=1}^{\infty} a_k$ converges, then $\lim_{n\to\infty} A_n = A$ exists by definition. But a convergent sequence is bounded. If $\langle A_n \rangle_{n=1}^{\infty}$ is bounded, then $A_{n+1} = A_n + a_{n+1} \geq A_n$ so the series is monotone increasing. A bounded monotone sequence is convergent. $\square$

Remark 6.11. When talking about series, $\sum_{k=1}^{\infty} a_k$, of non-negative terms we will use the following suggestive notation.

$$\sum_{k=1}^{\infty} a_k < \infty \iff \text{The series converges}$$
$$\sum_{k=1}^{\infty} a_k = \infty \iff \text{The series series diverges.}$$

This notation is not appropriate when talking about series with terms of mixed signs. For example the series $\sum_{k=1}^{\infty}(-1)^{k+1}$ has bounded partial sums, but is not convergent. □

6.3. **Tests for the convergence of series with monotone terms.** In general it is easier to understand the convergence of series with monotone decreasing terms. As a first example.

Theorem 6.12 (Cauchy Condensation Test)**.** *If $\langle a_k\rangle_{k=1}^{\infty}$ is a sequence of non-negative numbers that are monotone decreasing, then*

$$\sum_{k=1}^{\infty} a_k < \infty$$

if and only if

$$\sum_{k=0}^{\infty} 2^k a_{2^k} < \infty.$$

Proof. Let the partial sums of the two series be

$$A_n = \sum_{k=1}^{n} a_k, \qquad B_n = \sum_{k=0}^{n} 2^k a_{2^k}.$$

We will show

$$A_{2^{n+1}-1} \le B_n \tag{11}$$
$$B_n \le 2A_{2^n}. \tag{12}$$

If these hold the result is easy. If $\sum_{k=0}^{\infty} 2^k a_{2^k} < \infty$ then for any positive integer m choose n such that $m \le 2^{n+1}-1$. By (11),

$$A_m \le A_{2^{n+1}-1} \le B_n \le \sum_{k=0}^{\infty} 2^k a_{2^k} < \infty$$

and therefore the partial sums of $\sum_{k=1}^{\infty} a_k$ are bounded above and thus $\sum_{k=0}^{\infty} a_k < \infty$.

Conversely if $\sum_{k=1}^{\infty} a_k < \infty$ then for any positive integer n we use (12) to get

$$B_n \le 2A_{2^n} \le 2\sum_{k=1}^{\infty} a_k < \infty$$

which shows the partial sums of $\sum_{k=0}^{\infty} 2^k a_{2^k}$ are bounded above and thus $\sum_{k=0}^{\infty} 2^k a_{2^k}$ converges.

We now prove (11). Using that the terms are monotone decreasing,

$$\begin{aligned} A_{2^{n+1}-1} &= a_1 + \underbrace{(a_2 + a_3)}_{2^1 \text{ terms}} + \underbrace{(a_4 + \cdots + a_7)}_{2^2 \text{ terms}} + \cdots + \underbrace{(a_{2^n} + \cdots + a_{2^{n+1}-1})}_{2^n \text{ terms}} \\ &\le a_1 + \underbrace{(a_2 + a_2)}_{2^1 \text{ terms}} + \underbrace{(a_4 + \cdots + a_4)}_{2^2 \text{ terms}} + \cdots + \underbrace{(a_{2^n} + \cdots + a_{2^n})}_{2^n \text{ terms}} \\ &= a_1 + 2^2 a_{2^2} + 2^3 a_{2^3} + \cdots + 2^n a_{2^n} \\ &= B_n. \end{aligned}$$

The proof (12) is similar

$$\begin{aligned} A_{2^n} &= a_1 + a_2 + \underbrace{(a_3 + a_4)}_{2^1 \text{ terms}} + \underbrace{(a_5 + \cdots a_8)}_{2^2 \text{ terms}} + \cdots + \underbrace{(a_{2^{n-1}+1} + \cdots + a_{2^n})}_{2^{n-1} \text{ terms}} \\ &\ge a_1 + a_2 + \underbrace{(a_4 + a_4)}_{2^1 \text{ terms}} + \underbrace{(a_8 + \cdots a_8)}_{2^2 \text{ terms}} + \cdots + \underbrace{(a_{2^n} + \cdots + a_{2^n})}_{2^{n-1} \text{ terms}} \\ &= a_1 + a_2 + 2^1 a_{2^2} + 2^2 a_{2^3} + \cdots + 2^{n-1} a_{2^n} \\ &= 2^{-1} a_1 + 2^{-1} a_1 + a_2 + 2^1 a_{2^2} + 2^2 a_{2^3} + \cdots + 2^{n-1} a_{2^n} \\ &= 2^{-1} a_1 + 2^{-1}\left(2^0 a_1 + 2^1 a_2 + 2^2 a_{2^2} + 2^3 a_{2^3} + \cdots + 2^n a_{2^n}\right) \\ &= 2^{-1} a_1 + 2^{-1} B_n \\ &\ge \frac{1}{2} B_n. \end{aligned}$$

Multiplication by 2 completes the proof. □

Theorem 6.13. *For any real number $p > 0$ the series*

$$\sum_{k=1}^{\infty} \frac{1}{k^p}$$

converges if and only if $p > 1$.

Proof. We use the Cauchy-Condensation Test, which applies as the terms of the series are decreasing. The given series converges if and only if

$$\sum_{k=1}^{\infty} 2^k \frac{1}{(2^k)^p} = \sum_{k=1}^{\infty} \left(\frac{2}{2^p}\right)^k$$

converges. This is a geometric series with ratio

$$r = \frac{2}{2^p}.$$

Therefore the series converges if and only if $r = 2/2^p < 1$, that is if and only if $p > 1$. □

Anther method of dealing with series with monotone terms is by comparison with an integral. Let us start with an example. Let $f(x)$ be monotone decreasing on the interval $[0,6]$ and let

$$a_k = f(k) \qquad \text{for} \qquad 1 \le k \le 6$$

and

$$A_n = a_1 + \cdots + a_n = f(1) + \cdots + f(n).$$

Then, see Figure 5, we can compare the integral $\int_1^6 f(x)\,dx$ with some of the Riemann sums for the partition $\mathcal{P} = \{1,2,3,4,5,6\}$ to get

$$\int_1^6 f(x)\,dx \le A_5 \le A_6 \le f(1) + \int_1^6 f(x)\,dx.$$

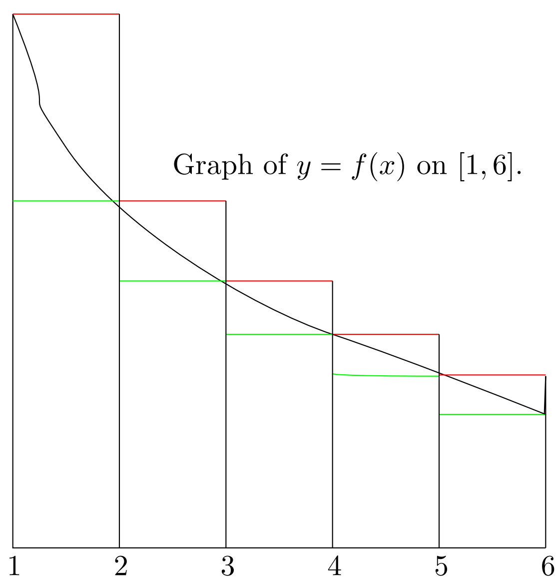


Figure 5. The area under the tall (with red tops) rectangles is $A_5 = f(1) + f(2) + f(3) + f(4) + f(5)$. The area under the short (with green tops) rectangles is $A_6 - f(1) = f(2) + f(3) + f(4) + f(5) + f(6)$. The area of the integral is clearly in between these two areas and therefore

$$A_6 - f(1) \le \int_1^6 f(x)\,dx \le A_5.$$

This can be rearranged to give

$$\int_1^6 f(x)\,dx \le A_5 \le A_6 \le f(1) + \int_1^6 f(x)\,dx = a_1 + \int_1^6 f(x)\,dx$$

which is a bit more aesthetic.

We could, and since this is a mathematics class, should be a bit more formal. Note that on any interval $[k, k+1]$ we have, because f is decreasing, that

$$f(k) \geq f(x) \geq f(k+1).$$

Then integration over $[k, k+1]$ and using that $\int_k^{k+1} f(k)\,dx = f(k)$ and $\int_k^{k+1} f(k+1)\,dx = f(k+1)$

$$f(k) \geq \int_k^{k+1} f(x)\,dx \geq f(k+1).$$

This can be summed it two ways to get

$$\int_1^6 f(x)\,dx = \sum_{k=1}^{5} \int_k^{k+1} f(x)\,dx \leq \sum_{k=1}^{5} f(k) = A_5$$

and

$$A_6 - a_1 = \sum_{k=2}^{6} f(k) \leq \sum_{k=1}^{5} \int_k^{k+1} f(x)\,dx = \int_1^6 f(x)\,dx.$$

Of course there is nothing special about $n = 6$ in this argument.

Proposition 6.14. *Let $f\colon [1,\infty) \to [0,\infty)$ be a monotone decreasing non-negative function. Let $a_k = f(k)$ and let*

$$A_n = \sum_{k=1}^{n} a_k$$

be the n-th partial sum of the series $\sum_{k=1}^{\infty} a_k$. Then

$$\int_1^n f(x)\,dx \leq A_n \leq f(1) + \int_1^n f(x)\,dx.$$

Problem 6.2. Use a variation of the argument given for $n = 6$ to prove this. □

Theorem 6.15 (The Integral Test)**.** *Let $f\colon [1,\infty) \to [0,\infty)$ be a monotone decreasing non-negative function. Let $a_k = f(k)$ and let*

$$A_n = \sum_{k=1}^{n} a_k$$

be the n-th partial sum of the series $\sum_{k=1}^{\infty} a_k$. Then

$$\sum_{k=1}^{\infty} a_k < \infty \qquad \Longleftrightarrow \qquad \lim_{n\to\infty} \int_1^n f(x)\,dx \quad \text{exists and is finite.}$$

(Note that $\left\langle \int_1^n f(x)\,dx \right\rangle_{n=1}^{\infty}$ is a monotone increasing sequence, thus the limit exists, but might be $+\infty$.)

