MAT2006: Elementary Real Analysis
Homework 3

W. Lee.

Due date: Tomorrow

Question 5.1-1. Show that

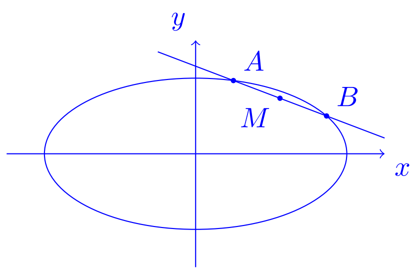
a) the tangent to the ellipse

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

at the point (x_0, y_0) has the equation

$$\frac{xx_0}{a^2}+\frac{yy_0}{b^2}=1$$

Let's use brutal method to deal with this problem. Consider the following ellipse,



We fix $A=(x_0,y_0)$, and take $B=(x_0+h, y_0+h')$, where $(h,h')\neq(0,0)$, so that A and B are both on the ellipse. We also denote the middle point of line segment AB as $C=(x_0+h/2, y_0+h'/2)$. Then we have

$$\begin{cases}\dfrac{x_0^2}{a^2}+\dfrac{y_0^2}{b^2}=1 & (1)\\[2mm] \dfrac{(x_0+h)^2}{a^2}+\dfrac{(y_0+h')^2}{b^2}=1 & (2)\end{cases}$$

Subtract (1) from (2), we obtain

$$\frac{(2x_0+h)h}{a^2}=-\frac{(2y_0+h')h'}{b^2}$$

If the line segment AB has a slope k, then

$$k=\frac{h'}{h}=-\frac{b^2}{a^2}\frac{2x_0+h}{2y_0+h'}$$

Notice that if $2y_0+h'=0$, k does not exist, line segment AB reduces to $x=x_0$. If not, then

$$AB:\ \ y-\left(y_0+\frac{1}{2}h'\right)=k\left[x-\left(x_0+\frac{1}{2}h\right)\right]$$

which can be rewritten as

$$AB:\quad \left[y-\left(y_0+\frac{1}{2}h'\right)\right]a^2(2y_0+h')=-b^2(2x_0+h)\left[x-\left(x_0+\frac{1}{2}h\right)\right]$$

We can see that for the new formula, even if k doesn't exist, i.e., $AB: x=x_0$ is still included, hence this formula can represent all AB as long as A,B are distinct points. To obtain tangent line passing A, by definition, we only need to make B go as closed as A. Hence, take the limit as $(h,h')\to(0,0)$ on both sides, we have

$$(y-y_0)a^2y_0=-b^2x_0(x-x_0)$$

Slightly change the order of the above equation, we have

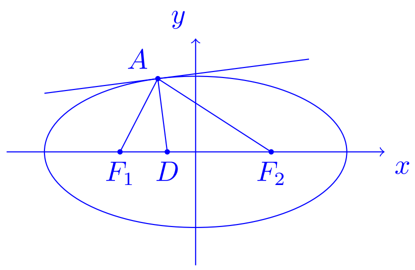
$$\frac{x_0x}{a^2}+\frac{y_0y}{b^2}=\frac{x_0^2}{a^2}+\frac{y_0^2}{b^2}$$

Since (x_0,y_0) is on the ellipse, the right hand side of the above equation is equal to 1, which yields the tangent line

$$\frac{xx_0}{a^2}+\frac{yy_0}{b^2}=1$$

b) light rays from a source located at a focus $F_1=\left(-\sqrt{a^2-b^2},0\right)$ or $F_1=\left(\sqrt{a^2-b^2},0\right)$ of an ellipse with semiaxes $a>b>0$ are gathered at the other focus by an elliptical mirror.

Again, we only use the most intelligible way to achieve our goal. Denote $F_1=(-c,0)$, $F_2=(c,0)$, consider



where $A=(x_0,y_0)$. Draw the tangent line going through A, and AD is the line segment perpendicular to the tangent line, we need to prove $\angle F_1AD=\angle DAF_2$. If $A=(\pm a,0)$, then $\angle F_1AD=\angle DAF_2=0$; if $A=(0,\pm b)$, then $AF_1=AF_2$, so $\angle F_1AD=\angle DAF_2$ is trivial. If $x_0=\pm c$, i.e., the slope of F_1A or F_2A doesn't exist, then we only consider one case, i.e., $A=(-c,b^2/a)$, because the other three cases are equivalent due to the symmetric property of ellipse.