Problem 6.3. Prove this. □

Problem 6.4. Use the Integral Test to give anther proof of Theorem 6.13. □

Problem 6.5. Use the Integral Test to show

$$\sum_{k=2}^{\infty} \frac{1}{n(\ln(n))^p}$$

converges if and only if $p > 1$. □

6.4. **Comparison tests.**

Proposition 6.16. *Let Let $\sum_{k=1}^{\infty} a_k$ and $\sum_{k=1}^{\infty} b_k$ be two series of non-negative terms. Assume there is a constant $C > 0$ such that*

$$a_k \leq C b_k$$

for all k. Then

(a) *If $\displaystyle\sum_{k=1}^{\infty} b_k$ converges, so does $\displaystyle\sum_{k=1}^{\infty} a_k$.*
(b) *If $\displaystyle\sum_{k=1}^{\infty} a_k$ diverges, so does $\displaystyle\sum_{k=1}^{\infty} b_k$.*

Problem 6.6. Prove this. *Hint:* Consider partial sums. □

Theorem 6.17 (Limit Comparison Test)**.** *Let $\sum_{k=1}^{\infty} a_k$ and $\sum_{k=1}^{\infty} b_k$ be two series of positive terms. Assume that*

$$L = \lim_{k\to\infty} \frac{a_k}{b_k}$$

exists. Then

(a) $\displaystyle\sum_{k=1}^{\infty} b_k < \infty$ *implies* $\displaystyle\sum_{k=1}^{\infty} a_k < \infty$
(b) *If $L \neq 0$ and $\displaystyle\sum_{k=1}^{\infty} a_k = \infty$, then $\displaystyle\sum_{k=1}^{\infty} b_k = \infty$.*

Often the following special case is enough.

Corollary 6.18. *Let $\sum_{k=1}^{\infty} a_k$ and $\sum_{k=1}^{\infty} b_k$ be two series of positive terms. Assume that*

$$L = \lim_{k\to\infty} \frac{a_k}{b_k}$$

exists and $L \neq 0$. Then $\displaystyle\sum_{k=1}^{\infty} a_k$ converges if and only if $\displaystyle\sum_{k=1}^{\infty} b_k$ converges. □

Problem 6.7. Prove Theorem 6.17. *Hint:* Recall that a convergent sequence is bounded. Thus $\langle a_k/b_k\rangle_{k=1}^{\infty}$ is bounded and therefore there is a constant C such that $a_k/b_k \leq C$. Thus Proposition 6.16 applies. □

Here some applications of these results.

Example 6.19. Does the series $\sum_{k=1}^{\infty} \frac{k^3+2k^2+7}{3k^5+2}$ converge? Let this series be $\sum_{k=1}^{\infty} a_k$ and let $\sum_{k=1}^{\infty} b_n$ be the p-series $\sum_{k=1}^{\infty} \frac{1}{k^2}$. Then it is not hard to check that

$$\lim_{k\to\infty} \frac{a_k}{b_k} = \frac{1}{3}.$$

Therefore, by Corollary 6.18, $\sum_{k=1}^{\infty} a_k$ converges if and only if $\sum_{k=1}^{\infty} b_k$ converges. But $\sum_{k=1}^{\infty} b_k$ is a p series with $p = 2 > 1$ and so both series converge. □

Example 6.20. Does the series $\sum_{k=1}^{\infty} = \sum_{k=1}^{\infty} \left(\sqrt[3]{n+5} - \sqrt[3]{n-2}\right)$ converge? Let $f(x) = \sqrt[3]{x} = x^{1/3}$. Then for $n > 2$ by the mean value theorem there is a ξ_n between -2 and 5 such that

$$a_n = f(n+5) - f(n-2) = f'(n+\xi_n)((n+5)-(n-2)) = \frac{1}{3}(n+\xi_n)^{-2/3}7.$$

Therefore if $\sum_{k=1}^{\infty} b_k$ is the divergent p-series $\sum_{k=1}^{\infty} 1/n^{2/3}$ we have

$$\lim_{k\to\infty} \frac{a_k}{b_k} = \frac{7}{3}.$$

So $\sum_{k=1}^{\infty} a_k$ diverges by limit comparison to $\sum_{k=1}^{\infty} b_k$.

Problem 6.8. For practice in these ideas do Problems 10 and 11 on Page 161 of the text. *Hint:* For Problem 11 it might help to notice that

$$\frac{1}{n} - \frac{1}{n+x} = \frac{x}{n(n+x)} \quad \text{and} \quad \lim_{n\to\infty} \frac{1/n^2}{1/(n(n+x))} = 1.$$ □

6.5. **The root and ratio tests.** These are basically just limit comparisons with a geometric series. To get started here is a version of the comparison were we only worry about the comparison for large values.

Lemma 6.21. *Let $\sum_{k=1}^{\infty} a_k$ and $\sum_{k=1}^{\infty} b_k$ be series of positive terms. Assume there is an N such that*

$$a_k \le b_k \qquad \textit{for all} \qquad k > N$$

and that $\sum_{k=1}^{\infty} b_k < \infty$. Then $\sum_{k=1}^{\infty} a_k < \infty$.

Proof. Let A_n and B_n be the partial sums of these series. Let

$$C_1 = \max\{A_n : 1 \le n \le N\}.$$

If $n > N$ then

$$\begin{aligned}
A_n &= (a_1 + \cdots a_N) + (a_{N+1} + \cdots + a_n) \\
&\le (a_1 + \cdots a_N) + (b_{N+1} + \cdots + b_n) \\
&= (a_1 + \cdots a_N) - (b_1 + \cdots + b_N) + (b_1 + \cdots + b_N + b_{N+1} + \cdots + b_n) \\
&= A_N - B_N + B_n \\
&\le A_N - B_N + \sum_{k=1}^{\infty} b_k < \infty.
\end{aligned}$$

Therefore if

$$C = \max\left\{C_1, A_N - B_N + \sum_{k=1}^{\infty} b_k\right\}$$

we have

$$A_n \le C$$

for all n. Thus the partial sums of $\sum_{k=1}^{\infty} a_k$ are bounded which implies that it is convergent. □

The following is a dressed up version of doing a comparison with a geometric series.

Theorem 6.22 (Root Test). *Let $\sum_{k=1}^{\infty} a_k$ be a series of positive terms and assume the limit*

$$\rho := \lim_{k\to\infty} (a_k)^{1/k}.$$

exists.

(a) If $\rho < 1$ then the series converges.
(b) If $\rho > 1$ then the series diverges.

Problem 6.9. Prove this. *Hint:* For (a) let r be any number such that $\rho < r < 1$. Then $\rho = \lim_{k\to\infty} (a_k)^{1/k} < r$ implies there is a N such that

$$k > N \qquad \Longrightarrow \qquad (a_k)^{1/k} < r.$$

Then

$$a_k < r^k \qquad \text{for all} \qquad k > N.$$

Now consider Lemma 6.21 and Theorem 6.9.

For (b) show that if $\rho > 1$ then $\lim_{k\to\infty} a_k \ne 0$. □

Here is anther dressed up version of comparison with a geometric series.

Theorem 6.23 (Ratio Test). *Let $\sum_{k=1}^{\infty} a_k$ be a series of positive terms assume the limit*

$$\rho := \lim_{k\to\infty} \frac{a_{k+1}}{a_k}$$

exists.

(a) If $\rho < 1$, then the series converges.
(b) If $\rho > 1$, then the series diverges.

Problem 6.10. Prove this. *Hint:* For (a) let r be a number such that $\rho < r < 1$. Then, by the definition of lim, there is a N such that

$$k > N \qquad \Longrightarrow \qquad \frac{a_{k+1}}{a_k} < r.$$

Thus for $k > N$ we have

$$a_k = a_{N+1} \frac{a_{N+2}}{a_{N+1}} \frac{a_{N+3}}{a_{N+2}} \cdots \frac{a_{k-1}}{a_{k-2}} \frac{a_k}{a_{k-1}} = (a_{N+1}) \prod_{j=N+1}^{k-1} \frac{a_{j+1}}{a_j} < a_{N+1} r^{k-N-1}.$$

The series

$$\sum_{k=1}^{\infty}(a_{N+1})r^{k-N-1} = \sum_{k=1}^{\infty}\left(a_{N+1}r^{-N-1}\right)r^k = \sum_{k=1}^{\infty}Cr^k$$

(where $C = \left(a_{N+1}r^{-N-1}\right)$) is a convergent geometric series. You should now be able to do a comparison by use of Lemma 6.21.

For (b) show $\rho > 1$ implies $\lim_{k\to\infty} a_k \neq 0$. □

The following shows that if the ratio test works, then the root test will also work.

Proposition 6.24. *Let $\langle a_n\rangle_{n=1}^{\infty}$ be a sequence of positive real numbers such that*

$$\lim_{n\to\infty}\frac{a_n+1}{a_n} = r$$

exists. Then also

$$\lim_{n\to\infty} a_n^{1/n} = r.$$

Problem 6.11. Prove this by filling in the details of the following outline of a proof. Let $\varepsilon > 0$. We start by using the same idea as the proof of 6.23. We choose an N such that

$$n \geq N \quad \text{implies} \quad \left|\frac{a_{n+1}}{a_n} - r\right| < \frac{\varepsilon}{2},$$

which implies

$$r - \frac{\varepsilon}{2} < \frac{a_{n+1}}{a_n} < r + \frac{\varepsilon}{2}.$$

Show for $n > N$

$$a_n = a_N\left(\frac{a_{N+1}}{a_N}\right)\left(\frac{a_{N+2}}{a_{N+1}}\right)\left(\frac{a_{N+3}}{a_{N+2}}\right)\cdots\left(\frac{a_{n-1}}{a_{n-2}}\right)\left(\frac{a_n}{a_{n-1}}\right)$$

and therefore

$$a_N\left(r-\frac{\varepsilon}{2}\right)^{n-N} < a_n < a_N\left(r+\frac{\varepsilon}{2}\right)^{n-N}$$

Taking n-th roots

$$a_N^{1/n}\left(r-\frac{\varepsilon}{2}\right)^{1-N/n} < a^{1/n} < a_N^{1/n}\left(r+\frac{\varepsilon}{2}\right)^{1-N/n}.$$

But

$$\lim_{n\to\infty} a_N^{1/n}\left(r-\frac{\varepsilon}{2}\right)^{1-N/n} = a_N^0\left(r-\frac{\varepsilon}{2}\right)^{1-0} = \left(r-\frac{\varepsilon}{2}\right).$$

This implies there is $N_1 > N$ such that

$$n \geq N_1 \quad \text{implies} \quad \left|a_N^{1/n}\left(r-\frac{\varepsilon}{2}\right)^{1-N/n} - \left(r-\frac{\varepsilon}{2}\right)\right|$$

which in turn implies

$$r - \varepsilon < a_N^{1/n}\left(r-\frac{\varepsilon}{2}\right)^{1-N/n}.$$

Do a similar argument to show there is a $N_2 > N$ such that

$$n \geq N_2 \quad \text{implies} \quad a_N^{1/n}\left(r + \frac{\varepsilon}{2}\right)^{1-N/n} < r + \varepsilon.$$

Set $N_3 = \max\{N_1, N_2\}$ and put the inequalities above together to get

$$n \geq N_3 \quad \text{implies} \quad \left|a_n^{1/n} - r\right| < \varepsilon$$

which finishes the proof. □

Problem 6.12. Here are a couple of applications of Proposition 6.24.

(a) For n a positive integer let

$$a_n = \frac{n!}{n^n}.$$

Show

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \frac{1}{e}.$$

Use this and Proposition 6.24 to show

$$\lim_{n\to\infty} \frac{(n!)^{1/n}}{n} = \frac{1}{e}.$$

(b) Let

$$b_n = \binom{2n}{n} = \frac{(2n)!}{n!\,n!}.$$

Show

$$\lim_{n\to\infty} \frac{b_{n+1}}{b_n} = 4$$

and use this to show

$$\lim_{n\to\infty} \binom{2n}{n}^{1/n} = 4.$$

□

6.6. **Absolutely and conditionally convergent series.**

Definition 6.25. The series $\sum_{k=1}^{\infty} a_k$ is ***absolutely convergent*** iff the series of absolute values $\sum_{k=1}^{\infty} |a_k|$ is convergent. □

Theorem 6.26. *If $\sum_{k=1}^{\infty} a_k$ is absolutely convergent, then it is convergent and*

$$\left|\sum_{k=1}^{\infty} a_k\right| \leq \sum_{k=1}^{\infty} |a_k|.$$

Problem 6.13. Prove this. *Hint:* Proposition 6.5 and the triangle inequality applied to partial sums. □

This, together with Proposition 6.16 implies

Proposition 6.27. *Let $\sum_{k=1}^{\infty} a_k$ and $\sum_{k=1}^{\infty} b_k$ be series with $|a_k| \leq Cb_k$ for some positive constant C. Assume $\sum_{k=1}^{\infty} b_k$ converges. Then $\sum_{k=1}^{\infty} a_k$ converges absolutely.* □

Example 6.28. The last proposition implies all the following

$$\sum_{k=1}^{\infty} \frac{\cos(k)}{k^2}, \qquad \sum_{k=1}^{\infty} \frac{(-1)^k}{n2^n}, \qquad \sum_{k=1}^{\infty} \frac{3+(-1)^k}{(k+1)\ln^2(k+1)}.$$

converge absolutely. □

Definition 6.29. The series $\sum_{k=1}^{\infty} a_k$ is ***conditional convergent*** iff $\sum_{k=1}^{\infty} a_k$ converges, but $\sum_{k=1}^{\infty} |a_k| = \infty$. □

The following gives one of the main methods of producing conditional convergent series.

Theorem 6.30. *Let $\langle a_k \rangle_{k=1}^{\infty}$ be a sequence of real numbers with*

(a) $a_k \geq a_{k+1}$ *(that is it is monotone decreasing),*
(b) $\lim_{k\to\infty} a_k = 0$.

Then

$$\sum_{k=1}^{\infty} (-1)^{k+1} a_k = a_1 - a_2 + a_3 - a_4 + \cdots$$

converges. If $A = \sum_{k=1}^{\infty}(-1)^{k+1}a_k$ is the sum and $A_n = \sum_{k=1}^{n} a_k$ is the n-th partial sum then

$$|A - A_n| \leq a_{n+1}.$$

That is the error at stopping at the n-th term is at most the $(n+1)$-st term.

Problem 6.14. Prove this. *Hint:* Note

$$A_3 = A_1 - a_2 + a_3 = A_1 - (a_2 - a_3) \leq A_1$$

as $a_2 \geq a_3$. Likewise

$$A_5 = A_3 - a_4 + a_5 = A_3 - (a_4 - a_5) \leq A_3$$

as $a_4 \geq a_5$. In general

$$A_{2m+3} = A_{2m+1} - (a_{2m} - a_{2m+1}) \leq A_{2m+1}$$

Give an analogous argument to show

$$A_{2m+2} = A_{2m} + (a_{2m+1} - a_{2m+2}) \geq A_{2m}.$$

Now use this to show that if $\ell \geq n$ then for n odd

$$A_{n+1} \leq A_\ell \leq A_n$$

and for n even

$$A_n \leq A_\ell \leq A_{n+1}.$$

Therefore if $\ell \geq n$ the partial sum A_ℓ is between A_n and A_{n+1}. Also show $|A_{n+1} - A_n| = a_{n+1}$. It should not be hard to finish from here. □

Problem 6.15. Show if $0 < p \leq 1$ the series

$$\sum_{k=1}^{\infty} \frac{(-1)^k}{k^p}$$

is conditional convergent. □

Therefore when $0 < p \leq 1$ (which implies $\sum_{k=1}^{\infty} \frac{1}{k^p}$ diverges) the series $\sum_{k=1}^{\infty} \frac{(-1)^k}{k^p}$ is conditionally convergent.

6.7. **Power series.**

Theorem 6.31. *Let $a_0, a_1, a_2, \ldots$ be a sequence of numbers and let $f(x)$ be defined on $\mathbb{R}$ by*

$$f(x) = \sum_{k=0}^{\infty} a_k x^k$$

for all x where this converges. If the series converges for $x = x_0$, then it converges absolutely for all x with $|x| < |x_0|$.

Problem 6.16. Prove this. *Hint:* As

$$f(x_0) = \sum_{k=0}^{\infty} a_k (x_0)^k$$

converges we have $\lim_{k\to\infty} a_k(x_0)^k = 0$ by Theorem 6.3. This implies that $\langle a_k(x_0)^k \rangle_{k=0}^{\infty}$ is bounded. So there is a constant C with

$$|a_k(x_0)^k| = |a_k||x_0|^k \leq C.$$

Then for $|x| < |x_0|$ we have

$$|a_k x^k| = |a_k||x|^k = |a_k||x_0|^k \left(\frac{|x|}{|x_0|}\right)^k \leq C \left(\frac{|x|}{|x_0|}\right)^k = Cr^k$$

where

$$r = \frac{|x|}{|x_0|} < 1.$$

□

Lemma 6.32. *Let $f(x)$ be as in the last theorem. If the series for $f(x)$ converges at $x = x_0$, then the series*

$$f^*(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}$$

*converges absolutely for all x with $|x| < |x_0|$. We call f^* the **formal derivative** of f as it is what the derivative would be if we knew that we could take it term at a time. (Shortly we will show that this the actual derivative.)*

Problem 6.17. Prove this. *Hint:* With notation as in Problem 6.16 show

$$|k a_k x^{k-1}| \leq kCr^{k-1}$$

and then show $\sum_{k=1}^{\infty} kCr^{k-1}$ converges by either the root or ratio test. □

Corollary 6.33. *With the same hypothesis as in the last lemma for $|x| < |x_0|$ the series*

$$f^{**}(x) = \sum_{k=2}^{\infty} k(k-1) a_k x^{k-2}$$

*converges absolutely. (This is the **formal second derivative**.)*

Proof. As $|x| < |x_0|$ there is a number r_0 such that $|x| < r_0 < |x_0|$. By the lemma the series $f^*(r_0)$ converges absolutely. But (with what I hope is not confusing notation) $(f^*)^*(x) = f^{**}(x)$ so this corollary follows by applying Lemma 6.32 to f^* (with r_0 replacing x_0). □

Lemma 6.34. *Let k be a positive integer and x, x_1, r_0 real numbers with $|x|, |x_0| < r_0$. Then*

$$\left| \frac{x^k - x_1^k}{x - x_1} - kx_1^{k-1} \right| \le \frac{k(k-1)}{2} {r_0}^{k-2} |x - x_0|.$$

Problem 6.18. Prove this. *Hint:* This is yet anther opportunity to use Taylor's theorem. Let $p(x)$ be any two times differentiable function. By Taylor's theorem

$$p(x) = p(x_1) + p'(x_1)(x - x_1) + \frac{p''(\xi)}{2}(x - x_1)^2$$

where ξ is between x and x_1. This can be rearranged as

$$\frac{p(x) - p(x_1)}{x - x_1} - p'(x_1) = \frac{p''(\xi)}{2}(x - x_1)$$

and so

$$\left| \frac{p(x) - p(x_1)}{x - x_1} - p'(x_1) \right| = \frac{|p''(\xi)|}{2} |x - x_1|.$$

Now consider the special case where $p(x) = x^k$. Then $|p''(\xi)| = k(k-1)|\xi|^{k-2} < k(k-1)r_0^{k-2}$ as ξ is between x and x_1 and $|x|, |x_1| < r_0$. □

Theorem 6.35. *Let $a_0, a_1, a_2, \ldots$ be a sequence of numbers and let $f(x)$ be defined on $\mathbb{R}$ by*

$$f(x) = \sum_{k=0}^{\infty} a_k x^k$$

for all x where this converges. If the series converges for $x = x_0$, then the function $f(x)$ exists and is differentiable for all x with $|x| < |x_0|$ and the derivative is given by the formal derivative

$$f'(x) = f^*(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}.$$

Problem 6.19. Prove this. *Hint:* That $f(x)$ exists for $|x| < |x_0|$ follows from Theorem 6.31. We need so show that if $|x_1| < |x_0|$ that f is differentiable at x_1 and the derivative is $f^*(x_1)$. Choose a number r_0 such that $|x_1| < r_0 < |x_0|$. Let x be such that $|x| < r_0$. Explain why the following hold.