From part a), we have known that the slope of tangent line going through A is $-x_0b^2/y_0a^2$. Denote the slope of AD as $k=y_0a^2/x_0b^2=-a/c$, and denote the slope of F_1A as k_1 (if exists), F_2A as k_2 (if exists). Then we have

$$\tan\angle DAF_2=\frac{k_2-k}{1+kk_2}=\frac{-b^2/2ac+a/c}{1+b^2/2c^2}$$

Substitute $b^2=a^2-c^2$, we have

$$\tan\angle DAF_2=\frac{-b^2/2ac+a/c}{1+b^2/2c^2}=\frac{c}{a}=\tan\angle F_1AD$$

Finally, we consider the remaining cases, where k, k_1, k_2 all exists.

$$k = \frac{y_0 a^2}{x_0 b^2}, \quad k_1 = \frac{y_0}{x_0 + c}, \quad k_2 = \frac{y_0}{x_0 - c}$$

We need to prove $\tan \angle F_1AD = \tan \angle DAF_2$, where

$$\tan \angle F_1AD = \frac{k - k_1}{1 + k_1 k} = \frac{\frac{y_0 a^2}{x_0 b^2} - \frac{y_0}{x_0 + c}}{1 + \frac{y_0 a^2}{x_0 b^2} \frac{y_0}{x_0 + c}}, \qquad \tan \angle DAF_2 = \frac{k_2 - k}{1 + k k_2} = \frac{\frac{y_0}{x_0 - c} - \frac{y_0 a^2}{x_0 b^2}}{1 + \frac{y_0 a^2}{x_0 b^2} \frac{y_0}{x_0 - c}}$$

Thus, we have

$$\begin{aligned}
&\frac{\frac{y_0 a^2}{x_0 b^2} - \frac{y_0}{x_0 + c}}{1 + \frac{y_0 a^2}{x_0 b^2} \frac{y_0}{x_0 + c}} = \frac{\frac{y_0}{x_0 - c} - \frac{y_0 a^2}{x_0 b^2}}{1 + \frac{y_0 a^2}{x_0 b^2} \frac{y_0}{x_0 - c}} \\
\Longleftarrow\ &\left(\frac{y_0 a^2}{x_0 b^2} - \frac{y_0}{x_0 + c}\right)\left(1 + \frac{y_0 a^2}{x_0 b^2} \frac{y_0}{x_0 - c}\right) = \left(1 + \frac{y_0 a^2}{x_0 b^2} \frac{y_0}{x_0 + c}\right)\left(\frac{y_0}{x_0 - c} - \frac{y_0 a^2}{x_0 b^2}\right) \\
\Longleftarrow\ &\frac{x_0}{x_0^2 - c^2} + \frac{a^2}{x_0 b^2} \frac{y_0^2}{x_0^2 - c^2} = \frac{a^2}{x_0 b^2} + \frac{a^4}{x_0 b^4} \frac{y_0^2}{x_0^2 - c^2} \\
\Longleftarrow\ &\frac{x_0^2 b^2 + a^2 y_0^2}{x_0 b^2 (x_0^2 - c^2)} = \frac{a^2 b^2 (x_0^2 - c^2) + a^4 y_0^2}{x_0 b^4 (x_0^2 - c^2)} \\
\Longleftarrow\ &\frac{x_0^2 b^2 + a^2 y_0^2}{b^2} = \frac{a^2 b^2 (x_0^2 - c^2) + a^4 y_0^2}{b^4} \\
\Longleftarrow\ &\frac{a^2 b^2}{b^2} = \frac{a^2 b^2 (x_0^2 - c^2) + a^4 y_0^2}{b^4} \\
\Longleftarrow\ &b^2 (x_0^2 - c^2) + a^2 y_0^2 = b^4 \\
\Longleftarrow\ &a^2 b^2 - b^2 c^2 = b^4
\end{aligned}$$

This shows that $\tan \angle F_1AD = \tan \angle DAF_2$, which means $\angle F_1AD = \angle DAF_2$. Hence, for any (x_0, y_0) on ellipse, $\angle F_1AD = \angle DAF_2$, which completes the proof.

Question 5.1-2. Write the formulas for approximate computation of the following values:

a) $\sin\left(\frac{\pi}{6} + \alpha\right)$ for values of α near 0;

We can use the best linear approximation, i.e.,

$$f(\alpha) = f(0) + f'(0)(\alpha - 0) = \frac{1}{2} + \frac{\sqrt{3}}{2}\alpha$$

b) $\sin(30^\circ + \alpha^\circ)$ for values of α° near 0;

Similarly, we also use the best linear approximation, but be careful about the unit of α.

$$f(\alpha^\circ) = f(0^\circ) + f'(0^\circ)(\alpha^\circ - 0^\circ) = \frac{1}{2} + \frac{\sqrt{3}}{2}\alpha^\circ$$

or equivalently,

$$f(\alpha) = \frac{1}{2} + \frac{\sqrt{3}\pi}{360}\alpha$$

c) $\cos\left(\frac{\pi}{4}+\alpha\right)$ for values of α near 0;

Similar to part a), we have

$$f(\alpha) = f(0) + f'(0)(\alpha - 0) = \frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}\alpha$$

d) $\cos(45^\circ + \alpha^\circ)$ for values of α° near 0.