(a) The series for the following all converge absolutely.

$$f(x), \quad f(x_1), \quad f^*(x_1), \quad f^{**}(r_0).$$

(b) We have

$$\frac{f(x) - f(x_1)}{x - x_1} - f^*(x_1) = \sum_{k=1}^{\infty} a_k \left(\frac{x^k - x_1^k}{x - x_1} - kx_1^{k-1} \right)$$

(c) The inequality

$$\left|\frac{f(x)-f(x_1)}{x-x_1} - f^*(x_1)\right| \le C|x-x_1|$$

holds, where

$$C = \frac{1}{2}\sum_{k=2}^{\infty} k(k-1)|a_k|r_0^{k-1} < \infty$$

holds. (Part of the problem is explaining why $C < \infty$. The hint here is that the series for $f^{**}(r_0)$ converges absolutely.)

(d) To finish show

$$f'(x_1) = \lim_{x\to x_1} \frac{f(x)-f(x_1)}{x-x_1} = f^*(x_1).$$

□

Now that we have differentiated we wish to integrate. Note that by Theorem 6.35 if the series $f(x) = \sum_{k=0}^{\infty} a_k x^k$ converges for $x = x_0$, then it is differentiable on the interval $(-|x_0|, ||x_0|)$ and therefore also continuous on this interval. Thus if $|x| < |x_0|$ this implies $\int_0^x f(t)\,dt$ is the integral of a continuous function and thus it exists.

Theorem 6.36. *Let $a_0, a_1, a_2, \ldots$ be a sequence of numbers and let $f(x)$ be defined on $\mathbb{R}$ by*

$$f(x) = \sum_{k=0}^{\infty} a_k x^k$$

for all x where this converges. If the series converges for $x = x_0$, then for any x with $|x| < |x_0|$

$$\int_0^x f(t)\,dt = \sum_{k=0}^{\infty} \frac{a_k}{k+1} x^{k+1} = \sum_{k=1}^{\infty} \frac{a_{k-1}}{k} x^k.$$

That is we can integrate the series for $f(x)$ term at a time.

Problem 6.20. Prove this. *Hint:* Let $F(x)$ be defined to be the ***formal integral*** of $f(x)$. That is

$$F(x) = \sum_{k=0}^{\infty} \frac{a_k}{k+1} x^{k+1}.$$

Choose r_0 with $|x| < r_0 < |x_0|$. Then as the series for $f(x)$ is convergent, its terms are bounded. That is there is a constant C such that

$$|a_k x_0^k| \le C.$$

Then

$$\left|\frac{a_k}{k+1} r_0^{k+1}\right| = \frac{r_0|a_k x_0^k|}{k+1}\left|\frac{r_0}{x_0}\right|^k \le \frac{r_0 C}{k+1}\left|\frac{r_0}{x_0}\right|^k = \frac{C_1}{k+1} r^k \le C_1 r^k$$

where

$$C_1 = r_0 C \qquad \text{and} \qquad r = \left|\frac{r_0}{x_0}\right| < 1.$$

Now

(a) Explain why the series for $F(r_0)$ converges absolutely. *Hint:* Compare the the geometric series $\sum_{k=0}^{\infty} C_1 r^k$.
(b) Explain why $F(x)$ is differentiable on the interval $(-r_0, r_0)$. *Hint:* Theorem 6.35 with x_0 replaced by r_0.
(c) The derivative of $F(x)$ on $(-r_0, r_0)$ is $f(x)$ *Hint:* Theorem 6.35 again.
(d) Finish the proof. *Hint:* Fundamental Theorem of Calculus. □

Now that we know that we can integrate and differentiate power series we can find new series form old ones.

Example 6.37. Find the series for $(1+x)^{-2}$ on the integral $(-1, 1)$. We know

$$(1+x)^{-1} = \frac{1}{1+x} = 1 - x + x^2 - x^3 + x^4 - x^5 + x^6 - \cdots$$

This can be differentiated term at a time to get

$$-(1+x)^{-2} = 0 - 1 + 2x - 3x^2 + 4x^3 - 5x^4 + 6x^5 - \cdots$$

so that

$$(1+x)^{-2} = 1 - 2x + 3x^2 - 4x^3 + 5x^4 - 6x^5 - \cdots = \sum_{k=0}^{\infty} (-1)^k (k+1) x^k.$$

□

Similar examples can be done by integrating term at a time. Here are some for you to try.

Problem 6.21. (a) Find a series for $\ln(1+x)$ valid on $(-1, 1)$. *Hint:*

$$\ln(1+x) = \int_0^x \frac{dt}{1+t}$$

and you know how to expand $1/(1+t)$ in a series.
(b) For any positive integer n find the series for $(1+x)^{-n}$ valid on $(-1, 1)$.
(c) On $(-1, 1)$ we have the convergent geometric series:

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + x^8 - x^{10} + \cdots$$

Use this to find a power series for $\arctan(x)$ valid on $(-1, 1)$. □

6.8. **The product of two series.** Consider two power series

$$f(x) = \sum_{k=0}^{\infty} a_k x^k$$

$$g(x) = \sum_{k=0}^{\infty} b_k x^k$$

It we assume that we can multiply these the same way we would polynomials we get

$$\begin{aligned} f(x)g(x) &= \left(a_1 + a_1 x + a_2 x^2 + a^3 x^3 + \cdots\right)\left(b_1 + b_1 x + b_2 x^2 + b^3 x^3 + \cdots\right) \\ &= a_0 b_1 + (a_0 b_1 + a_1 b_0)\, x + (a_0 b_2 + a_1 b_1 + a_2 b_0)\, x^2 + \cdots \\ &= \sum_{k=0}^{\infty} \left(\sum_{j=0}^{k} a_k b_{k-j} \right) x^k \end{aligned}$$

Problem 6.22. Here is anther way to see this. Let

$$h(x) = f(x)g(x).$$

Then the first couple of derivatives of h are

$$\begin{aligned} h'(x) &= f'(x)g(x) + f(x)g'(x) \\ h''(x) &= f''(x)g(x)f'(x)g'(x) + f(x)g''(x) \\ h'''(x) &= f'''(x)g(x) + 3f''(x)g'(x) + 3f'g''(x) + g'''(x) \end{aligned}$$

which reminds use of the Binomial Theorem.

(a) Prove that k-th derivative of $h(x)$ is

$$h^{(k)}(x) = \sum_{j=0}^{k} \binom{k}{j} f^{(j)}(x) g^{(k-j)}(x).$$

(b) Let

$$a_k = \frac{f^{(k)(0)}}{k!} \quad \text{and} \quad b_k = \frac{g^{(k)(0)}}{k!}$$

and use the formula for $h^{(k)}(0)$ to show

$$\frac{h^{(k)}(0)}{k!} = \sum_{j=0}^{k} a_j b_{k-j}.$$

□

If we assume that both series for $f(x)$ and $g(x)$ both converge for $x = 1$ we can let $x = 1$ the result is

$$\left(\sum_{k=0}^{\infty} a_k\right)\left(\sum_{n=0}^{\infty} b_n\right) = \sum_{k=0}^{\infty}\left(\sum_{j=0}^{k} a_k b_{k-j}\right).$$

This motivates:

Definition 6.38. Let

$$\sum_{k=0}^{\infty} a_k \qquad \sum_{k=0}^{\infty} b_k$$

be two series then the ***Cauchy product*** of these series is the series

$$\sum_{k=0}^{\infty} c_k$$

where

$$c_k = \sum_{j=0}^{k} a_j b_{k-j} = \sum_{i+j=k} a_i b_j.$$

□

Theorem 6.39. *Let*

$$A = \sum_{k=0}^{\infty} a_k \qquad B = \sum_{k=0}^{\infty} b_k$$

we convergent series with at least one of the two absolutely convergent. Let

$$c_k = \sum_{j=0}^{k} a_j b_{k-j}.$$

Then the series $\sum_{k=0}^{\infty} c_k$ *converges and*

$$\sum_{k=0}^{\infty} c_k = AB.$$

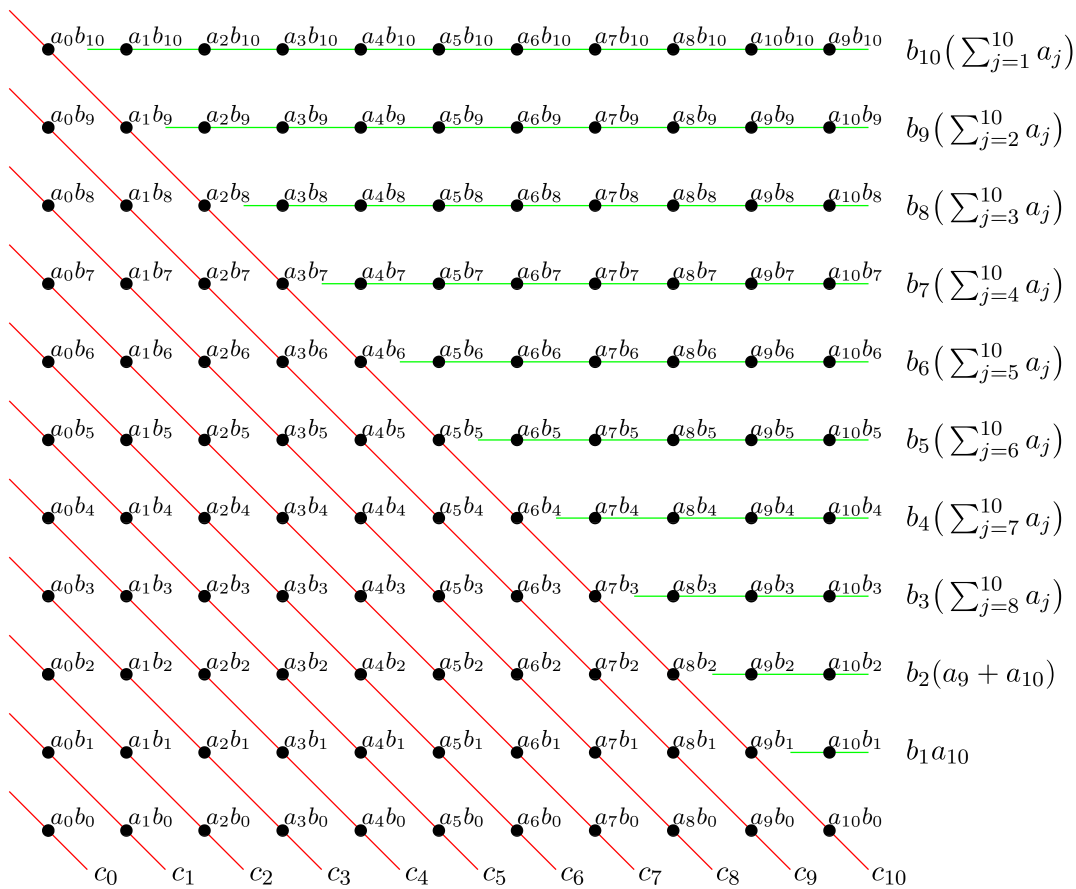
Problem 6.23. Prove this along the following lines. Let

$$A_n = \sum_{k=0}^{n} a_k \qquad B_n = \sum_{k=0}^{n} b_k \qquad C_n = \sum_{k=0}^{n} c_k$$

be the partial sums. Note that the product

$$A_n B_n = (a_0 + a_1 + \cdots + a_n)(b_1 + b_2 + \cdots + b_n)$$

is the sum of the $(n+1)^2$ products $a_j b_k$ with $0 \leq j, k \leq n$. For $n = 10$ these terms are shown in the following figure.



(a) Using the figure above as a guide show for all $n = 1, 2, \cdots$ that

$$A_nB_n - C_n = \sum_{k=1}^{n} b_k \left(\sum_{j=n-k+1}^{n} a_j \right).$$

(b) Let $0 < m < n$ (the integer m will be chosen later). Explain why

$$\begin{aligned}
|A_nB_n - C_n| &= \left| \sum_{k=1}^{n} b_k \left(\sum_{j=n-k+1}^{n} a_j \right) \right| \\
&\le \sum_{k=1}^{n} |b_k| \left| \sum_{j=n-k+1}^{n} a_j \right| \\
&= \sum_{k=1}^{m} |b_k| \left| \sum_{j=n-k+1}^{n} a_j \right| + \sum_{k=m+1}^{n} |b_k| \left| \sum_{j=n-k+1}^{n} a_j \right|
\end{aligned}$$

(c) Without loss of generality we may assume that $\sum_{k=0} b_k$ is absolutely convergent. Explain why there is constant $\beta \geq 0$ such that for all m

$$\sum_{k=1}^{m} |b_k| \leq \beta.$$

(d) The series $\sum_{j=1}^{\infty} a_j$ is convergent. That is $\lim_{n\to\infty} A_n$ exists. Show this implies there is a constant C such that $|A_n| \leq C$ for all n and then use

$$\sum_{j=n-k+1}^{n} a_j = A_n - A_{n-k}$$

to show there is a constant $\alpha \geq 0$ such that

$$\left| \sum_{j=n-k+1}^{n} a_j \right| \leq \alpha$$

for all n and k with $0 \leq k \leq n$.

(e) Combine parts (b), (c), and (d) to show

$$\begin{aligned} |A_nB_n - C_n| &\leq \beta \left| \sum_{j=n-k+1}^{n} a_j \right| + \alpha \sum_{k=m+1}^{n} |b_k| \\ &= \beta|A_n - A_{n-k}| + \alpha \sum_{k=m+1}^{n} |b_k| \end{aligned}$$

when $0 \leq k \leq m \leq n$.

(f) Let $\varepsilon > 0$. Explain where are are $N_1, N_2 > 0$ such that

$$m \geq N_1 \quad \text{implies} \quad \sum_{k=m+1}^{n} |b_k| < \frac{\varepsilon}{2\alpha},$$

and

$$n \geq N_2 \text{ and } n - k \geq N_2 \qquad \text{implies} \qquad |A_n - A_{n-k}| < \frac{\varepsilon}{2\beta}.$$

(g) Let $n \geq N_1 + N_2 + 2$, set $m = N_1$, and show that for any k with $0 \leq k \leq m$ that the inequalities

$$m \geq N_1, \quad n \geq N_2, \quad n - k \geq N_2$$

all hold and that this in turn yields

$$n \geq N_1 + N_2 + 2 \quad \text{implies} \quad |A_nB_n - C_n| < \varepsilon.$$

(h) Conclude from part (f) that

$$\lim_{n\to\infty} (A_nB_n - C_n) = 0.$$

(i) Complete the proof by showing

$$\lim_{n\to\infty} C_n = AB.$$

□

Problem 6.24. Here is an example to show that it is important that in Theorem 6.39 at least one of the two series is absolutely convergent. Let $\sum_{k=0}^\infty a_k = \sum_{k=0}^\infty b_k = \sum_{k=0}^\infty (-1)^k/\sqrt{k+1}$. The Cauchy product is $\sum_{k=0}^\infty c_k$ where

$$c_k = (-1)^k \sum_{j=0}^k \frac{1}{\sqrt{(j+1)(k-j+1)}}.$$

Show

$$|c_k| = \sum_{j=0}^k \frac{1}{\sqrt{(j+1)(k-j+1)}} \geq 1$$

and therefore the series $\sum_{k=0}^\infty c_n$ diverges. □

Theorem 6.40. *Let*

$$f(x) = \sum_{k=0}^\infty a_k x^k, \qquad g(x) = \sum_{k=0}^\infty b_k x^k$$

be power series with radius convergence at least R. Let

$$c_n = \sum_{j=0}^n a_j b_{n-j}$$

then the power series

$$h(x) = \sum_{n=0}^\infty c_n x^n$$

also has radius of convergence at least R and

$$h(x) = f(x)g(x)$$

for $|x| < R$.

Proof. This follows easily form Theorem 6.39. □

We now give a short indication of how to divide power series. Assume that we wish to find the power series expansion of

$$f(x) = \frac{h(x)}{g(x)}$$

where

$$h(x) = \sum_{k=0}^\infty c_k x^k \qquad g(x) = \sum_{k=0}^\infty b_k x^k.$$

and we wish to find the series for $f(x)$. Assume that $f(x)$ has an expansion

$$f(x) = \sum_{k=0}^\infty a_k x^k.$$

Then $h(x) = f(x)g(x)$ and so we have

$$c_k = \sum_{j=0}^{k} a_j b_{k-j}.$$

Assume that $g(0) \neq 0$, that is $b_0 \neq 0$. Then the last equation can be rewritten as

$$a_k = \frac{1}{b_0}\left(c_k - \sum_{j=0}^{n-1} a_j b_{k-j}\right).$$

For small values of k these formulas are

$$\begin{aligned}
a_0 &= \frac{c_0}{b_0} \\
a_1 &= \frac{1}{b_0}\big(c_1 - a_0 b_1\big) \\
a_2 &= \frac{1}{b_0}\big(c_2 - a_0 b_2 - a_1 b_1\big) \\
a_3 &= \frac{1}{b_0}\big(c_3 - a_0 b_3 - a_1 b_2 - a_2 b_1\big) \\
a_4 &= \frac{1}{b_0}\big(c_4 - a_0 b_4 - a_1 b_3 - a_2 b_2 - a_3 b_1\big)
\end{aligned}$$

This allows us to find the coefficients $a_0, a_1, a_2, \ldots$ of $f(x)$ recursively. Unfortunately this method does not tell us anything about the radius of convergence of $f(x)$ in terms of the radii of convergence of $g(x)$ and $h(x)$. But if we already know that all three have positive radius of convergence, it does give us a method for finding the coefficients of $f(x)$ from the coefficients of $g(x)$ and $h(x)$.

Problem 6.25. Find the first three nonzero terms in the power series of

$$f(x) = \frac{e^{2x}}{\cos(x)}.$$

□

Problem 6.26. Find the first couple terms of the power series of the following and thus convince yourself that using series tells you more than using L'Hôspital's rule.

(a) $\dfrac{\sin(2x)}{4x}$

(b) $\dfrac{1-\cos(5x)}{x^2}$

(c) $\dfrac{e^x - 1 - x}{1-\cos(2x)}$. □

Problem 6.27. Find the power series of the following functions around the indicated points x_0.

(a) $f(x) = \sin(x)$ around $x_0 = \pi/4$.

(b) $f(x) = e^{2x}$ around the point $x_0 = 1$.

(c) $f(x) = \sqrt{4-x}$ around the point $x = 4$.

7. Uniform convergence

7.1. **Uniform Convergence and uniform limits of continuous functions.** For convergence of sequences of functions between metric spaces there are (at least) two natural notions of convergence.

Definition 7.1. Let $f_1, f_2, f_2, \ldots X \to Y$ be a sequence of functions between metric spaces and $f\colon X \to Y$. Then $\lim_{n\to\infty} f_n = f$ ***pointwise*** (or f is the ***pointwise limit*** of the sequence $\langle f_n \rangle_{n=1}^{\infty}$ if and only if for for each $x \in X$

$$\lim_{n\to\infty} f_n(x) = f(x).$$

□

Definition 7.2. Let $f_1, f_2, f_2, \ldots X \to Y$ be a sequence of functions between metric spaces and $f\colon X \to Y$. Then $\lim_{n\to\infty} f_n = f$ ***uniformly*** (or f is the ***uniform limit*** of the sequence $\langle f_n \rangle_{n=1}^{\infty}$ if and only if for all $\varepsilon > 0$ there is a N such that

$$n \ge N \qquad d_Y(f_n(x), f(x)) < \varepsilon \qquad \text{for all } x \in X.$$

□

Problem 7.1. Prove that if a sequence converges uniformly, then it converges pointwise. □

Theorem 7.3. *Let $f_n\colon X \to Y$, $n = 1, 2, \ldots$ be a sequence of continuous functions which converge uniformly to the function f. Then f is also continuous.*

That is *the uniform limit of continuous functions is continuous.*

Problem 7.2. Prove this. *Hint:* Let $x_0 \in X$ and we want to show that f is continuous at x_0. Let $\varepsilon > 0$ Then there is an integer n such that

$$d_Y(f(x), f_n(x)) < \frac{\varepsilon}{3} \qquad \text{for all } x \in X.$$

As f_n is continuous at x_0 there is a $\delta > 0$ such that

$$d_X(x, x_0) < \delta \quad \text{implies} \quad d_Y(f_n(x), f_n(x_0)) < \frac{\varepsilon}{3}.$$

Use the triangle inequality to show

$$d_Y(f(x), f(x_0)) \le d_Y(f(x), f_n(x_0)) + d_Y(f_n(x), f_n(x_0)) + d_Y(f_n(x_0), f(x_0))$$

and use this to complete the proof.

Theorem 7.3 makes is easy to give examples of pointwise convergent sequences that are not uniformly convergent.

Problem 7.3. Let $f_n\colon [0,1] \to \mathbb{R}$ be

$$f_n(x) = x^n.$$

Show this converges pointwise to a discontinuous function f. Therefore the convergence can not be uniform as if it were the limit would be continuous.

Problem 7.4. Here is a variant on the previous problem. Let $\arctan \mathbb{R} \to (-\pi/2, \pi/2)$ be the inverse of the restriction of the tangent function to $(-\pi/2, \pi/2)$. Find the pointwise limit

$$f(x) = \lim_{n\to\infty} \arctan(nx)$$

and show it is discontinuous and therefore this is also an example of pointwise convergent sequence that is not uniformly convergent. □

Here is anther nice property of uniform convergence.

Theorem 7.4. *Let $f_1, f_2, f_3, \ldots$ be a sequence of Riemann integrable functions on $[a, b]$ and assume that $\lim_{n\to\infty} f_n = f$ where f is also Riemann integrable. Then*

$$\lim_{n\to\infty} \int_a^b f_n(x)\, dx = \int_a^b f(x)\, dx.$$

Problem 7.5. Prove this. *Hint:* First note

$$\left| \int_a^b f_n(x)\, dx - \int_a^b f_n(x)\, dx \right| \le \int_a^b |f_n(x) - f(x)|\, dx$$

so that it is enough to show

$$\lim_{n\to\infty} \int_a^b |f_n(x) - f(x)|\, dx = 0.$$

Toward this end let $\varepsilon > 0$ and choose N so that

$$n \ge N \quad \text{implies} \quad |f_n(x) - f(x)| < \frac{\varepsilon}{b-a} \quad \text{for all } x \in [a, b]$$

and show for $n \ge N$ that

$$\int_a^b |f_n(x) - f(x)|\, dx < \varepsilon.$$

□

Definition 7.5. Let for each $k = 1, 2, 3, \ldots$ let $f_k \colon X \to \mathbb{R}$ be a function. Then the series $\displaystyle\sum_{k=1}^{\infty} f_k(x)$ ***converges uniformly*** if and only if the sequence of partial sums $F_n(x) := \displaystyle\sum_{k=1}^{n} f_k(x)$ converges uniformly. □

The following is probably the most used for showing a series of functions is uniformly convergent.

Theorem 7.6 (Weierstrass M Test)**.** *Let X be a metric space and $f_1, f_2, f_3, \ldots$ be functions $f_k \colon X \to \mathbb{R}$. Assume there are constants M_n so that*

$$|f_n(x)| \le M_n \quad \textit{for all } x \in X$$

and

$$\sum_{k=1}^{\infty} M_k < \infty.$$

Then the series $\sum_{k=1}^{\infty} f_k(x)$ *converges absolutely and uniformly.*

Problem 7.6. Prove this. For each $x \in X$ we have that the series $\sum_{k=1}^{\infty} f_k(x)$ converges absolutely be comparison to the convergent series $\sum_{k=1}^{\infty} M_k$. Therefore the function

$$F(x) = \sum_{k=1}^{\infty} f_k(x)$$

is defined on X. To show uniform convergence let $\varepsilon >$. Then as $\sum_{k=1}^{\infty} M_k$ is convergent there is a N such that

$$\sum_{k=N+1}^{\infty} M_k < \varepsilon.$$

Let $F_n = \sum_{k=1}^{n} f_k$ be the n-th partial sum of the series $\sum_{k=1}^{\infty} f_k$. Let $n \geq N$ and justify the calculatioin

$$\begin{aligned} |F(x) - F_n(x)| &= \left| \sum_{k=n+1}^{\infty} f_k(x) \right| \\ &\leq \sum_{k=n+1}^{\infty} |f_k(x)| \\ &\leq \sum_{k=n+1}^{\infty} M_k \\ &\leq \sum_{k=N+1}^{\infty} M_k \\ &< \varepsilon. \end{aligned}$$

and explain why this completes the proof. □

Problem 7.7. Use the Weierstrass M test show the series

$$S(x) = \sum_{k=1}^{\infty} \frac{\sin(4^k x)}{2^k}$$

converges uniformly and therefore is a continuous function. □

Remark 7.7. The function $S(x)$ in the previous problem is an example, due to Weierstrass, of a continuous that does not have a derivative at any point. □

Theorem 7.8. *Let* $f_k \colon [a, b] \to \mathbb{R}$ *for* $k = 0, 1, 2, \ldots$ *be continuous functions such that the series*

$$F(x) = \sum_{k=0}^{\infty} f_k(x)$$

converges uniformly on $[a, b]$*. Then* $F(x)$ *is continuous and*

$$\int_a^b F(x)\,dx = \sum_{k=0}^{\infty} \int_a^b f_k(x)\,dx.$$

Problem 7.8. Prove this. □

Theorem 7.9. *Let* $f(x)$ *be defined by a power series*

$$f(x) = \sum_{k=0}^{\infty} c_k (x - x_0)^k$$

with a positive radius of convergence R*. Let* $0 < R_0 < R$*. Then series for* $f(x)$ *converges absolutely and uniformly on the interval* $[x_0 - R_0, x_0 + R_0]$*.*

Problem 7.9. Prove this. *Hint:* Let $R_0 < R_1 < R$. Then $\sum_{k=0}^{\infty} c_k R_1^k$ converges and therefore the terms of this series are bounded, say $|c_k R_1^k| \le B$ for some $B > 0$. Let $M_k = B(R_0/R_1)^k$ and show

$$\sum_{k=0}^{\infty} M_k < \infty \qquad \text{and} \qquad |c_k||x - x_0|^k \le M_k$$

hold for all k and $x \in [x_0 - R_0, x_0 + R_0]$. Now use some of the results above to complete the proof. □

Theorem 7.10. *Let* $f(x)$ *be defined by a power series*

$$f(x) = \sum_{k=0}^{\infty} c_k (x - x_0)^k$$

with a positive radius of convergence R*. Show that for* x *with* $|x - x_0| < R$ *that we can compute the integral of* f *by*

$$\int_{x_0}^{x} f(t)\,dt = \sum_{k=0}^{\infty} \frac{c_k (x - x_0)^{k+1}}{k+1} = \sum_{k=1}^{\infty} \frac{c_{k-1}(x - x_0)^k}{k}.$$

Problem 7.10. Prove this using the results above. *Note:* We gave a different proof of this when first talking about power series. This gives anther, and maybe more natural, proof. □

Theorem 7.11. *Let* $f(x)$ *be defined by a power series*

$$f(x) = \sum_{k=0}^{\infty} c_k (x - x_0)^k$$

with a positive radius of convergence R*. Show that for* $|x - x_0| < R$ *the derivative of* f *is given by*

$$f'(x) = \sum_{k=0}^{\infty} k c_k (x - x_0)^{k-1} = \sum_{k=0}^{\infty} (k+1) c_{k+1} x^k.$$

Problem 7.11. Prove this. *Hint:* This is anther one we have proven before, but with a messy proof. Now that we have more machinery to use we can give an easier proof. Let

$$f^*(x) = \sum_{k=0}^{\infty} kc_k(x - x_0)^{k-1}$$

be the formal derivative of $f(x)$. Then we have shown this has the same radius of convergence as the original series (you can assume this without proof). Now just use that we can integrate this term by term and that by the Fundamental Theorem of Calculus and differentiation are inverse to each other. □

Theorem 7.12. *Let X be a compact metric space and for $k = 1, 2, 3, \ldots$ let $f_k \to X\colon \mathbb{R}$ be continuous and assume for each $x \in X$ that $\langle f_k(x)\rangle_{k=1}^{\infty}$ is monotone decreasing (that is $f_{k+1} \leq f_k(x)$ for all k.) Assume there is a continuous function $f\colon X \to \mathbb{R}$ such that*

$$\lim_{k\to\infty} f_k(x) = f(x)$$

for all $x \in X$. Then $\lim_{k\to\infty} f_k = f$ uniformly.