Similar to part b), we have

$$f(\alpha^\circ) = f(0^\circ) + f'(0^\circ)(\alpha^\circ - 0^\circ) = \frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}\alpha^\circ$$

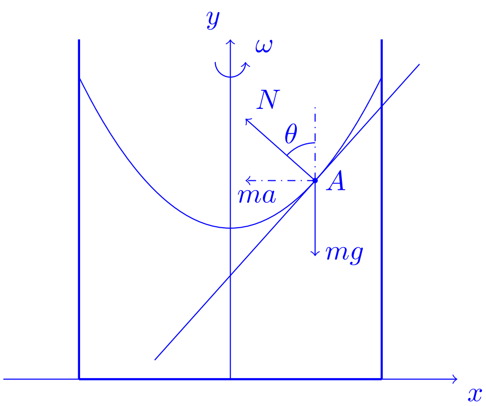
or equivalently,

$$f(\alpha) = \frac{\sqrt{2}}{2} - \frac{\sqrt{2}\pi}{360}\alpha$$

Question 5.1-3. A glass of water is rotating about its axis at constant angular velocity ω. Let $y = f(x)$ denote the equation of the curve obtained by cutting the surface of the liquid with a plane passing through its axis of rotation.

a) Show that $f'(x) = \frac{\omega^2}{g}x$, where g is the acceleration of free fall.

Consider the following free-body diagram



At dynamic equilibrium, for arbitrary point $A = (x, y)$ on the curve, we have

$$\begin{cases} N\sin\theta = m\omega^2 x \\ N\cos\theta = mg \end{cases}$$

Since we also have

$$\tan\theta = \frac{dy}{dx} = \frac{\omega^2}{g}$$

Hence, we have $f'(x) = \frac{\omega^2}{g}x$.

b) Choose a function $f(x)$ that satisfies the condition given in part a).

As we proved in part a), we can just take

$$f(x) = \frac{\omega^2}{2g}x^2$$

Then $f'(x) = \frac{\omega^2}{g}x$.

c) Does the condition on the function $f(x)$ given in part a) change if its axis of rotation does not coincide with the axis of the glass?

No, the condition will change to $f'(x) = \frac{\omega^2}{g}(x+d)$, where $x = d$ ($d \neq 0$) is its axis of rotation and y-axis is the vertical axis of the glass.

Question 5.1-4. A body that can be regarded as a point mass is sliding down a smooth hill under the influence of gravity. The hill is the graph of a differentiate function $y = f(x)$.

a) Find the horizontal and vertical components of the accleration vector that the body has at the point (x_0, y_0).

First, the tangent line at (x_0, y_0) is

$$y - y_0 = f'(x_0)(x - x_0)$$

Denote the included angle of the tangent line and horizontal line as θ, then $\tan\theta = f'(x_0)$. According to the free body diagram, we know that

$$\begin{cases} mg\cos\theta = N \\ mg\sin\theta = ma \end{cases}$$

Thus, we can solve that $a = g\sin\theta$, and the horizontal and vertical components are

$$\begin{cases} a_x = g\sin\theta\cos\theta \\ a_y = g\sin^2\theta \end{cases}$$

Since $\tan\theta = f'(x_0)$, we have

$$\begin{cases} a_x = \dfrac{f'(x_0)}{1+[f'(x_0)]^2}g \\ a_y = \dfrac{[f'(x_0)]^2}{1+[f'(x_0)]^2}g \end{cases}$$

b) For the case $f(x) = x^2$ when the body slides from a great height, find the point of the parabola $y = x^2$ at which the horizontal component of the acceleration is maximal.

Let $f(x) = x^2$, $f'(x) = 2x$. Consider horizontal component

$$a_x = \frac{f'(x_0)}{1+[f'(x_0)]^2}g = \frac{g}{\frac{1}{f'(x_0)} + f'(x_0)}$$

Since the sign of $f'(x_0)$ only indicates the direction of acceleration, we can only consider positive case. In this way, when $f'(x_0) = 1$, a_x will obtain its maximal value $g/2$. In this case, $x_0 = \frac{1}{2}$, and $y_0 = \frac{1}{4}$. Similarly, for negative case, $x_0 = -\frac{1}{2}$, and $y_0 = \frac{1}{4}$.

Question 5.1-5. Set

$$\Psi_0(x) = \begin{cases} x, & \text{if } 0 \leq x \leq \dfrac{1}{2} \\ 1 - x, & \text{if } \dfrac{1}{2} \leq x \leq 1 \end{cases}$$

and extend this function to the entire real line so as to have period 1. We denote the extended function by φ_n. Further, let

$$\varphi_n(x) = \frac{1}{4^n}\varphi_0\left(4^n x\right)$$

The function $\varphi_n(x)$ has period 4^{-n} and a derivative equal to $+1$ or —1 everywhere except at the points $x = \frac{k}{2^n+1}$, $n \in \mathbb{Z}$. Let

$$f(x) = \sum_{n=1}^{\infty} \varphi_n(x)$$

Show that the function f is defined and continuous on $\mathbb{R}$, but does not have a derivative at any point.

It is easy to prove that $\varphi_0(x)$ is continuous on $\mathbb{R}$, because it is continuous on every interval $(n, n+1)$, and we can see $\varphi_0(n+) = \varphi_0(n-) = \varphi_0(0) = 0$ for any $n \in \mathbb{Z}$.

Therefore, since every $\varphi_n(x)$ can be regarded as a composite function of $\varphi_0(x)$ and continuous function, $\varphi_n(x)$ must be continuous on $\mathbb{R}$.

We can easily see that every $\varphi_n(x)$ is bounded, i.e.,$|\varphi_n(x)| \leq 0.5 \cdot 4^{-n}$. Hence, by Weierstrass M-Test, the convergence of $\sum_{n=1}^{\infty} 0.5 \cdot 4^{-n}$ implies that f_k converges to f uniformly, where

$$f_k(x) = \sum_{n=1}^{k} \varphi_n(x)$$

The convergence of f_k shows that f is well-defined on $\mathbb{R}$.

Also, $f_k(x)$ is continuous because it is the sum of finitely many continuous function. In this case, f is also continuous.

Fix a real number x and positive integer m, and put

$$\delta_m = \pm\frac{1}{4} \cdot 4^{-m}$$

where the sign is well-chosen so that no $r/2$ ($r \in \mathbb{Z}$) lies in the interior between $4^m x$ and $4^m(x+\delta_m)$. This can be obtained because $4^m|\delta_m| = \frac{1}{4}$. If x coincides with $r/2$, take positive sign for all m. Define

$$\gamma_n = \frac{\varphi_n(4^n(x+\delta_m)) - \varphi_n(4^n x)}{\delta_m}$$

When $n > m$, $4^n\delta_m$ is an integer, by the periodicity of $\varphi_n(x)$, $\gamma_n = 0$. When $1 \leq n \leq m$,

$$|\gamma_n| = \frac{|\varphi_n(4^n(x+\delta_m)) - \varphi_n(4^n x)|}{|\delta_m|} = \frac{(1/4)^{m+1-n}}{|\delta_m|} = \frac{(1/4)^{m+1-n}}{(1/4)4^{-m}} = 4^n$$

Since $|\gamma_n| = 4^n$, we conclude that

$$\frac{f(x+\delta_m) - f(x)}{\delta_m} = \sum_{n=1}^{\infty} \frac{1}{4^n}\gamma_n = \sum_{n=1}^{m} \frac{1}{4^n}\gamma_n$$

We can see that $\frac{1}{4^n}\gamma_n = \pm 1$, which means

$$\lim_{n\to\infty} \frac{1}{4^n}\gamma_n \neq 0$$

Therefore, the series

$$\sum_{n=1}^{\infty} \frac{1}{4^n}\gamma_n$$

will not converge, and since as $m \to \infty$, $\delta_m \to 0$, the derivative of $f(x)$ at any x does not exist.

Question 5.2-1. Let $\alpha_0, \alpha_1, \ldots, \alpha_n$ be given real numbers. Exhibit a polynomial $P_n(x)$ of degree n having the derivatives $P_n^{(k)}(x_0) = \alpha_k$, $k = 0, 1, \ldots, n$, at a given point $x_0 \in \mathbb{R}$.

Suppose the polynomial of degree n is defined as

$$P_n(x) = \lambda_n x^n + \lambda_{n-1}x^{n-1} + \lambda_1 x + \lambda_0$$

Since $P_n^{(k)}(x_0) = \alpha_k$, $k = 0, 1, \ldots, n$, we have

$$\begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^{n-1} & x_0^n \\ 0 & 1 & 2x_0 & \cdots & (n-1)x_0^{n-2} & nx_0^{n-1} \\ 0 & 0 & 2 & \cdots & (n-1)(n-2)x_0^{n-3} & n(n-1)x_0^{n-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & (n-1)! & n\cdot(n-1)\cdots 2\cdot x_0 \\ 0 & 0 & 0 & \cdots & 0 & n! \end{bmatrix} \cdot \begin{bmatrix} \lambda_0 \\ \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_{n-1} \\ \lambda_n \end{bmatrix} = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{n-1} \\ \alpha_n \end{bmatrix}$$

Notice that the matrix U on the left hand side is an upper triangular matrix, thus, to solve this system, we could use backward substitution (Tom Luo taught this technique in MAT2004). Here $U_{i,j}$ denote the entry of matrix U at i-th row, j-th column.

$$\begin{aligned} \lambda_n &= \frac{\alpha_n}{U_{n+1,n+1}} \\ &\vdots \\ \lambda_m &= \frac{\alpha_m - \sum_{i=m+1}^{n} U_{m+1,i+1}\lambda_i}{U_{m+1,m+1}} \quad \text{for } 0 < m < n \\ &\vdots \\ \lambda_0 &= \frac{\alpha_0 - \sum_{i=1}^{n} U_{1,i+1}\lambda_i}{U_{1,1}} \end{aligned}$$

Therefore, the polynomial $P_n(x)$ with the above coefficients will be the required answer.