Problem 7.12. Prove this. *Hint:* Since each sequence $f_1(x), f(2), f_3(x), \ldots$ is monotone decreasing we have that $f_k(x) \geq f(x)$ for all $x \in X$. Therefore to prove uniform convergence we only need show that for each $\varepsilon > 0$ there is a N such that $f_k(x) - f(x) < \varepsilon$ for all $k \geq N$. (For then $0 \leq f_k(x) - f(x) < \varepsilon$ which implies $|f_k - f(x)| < \varepsilon$).

Let $\varepsilon > 0$ and let

$$U_k = \{x \in X : f_k(x) - f(x) < \varepsilon\}.$$

Explain why U_k is open and show $U_k \subseteq U_{k+1}$. Use $\lim_{k\to\infty} f_k(x) = f(x)$ to show

$$\bigcup_{k=1}^{\infty} U_k = X.$$

Finally use compactness to finish the proof. □

7.2. **Uniform approximation of continuous functions by polynomials and other smooth functions.** Given a function, f, if is often useful to it is often useful to approximate it by "nicer" functions. For example give a continuous function, f, it can be useful to find a sequence of differentiable functions $f_1, f_2, f_3, \ldots$ that converge to f uniformly. Here we give one of the basic methods for doing this.

Definition 7.13. A sequence of functions $K_1, K_2, K_3, \ldots$ defined on $\mathbb{R}$ is a ***Dirac sequence***, or an ***approximation to the identity*** iff it satisfies the following conditions.

(a) $K_n \geq 0$ for all k,

(b) For all n

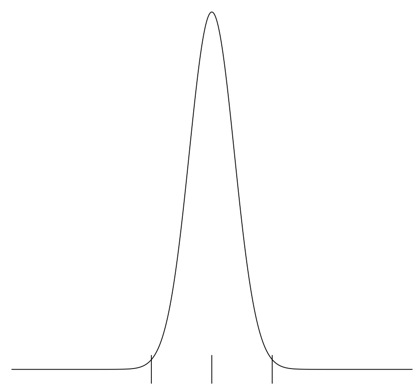
$$\int_{-\infty}^{\infty} K_n(x)\,dx = 1.$$

(c) For all $\delta > 0$

$$\lim_{n\to\infty} \int_{|x|\geq\delta} K_n(x)\,dx = 0.$$

□

The condition (c) say that all most all of the mass of K_n is in $(-\delta, \delta)$.



For large n almost all of the area under the graph of $y = K_n(x)$ is between $-\delta$ and δ.

Here is a standard method of constructing Dirac sequences.

Proposition 7.14. *Let $\varphi\colon \mathbb{R} \to \mathbb{R}$ be a Riemann integrable function with*

$$\varphi \geq 0, \qquad \textit{and} \qquad \int_{-\infty}^{\infty} \varphi(x)\,dx = 1.$$

Then

$$K_n(x) = n\varphi(nx)$$

is a Dirac sequence.

Problem 7.13. Prove this. □

Theorem 7.15. *Let f be a bounded continuous function on $\mathbb{R}$ and $\langle K_n\rangle_{n=1}^{\infty}$ be a Dirac sequence. Let*

$$f_n(x) = \int_{-\infty}^{\infty} f(x-y)K_n(y)\,dy$$

then

$$\lim_{n\to\infty} f_n(x) = f(x)$$

pointwise.

Problem 7.14. Prove this. *Hint:* The basic trick is to note that as $\int_{-\infty}^{\infty} K_n(y)\,dy = 1$ we have

$$f(x) = f(x)\cdot 1 = f(x)\int_{-\infty}^{\infty} K_n(y)\,dy = \int_{-\infty}^{\infty} f(x)K_n(y)\,dy.$$

Therefore for any $\delta > 0$ we have

$$\begin{aligned} f(x) - f_n(x) &= \int_{-\infty}^{\infty} f(x)K_n(y)\,dy - \int_{-\infty}^{\infty} f(x-y)K_n(y)\,dy \\ &= \int_{-\infty}^{\infty} (f(x) - f(x-y))K_n(y)\,dy \\ &= \int_{|y|<\delta} (f(x) - f(x-y))K_n(y)\,dy + \int_{|y|\geq\delta} (f(x) - f(x-y))K_n(y)\,dy \\ &= I_{\delta,n}(x) + J_{\delta,n}(x). \end{aligned} \tag{13}$$

Now let $\varepsilon > 0$. Then as f is continuous at x there is a $\delta > 0$ such that

$$|y - x| < \delta \quad\implies\quad |f(x) - f(y)| < \frac{\varepsilon}{2}.$$

Explain why the following holds

$$|I_{\delta,n}(x)| \leq \int_{|y|<\delta} |f(x) - f(x-y)|K_n(y)\,dy < \int_{|y|<\delta} \left(\frac{\varepsilon}{2}\right) K_n(y)\,dy \leq \frac{\varepsilon}{2}.$$

Using this in the displayed sequence of equalities (13) gives

$$|f(x) - f_n(x)| \leq |I_{\delta,n}(x)| + |J_{\delta,n}(x)| < \frac{\varepsilon}{2} + |J_{\delta,n}(x)|.$$

This holds for all n. The function f is bounded thus there is a constant B such that $|f(x)| \leq B$ for all x. It follows that for all $x, y \in \mathbb{R}$ that

$$|f(x) - f(x-y)| \leq |f(x)| + |f(x-y)| \leq 2B.$$

Therefore

$$|J_{\delta,n}| \leq \int_{|y|\geq\delta} |f(x) - f(x-y)|K_n(y)\,dy \leq 2B\int_{|y|\geq\delta} K_n(y)\,dy.$$

If you now look back at the definition of a Dirac sequence you should be able to use the last inequality to show

$$\lim_{n\to\infty} |J_{\delta,n}(x)| = 0$$

and thus there is a N such that $n > N$ implies $|J_{\delta,n}(x) < \varepsilon/2$. □

We can do a bit better.

Theorem 7.16. *Let f function on $\mathbb{R}$ that is both bounded and uniformly continuous and let $\langle K_n\rangle_{n=1}^{\infty}$ be a Dirac sequence. Define*

$$f_n(x) = \int_{-\infty}^{\infty} f(x-y)K_n(y)\,dy.$$

Then

$$\lim_{n\to\infty} f_n(x) = f(x)$$

uniformly on $\mathbb{R}$.

Problem 7.15. Prove this. *Hint:* This is just a matter of rewriting the proof of Theorem 7.15 and making sure that you can make the choices of quantities such as δ and N in a way that is independent of x. □

The following gives a large number of examples of functions where Theorem 7.16 applies.

Proposition 7.17. *Let f be a continuous function such that for some interval $[\alpha, \beta]$ we have $f(x) = 0$ for all $x \notin [\alpha, \beta]$. Then f is bounded and uniformly continuous.*

Problem 7.16. Prove this. *Hint:* This is a good problem to review several of the results we have been working with. (Continuous functions on closed bounded intervals are bounded and uniformly continuous). □

Proposition 7.18. *Let f be bounded and continuous on $\mathbb{R}$ and let $\langle K_n\rangle_{n=1}^{\infty}$ be a Dirac sequence and*

$$f_n(x) = \int_{-\infty}^{\infty} f(x-y)K_n(y)\,dy.$$

Then f_n can be rewritten as

$$f_n(x) = \int_{-\infty}^{\infty} f(y)K_n(x-y)\,dy$$

Problem 7.17. Prove this. *Hint:* As far as y is concerned, x is a constant. So if we do the change of variable $z = x - y$ we have $dz = -dy$. □

Remark 7.19. In what follows we will use whichever formula for f_n given by Proposition 7.18 that it convenient without refering Proposition 7.18.

We are now in a position to prove one of the most famous theorems in analysis, the *Weierstrass Approximation Theorem*, which says that a continuous function on a closed bounded interval can be uniformly approximated by a polynomial. To start we need a Dirac sequence that is constructed from polynomials.

Lemma 7.20. *Let*

$$K_n(x) := \begin{cases} c_n(1-x^2)^n, & |x| \le 1; \\ 0, & |x| > 0. \end{cases}$$

where

$$c_n := \frac{1}{\displaystyle\int_{-1}^{1} (1-x^2)^n\,dx}.$$

Then $\langle K_n\rangle_{n=1}^{\infty}$ is a Dirac sequence.

Proof. That $K_n \geq 0$ and $\int_{-\infty}^{\infty} K_n(x)\,dx = 1$ are easy, so it remains to show that for $\delta > 0$ the limit $\lim_{n\to\infty} \int_{|x|\geq\delta} K_n(x)\,dx = 0$. We first give a bound on c_n.

$$\int_{-1}^{1}(1-x^2)^n\,dx = 2\int_0^1 (1+x)^n(1-x)^n\,dx \geq 2\int_0^1 (1-x)^n\,dx = \frac{2}{n+1}.$$

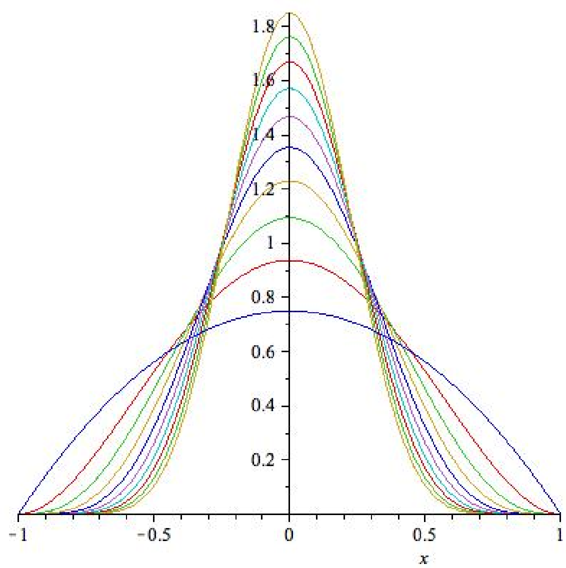
Thus

$$c_n \leq \frac{n+1}{2}.$$

Let $0 < \delta < 1$. Then

$$\int_{|x|\geq\delta} K_n(x)\,dx = 2c_n\int_\delta^1 (1-x^2)^n\,dx \leq 2c_n\int_\delta^1 (1-\delta)^n\,dx \leq (n+1)(1-\delta^2)^n.$$

But $(1-\delta^2) < 1$ so $\lim_{n\to\infty}(n+1)(1-\delta^2)^n = 0$ which completes the proof. □



The graphs of $y = K_n(x)$ for $n = 1, 2, \ldots, 10$.