Question 5.2-2. Compute $f'(x)$ if

a) $f(x) = \begin{cases} \exp\left(-\dfrac{1}{x^2}\right) & \text{for } x \neq 0 \\ 0 & \text{for } x = 0 \end{cases}$

When $x \neq 0$,

$$f'(x) = \left(\frac{2}{x^3}\right) \exp\left(-\frac{1}{x^2}\right)$$

When $x = 0$,

$$\begin{aligned} f'_+(0) &= \lim_{h \to +0} \frac{f(0+h) - f(0)}{h} \\ &= \lim_{h \to +0} \frac{e^{-1/h^2}}{h} && \text{Take } t = 1/h \\ &= \lim_{t \to +\infty} \frac{t}{e^{t^2}} \\ &= \lim_{t \to +\infty} \frac{1}{2te^{t^2}} = 0 \end{aligned}$$

Similarly,

$$f'_-(0) = \lim_{t \to -\infty} \frac{1}{2te^{t^2}} = 0$$

Hence, $f'(0) = f'_+(0) = f'_-(0) = 0$. In conclusion, $f'(x)$ is given by

$$f'(x) = \begin{cases} \left(\dfrac{2}{x^3}\right) \exp\left(-\dfrac{1}{x^2}\right) & \text{for } x \neq 0 \\ 0 & \text{for } x = 0 \end{cases}$$

b) $f(x) = \begin{cases} x^2 \sin \dfrac{1}{x} & \text{for } x \neq 0 \\ 0 & \text{for } x = 0 \end{cases}$

When $x \neq 0$,

$$f'(x) = 2x \sin \frac{1}{x} + x^2 \cos \frac{1}{x} \left(-\frac{1}{x^2}\right) = 2x \sin \frac{1}{x} - \cos \frac{1}{x}$$

When $x = 0$,

$$\begin{aligned} f'_+(0) &= \lim_{h \to +0} \frac{f(0+h) - f(0)}{h} \\ &= \lim_{h \to +0} \frac{h^2 \sin \frac{1}{h}}{h} \\ &= \lim_{h \to +0} h \sin \frac{1}{h} && \text{Take } t = 1/h \\ &= \lim_{t \to +\infty} \frac{\sin t}{t} = 0 \end{aligned}$$

Similarly,

$$f'_-(0) = \lim_{t \to -\infty} \frac{\sin t}{t} = 0$$

Hence, $f'(0) = f'_+(0) = f'_-(0) = 0$. In conclusion, $f'(x)$ is given by

$$f'(x) = \begin{cases} 2x \sin \dfrac{1}{x} - \cos \dfrac{1}{x} & \text{for } x \neq 0 \\ 0 & \text{for } x = 0 \end{cases}$$

c) Verify that the function in part a) is infinitely differentiable on $\mathbb{R}$, and that $f^{(n)}(0) = 0$.

Since $f(x)$ is differentiable and its derivative has the form of

$$f'(x) = \begin{cases} \left(\frac{2}{x^3}\right) \exp\left(-\frac{1}{x^2}\right) & \text{for } x \neq 0 \\ 0 & \text{for } x = 0 \end{cases}$$

We assume that $f^{(n)}(x)$ exists and has the form of

$$f^{(n)}(x) = \begin{cases} \left[\sum_{k=1}^{n} \frac{a_k}{x^{n+2k}}\right] \exp\left(-\frac{1}{x^2}\right) & \text{for } x \neq 0 \\ 0 & \text{for } x = 0 \end{cases}$$

where a_k is some certain constant (we don't care the exact value, as long as it is fixed for each k).

By induction, we can compute $f^{(n+1)}(x)$. When $x \neq 0$,

$$\begin{aligned} f^{(n+1)}(x) = [f^{(n)}(x)]' &= \left[\sum_{k=1}^{n} \frac{2a_k}{x^{n+2k+3}} + \sum_{k=1}^{n} \frac{-(n+2k)a_k}{x^{n+2k+1}}\right] \exp\left(-\frac{1}{x^2}\right) \\ &= \left[\sum_{k=1}^{n+1} \frac{b_k}{x^{(n+1)+2k}}\right] \exp\left(-\frac{1}{x^2}\right) \end{aligned}$$

where

$$b_k \triangleq \begin{cases} -(n+2k)a_k & k = 1 \\ [2-(n+2k)]a_k & 2 \leq k \leq n \\ 2a_k & k = n+1 \end{cases}$$

We can see that $f^{(n+1)}(x)$ is also differentiable (composition, summation and multiplication of elementary function) and has the same form of our assumption, hence our assumption is true when $x \neq 0$, for arbitrary n.

When $x = 0$, by definition,

$$\begin{aligned} \left[f_+^{(n+1)}\right](0) &= \lim_{h \to +0} \frac{f^{(n)}(0+h) - f^{(n)}(0)}{h} \\ &= \lim_{h \to +0} \frac{\left[\sum_{k=1}^{n} \frac{a_k}{h^{n+2k}}\right] \exp\left(-\frac{1}{h^2}\right)}{h} \\ &= \lim_{h \to +0} \left[\sum_{k=1}^{n} \frac{a_k}{h^{n+2k+1}}\right] \exp\left(-\frac{1}{h^2}\right) \qquad \text{Take } t = 1/h \\ &= \lim_{t \to +\infty} \frac{\sum_{k=1}^{n} a_k t^{n+2k+1}}{e^{t^2}} = 0 \end{aligned}$$

Similarly,

$$f_-^{(n+1)}(0) = \lim_{t \to -\infty} \frac{\sum_{k=1}^{n} a_k t^{n+2k+1}}{e^{t^2}} = 0$$

Hence, $f^{(n+1)}(0) = f_+^{(n+1)}(0) = f_-^{(n+1)}(0) = 0$. In conclusion, our assumption is true for all $x \in \mathbb{R}$ and arbitrary n. This shows that $f(x)$ is real smooth function (infinitely differentiable) and $f^{(n)}(0) \equiv 0$.

d) Show that the derivative in part b) is defined on $\mathbb{R}$ but is not a continuous function on $\mathbb{R}$.

Since the derivative in part b) is defined as

$$f'(x) = \begin{cases} 2x\sin\dfrac{1}{x} - \cos\dfrac{1}{x} & \text{for } x \neq 0 \\ 0 & \text{for } x = 0 \end{cases}$$

It is well-defined on $\mathbb{R}$. Now let's check its continuity at $x = 0$. Consider the right hand side limit of $f'(x)$,

$$\lim_{x\to 0+} f'(x) = \lim_{x\to 0+}\left[2x\sin\frac{1}{x} - \cos\frac{1}{x}\right]$$

We can see $2x\sin\frac{1}{x}$ converges to zero as $x \to 0+$, but $\cos\frac{1}{x}$ diverges because when $x \to 0+$, $\frac{1}{x}$ goes to $+\infty$, and $\cos(+\infty)$ vibrates between $[-1, 1]$. Hence, $f'(x)$ is not continuous at $x = 0$.

e) Show that the function

$$f(x) = \begin{cases} \exp\left(-\dfrac{1}{(1+x)^2} - \dfrac{1}{(1-x)^2}\right) & \text{for } -1 < x < 1 \\ 0 & \text{for } 1 \leq |x| \end{cases}$$

is infinitely differentiable on $\mathbb{R}$.

The original function can be transformed into

$$f(x) = \begin{cases} \exp\left(-\dfrac{2}{(1-x^2)^2}\right) & \text{for } -1 < x < 1 \\ 0 & \text{for } 1 \leq |x| \end{cases}$$

When $-1 < x < 1$, it's easy to see $f(x)$ is differentiable, i.e.,

$$f'(x) = -\frac{8x}{(1-x^2)^3}\exp\left(-\frac{2}{(1-x^2)^2}\right)$$

Assume that when $-1 < x < 1$, $f^{(n)}(x)$ exists, and has the form (which satisfies the case when $n = 1$)

$$f^{(n)}(x) = \sum_{k=1}^{n}\frac{\sum_{L=0}^{n} a_{k,L}x^L}{(1-x^2)^{n+2k}}\exp\left(-\frac{2}{(1-x^2)^2}\right)$$

where $a_{k,L}$ is arbitrary fixed coefficients (some may be zero).

We can see $f^{(n)}(x)$ is differentiable, i.e., $f^{(n+1)}(x)$ exists, and has the form of

$$\begin{aligned} f^{(n+1)}(x) &= \left[\sum_{k=1}^{n} -\frac{8x}{(1-x^2)^3}\cdot\frac{\sum_{L=0}^{n} a_{k,L}x^L}{(1-x^2)^{n+2k}} + \sum_{k=1}^{n}\frac{\sum_{L=0}^{n+1} b_{k,L}x^L}{(1-x^2)^{n+2k+1}}\right]\exp\left(-\frac{2}{(1-x^2)^2}\right) \\ &= \left[\sum_{k=1}^{n}\frac{\sum_{L=0}^{n} -8a_{k,L}x^{L+1}}{(1-x^2)^{n+2k+3}} + \sum_{k=1}^{n}\frac{\sum_{L=0}^{n+1} b_{k,L}x^L}{(1-x^2)^{n+2k+1}}\right]\exp\left(-\frac{2}{(1-x^2)^2}\right) \\ &= \sum_{k=1}^{n+1}\frac{\sum_{L=0}^{n+1} c_{k,L}x^L}{(1-x^2)^{(n+1)+2k}}\exp\left(-\frac{2}{(1-x^2)^2}\right) \end{aligned}$$

where $b_{k,L}$ and $c_{k,L}$ are coefficients that you don't need to compute. By induction, we conclude that our assumption is correct, that is, when $-1 < x < 1$, $f(x)$ is differentiable for arbitrary n.