Problem 7.18. While the exact value of $\int_{-1}^{1}(1-x^2)^n\,dx$ is not needed in the last proof, it is fun to compute it. So find $\int_{-1}^{1}(1-x^2)^n\,dx$. *Hint:* This is a case where it pays to generalize. Let

$$I(m,n) := \int_{-1}^{1}(1-x)^m(1+x)^n\,dx.$$

Then we are trying to compute $I(n,n)$. Use integration by parts to show

$$I(m,n) = \frac{m}{n+1}I(m-1,n+1)$$

when $m \geq 1$ and $n \geq 0$ and note that $I(0,k) = \int_{-1}^{1}(1+x)^k\,dx$ is easy to compute. □

Proposition 7.21. *Let* $f\colon \mathbb{R} \to \mathbb{R}$ *be a continuous function so such that* $f(x) = 0$ *for all* $x \notin [0,1]$ *and let* K_n *be as in Lemma 7.20. Set*

$$p_n(x) = \int_{-1}^{1} K_n(x-y) f(y)\, dy$$

then $p_n \to f$ *uniformly and the restriction of* p_n *to* $[0,1]$ *is a polynomial.*

Proof. By Proposition 7.17 f is bounded and uniformly continuous. Let B be a bound for f, that is $|f(x)| \leq B$ for all $x \in \mathbb{R}$. As f is uniformly continuous for $\varepsilon > 0$ here is a $\delta > 0$, such that

$$|x-y| \leq \delta \text{ and } x, y \in [0,1] \qquad \Longrightarrow \qquad |f(x) - f(y)| \leq \varepsilon.$$

As f is bounded and uniformly continuous, Theorem 7.16 implies $p_n \to f$ uniformly. All that remains is to show that p_n restricted to $[0,1]$ is a polynomial. If $x, y \in [0,1]$, then $x - y \in [-1,1]$ and therefore

$$\begin{aligned} K_n(x-y) &= c_n(1-(x-y)^2)^n \\ &= g_0(y) + g_1(y)x + g_2(y)x^2 + \cdots + g_{2n}(y)x^{2n} \\ &= \sum_{k=0}^{2n} g_k(y)x^k \end{aligned}$$

where we have just expanded $c_n(1-(x-y)^2)^n$ and grouped by powers of x. (Each $g_k(y)$ is a polynomial in y, but this does not really matter for us.) As $f(y) = 0$ for $y \notin [0,1]$ if $x \in [0,1]$ we have

$$\begin{aligned} f_n(x) &:= \int_0^1 K_n(x-y) f(y)\, dy \\ &= \int_0^1 \sum_{k=0}^{2n} g_k(y) x^k f(y)\, dy \\ &= \sum_{k=0}^{2n} \left(\int_0^1 g_k(y)\, dy \right) x^k \end{aligned}$$

which is clearly a polynomial. □

Lemma 7.22. *Let* $f\colon [\alpha, \beta] \to \mathbb{R}$ *be a continuous function with* $f(x) = 0$ *for* $x \notin [\alpha, \beta]$*. Define* $F\colon [0,1] \to \mathbb{R}$ *to be the function*

$$F(x) := f\left(\alpha + (\beta - \alpha)x\right)$$

and let $P_n\colon [0,1] \to \mathbb{R}$ *be polynomials such that* $P_n \to F$ *uniformly and set*

$$p_n(x) = P_n\left(\frac{x-\alpha}{\beta-\alpha}\right).$$

Then each p_n *is a polynomial and* $p_n \to f$ *uniformly.*

Problem 7.19. Prove this. *Hint:* This is not hard, so don't be long winded.

Theorem 7.23 (**Weierstrass Approximation Theorem**). *Let* $f\colon [a,b] \to \mathbb{R}$ *be continuous. Then there is a sequence of polynomial* $p_n\colon [a,b] \to \mathbb{R}$ *with* $p_n \to f$ *uniformly.*

Problem 7.20. Prove this. *Hint:* Extend f to $\mathbb{R}$ (we still denote the extended function by f) by

$$f(x) := \begin{cases} 0, & x < a-1; \\ (x-(a-1))f(a), & a-1 \le x < a; \\ f(x), & a \le x \le b; \\ ((b+1)-x)f(b), & b < x \le b+1; \\ 0, & b+1 < x. \end{cases}$$

This is continuous (don't prove this, just draw the picture and say it is clear). Let $\alpha := a-1$ and $\beta = b+1$. Then use Proposition 7.21 and Proposition 7.17 to complete the proof.

We now give some applications of these results.

Problem 7.21. Let $f\colon [a,b] \to \mathbb{R}$ be continuous and assume that

$$\int_a^b f(x)x^n\,dx = 0$$

for all $n = 0, 1, 2, 3, \ldots$. Then show $f(x) = 0$ for all $x \in [a,b]$. *Hint:* Show that $\int_a^b f(x)p(x)\,dx = 0$ all polynomials. Then choose a sequence of polynomials $p_n \to f$ uniformly. Use this sequence to conclude $\int_a^b f(x)^2\,dx = 0$. □

Problem 7.22. Let $f, g\colon [a,b] \to \mathbb{R}$ be continuous functions such that

$$\int_a^b f(x)x^n\,dx = \int_a^b g(x)x^n\,dx$$

for all $n = 0, 1, 2, 3, \ldots$. Show that $f(x) = g(x)$ for $x \in [a,b]$. *Hint:* Reduce this to the last problem. □

Convention. For the rest of this homework $f\colon \mathbb{R} \to \mathbb{R}$ is a function such that for some $b > 0$ we have $f(x) = 0$ for all x with $|x| \ge b$ and f is Riemann integrable on $[-b,b]$ and that there is a constant B such that $|f(x)| \le B$ for all x. □

Theorem 7.24. *If* $\langle K_k \rangle_{n=1}^{\infty}$ *is a Dirac sequence and*

$$f_n(x) = \int_{-\infty}^{\infty} K_n(y)f(x-y)\,dy = \int_{-\infty}^{\infty} K_n(x-y)f(y)\,dy$$

then at any point x *where* f *is continuous*

$$\lim_{n\to\infty} f_n(x) = f(x).$$

Problem 7.23. Prove this. *Hint:* This is an easier version of an earlier theorem.

Definition 7.25. A Dirac sequence $\langle K_n\rangle_{n=1}^{\infty}$ is ***differentiable*** iff for each n K_n is differentiable and

$$\lim_{h\to 0}\frac{K_n(x+h)-K_n(x)}{h}=K_n'(x)$$

uniformly. Explicitly this means that for each n and $\varepsilon>0$ there is a $\delta>0$ such that

$$|h|\le\delta \quad\Longrightarrow\quad \left|\frac{K_n(x+h)-K_n(x)}{h}-K_n'(x)\right|\le\varepsilon \tag{14}$$

for all $x\in\mathbb{R}$ □

Proposition 7.26. *Let f be as in the convention and $\langle K_k\rangle_{n=1}^{\infty}$ a differentiable Dirac sequence. Then for each n*

$$f_n(x)=\int_{-\infty}^{\infty}K_n(x-y)f(y)\,dy$$

is differentiable and

$$f_n'(x)=\int_{-\infty}^{\infty}K_n'(x-y)f(y)\,dy.$$

(It is not being assumed that f is differentiable.)

Problem 7.24. Prove this. *Hint:* First show

$$\begin{aligned}\left(\frac{f_n(x+h)-f_n(x)}{h}\right)&-\int_{-\infty}^{\infty}K_n'(x-y)f(y)\,dy\\ &=\int_{-\infty}^{\infty}\left(\frac{K_n(x-y+h)-K_n(x-y)}{h}-K_n'(x-y)\right)f(y)\,dy\\ &=\int_{-b}^{b}\left(\frac{K_n(x-y+h)-K_n(x-y)}{h}-K_n'(x-y)\right)f(y)\,dy\end{aligned}$$

take absolute values and then use (14).

Lemma 7.27. *Let f be as in the convention and also assume that f is differentiable with f' uniformly continuous and let $\langle K_k\rangle_{n=1}^{\infty}$ be a differentiable Dirac sequence. Then the derivative of*

$$f_n(x)=\int_{-\infty}^{\infty}K_n(x-y)f(y)\,dy$$

can be written as

$$f_n'(x)=\int_{-\infty}^{\infty}K_n(x-y)f'(y)\,dy$$

Problem 7.25. Prove this. *Hint:* Starting with Proposition 7 show

$$\begin{aligned} f_n'(x) &= \int_{-\infty}^{\infty} K_n'(x-y)f(y)\,dy \\ &= -\int_{-\infty}^{\infty} \left(\frac{d}{dy}K_n(x-y)\right) f(y)\,dy \\ &= -\int_{-b}^{b} \left(\frac{d}{dy}K_n(x-y)\right) f(y)\,dy \end{aligned}$$

and use integration by parts along with $f(-b) = f(b) = 0$. □

Theorem 7.28. *Let f be as in the convention and also assume that f is differentiable with f' uniformly continuous and let $\langle K_k\rangle_{n=1}^{\infty}$ be a differentiable Dirac sequence. Then if*

$$f_n(x) = \int_{-\infty}^{\infty} K_n(x-y)f(y)\,dy$$

the limit

$$\lim_{n\to\infty} f_n' = f'$$

holds uniformly.

Problem 7.26. Prove this. □

Department of Mathematics,
University of South Carolina,
Columbia, S.C. 29208, USA
Email address: howard@math.sc.edu
URL: www.math.sc.edu/∼howard