When $|x| > 1$, $f(x)$ is constant function zero, so it is infinitely differentiable, and the derivative is zero.

Finally, we need to deal with two points $x = \pm 1$. Considet $x = 1$, by definition, using similar technique as part c), we can find that $f'(1) = 0$. Also $f'(-1) = 0$ for the same reason. Suppose $f^{(n)}(\pm 1)$ exists and also equal to zero, We want to prove $f^{(n+1)}(\pm 1) = 0$.

By definition,

$$\begin{aligned}
f_+^{(n+1)}(1) &= \lim_{h\to 0+} \frac{f^{(n)}(1+h) - f^{(n)}(1)}{h} \\
&= \lim_{h\to 0+} \frac{1}{h} \sum_{k=1}^{n} \frac{\sum_{L=0}^{n} a_{k,L}(1+h)^L}{(1-(1+h)^2)^{n+2k}} \exp\left(-\frac{2}{(1-(1+h)^2)^2}\right) \\
&= \lim_{h\to 0+} \sum_{k=1}^{n} \frac{P_n(h)}{P_{2n+4k+1}(h)} \exp\left(-\frac{2}{(1-(1+h)^2)^2}\right) \qquad \text{Take } t = 1/h \\
&= \lim_{t\to +\infty} \sum_{k=1}^{n} \frac{P_n(1/t)}{P_{2n+4k+1}(1/t)} \exp\left(-\frac{2t^4}{(2t+1)^2}\right) \\
&= 0
\end{aligned}$$

Similarly, $f_+^{(n+1)}(1) = 0$. Therefore, $f^{(n+1)}(1) = 0$. You can also show that $f^{(n+1)}(-1) = 0$. Thus, our assumption is correct, and we conclude that $f(x)$ is infinitely differentiable at $x \in \mathbb{R}$, with

$$f^{(n)}(x) = \begin{cases} \displaystyle\sum_{k=1}^{n} \frac{\sum_{L=0}^{n} a_{k,L} x^L}{(1-x^2)^{n+2k}} \exp\left(-\frac{2}{(1-x^2)^2}\right) & \text{for } -1 < x < 1 \\ 0 & \text{for } 1 \le |x| \end{cases}$$

The computation of derivative is very tedious, so you can try to avoid such mechanical calculations.

Question 5.2-3. Let $f \in \mathcal{C}^\infty(\mathbb{R})$. Show that for $x \neq 0$

$$\frac{1}{x^{n+1}} f^{(n)}\left(\frac{1}{x}\right) = (-1)^n \frac{d^n}{dx^n}\left(x^{n-1} f\left(\frac{1}{x}\right)\right)$$

When $n = 1$,

$$\frac{d}{dx} f\left(\frac{1}{x}\right) = -\frac{1}{x^2} f'\left(\frac{1}{x}\right)$$

When $n = 2$,

$$\begin{aligned}
\frac{d^2}{dx^2}\left[xf\left(\frac{1}{x}\right)\right] &= \frac{d}{dx}\left[f\left(\frac{1}{x}\right) - \frac{1}{x}f'\left(\frac{1}{x}\right)\right] \\
&= \frac{d}{dx}f\left(\frac{1}{x}\right) + \frac{d}{dx}\left[-\frac{1}{x}f'\left(\frac{1}{x}\right)\right] \\
&= -\frac{1}{x^2}f'\left(\frac{1}{x}\right) + \frac{1}{x^2}f'\left(\frac{1}{x}\right) + \frac{1}{x^3}f''\left(\frac{1}{x}\right) \\
&= (-1)^2\frac{1}{x^3}f''\left(\frac{1}{x}\right)
\end{aligned}$$

By strong induction, we suppose for all $n = 1, 2, \ldots, k$, we have

$$\frac{d^n}{dx^n}\left(x^{n-1}f\left(\frac{1}{x}\right)\right) = (-1)^n\frac{1}{x^{n+1}}f^{(n)}\left(\frac{1}{x}\right)$$

Consider when $n = k + 1$,

$$\begin{aligned}
\frac{d^{k+1}}{dx^{k+1}}\left(x^k f\left(\frac{1}{x}\right)\right) &= \frac{d^k}{dx^k}\left[\frac{d}{dx}\left(x^k f\left(\frac{1}{x}\right)\right)\right] \\
&= \frac{d^k}{dx^k}\left[kx^{k-1}f\left(\frac{1}{x}\right) - x^{k-2}f'\left(\frac{1}{x}\right)\right] \\
&= k\frac{d^k}{dx^k}\left[x^{k-1}f\left(\frac{1}{x}\right)\right] - \frac{d}{dx}\left\{\frac{d^{k-1}}{dx^{k-1}}\left[x^{k-2}f'\left(\frac{1}{x}\right)\right]\right\} \\
&= k(-1)^k\frac{1}{x^{k+1}}f^{(k)}\left(\frac{1}{k}\right) + \frac{d}{dx}\left[(-1)^k\frac{1}{x^k}f^{(k)}\left(\frac{1}{x}\right)\right] \qquad (1) \\
&= k(-1)^k\frac{1}{x^{k+1}}f^{(k)}\left(\frac{1}{k}\right) - k(-1)^k\frac{1}{x^{k+1}}f^{(k)}\left(\frac{1}{x}\right) - (-1)^k\frac{1}{x^{k+2}}f^{(k+1)}\left(\frac{1}{x}\right) \\
&= (-1)^{k+1}\frac{1}{x^{k+2}}f^{(k+1)}\left(\frac{1}{x}\right)
\end{aligned}$$

Notice that in step (1), we substitute the original terms for another two terms in our assumption. One can easily see that when $n = k + 1$, the formula still holds. Therefore, we finish the proof.

Question 5.2-4. Let f be a differentiable function on $\mathbb{R}$. Show that

a) if f is an even function, then f' is an odd function;

If f is an even function, then $f(x) = f(-x)$ for $x \in \mathbb{R}$. Consider

$$\begin{aligned}
f'(-x) &= \lim_{h\to 0}\frac{f(-x+h) - f(-x)}{h} \\
&= \lim_{h\to 0}\frac{f(x-h) - f(x)}{h} \qquad \text{Take } t = -h \\
&= \lim_{t\to 0}\frac{f(x+t) - f(x)}{-t} \\
&= \lim_{t\to 0} -\frac{f(x+t) - f(x)}{t} = -f'(x)
\end{aligned}$$

Therefore, $f'(x)$ is an odd function.

b) if f is an odd function, then f' is an even function;

If f is an odd function, then $f(-x) = -f(x)$ for $x \in \mathbb{R}$. Consider

$$\begin{aligned} f'(-x) &= \lim_{h\to 0} \frac{f(-x+h) - f(-x)}{h} \\ &= \lim_{h\to 0} \frac{-f(x-h) + f(x)}{h} \quad \text{Take } t = -h \\ &= \lim_{t\to 0} \frac{-f(x+t) + f(x)}{-t} \\ &= \lim_{t\to 0} \frac{f(x+t) - f(x)}{t} = f'(x) \end{aligned}$$

Therefore, $f'(x)$ is an even function.

c) (f' is odd) $\Longleftrightarrow$ (f is even).

You really need to be careful that for this question, you cannot exchange the "odd" with "even".

Since f' is odd, we have $f'(-x) = -f'(x)$. We consider

$$\begin{aligned} [f(x) - f(-x)]' &= f'(x) - f'(-x)(-1) \\ &= f'(x) + f'(-x) = 0 \end{aligned}$$

Therefore, we can see $f(x) - f(-x) = c$ for all $x \in \mathbb{R}$, where c is constant.

Since $f(x)$ is differentiable, $f(x) - f(-x)$ must be continuous at $x = 0$, i.e.,

$$f(0) - f(-0) = 0 = c$$

Thus, $c = 0$, and we have $f(x) = f(-x)$ for all $x \in \mathbb{R}$, meaning that $f(x)$ is even function.

The other direction we have proved it in part a), so we finally have (f' is odd) $\Longleftrightarrow$ (f is even).

Question 5.2-5. Show that

a) the function $f(x)$ is differentiable at the point x_0 if and only if $f(x) - f(x_0) = \varphi(x)(x - x_0)$, where $\varphi(x)$ is a function that is continuous at x_0 (and in that case $\varphi(x_0) = f'(x_0)$);

We first prove the "if" part. When $x \to x_0$, $x \neq x_0$, we have

$$\lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x\to x_0} \varphi(x)$$

Since $\varphi(x)$ is continuous at x_0, so the limit exist, and

$$\lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x\to x_0} \varphi(x) = \varphi(x_0)$$

Then we prove the "only if" part. Since $f(x)$ is differentiable at x_0, we can define a function $\varphi(x)$ as

$$\varphi(x) = \begin{cases} \dfrac{f(x) - f(x_0)}{x - x_0} & \text{if } x \neq x_0 \\ f'(x_0) & \text{if } x = x_0 \end{cases}$$

We only need to prove such $\varphi(x)$ is continuous at x_0, which is equivalent to show

$$\lim_{x\to x_0} \varphi(x) = \varphi(x_0)$$

Notice that the L.H.S. is

$$\lim_{x\to x_0} \varphi(x) = \lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0)$$

which is by definition equal to the R.H.S., therefore, we finish the proof.

b) if $f(x) - f(x_0) = \varphi(x)(x - x_0)$ and $\varphi \in \mathcal{C}^{(n-1)}(U(x_0))$, where $U(x_0)$ is a neighborhood of x_0, then $f(x)$ has a derivative $(f^{(n)}(x_0))$ of order n at x_0.

Since $\varphi \in \mathcal{C}^{(n-1)}(U(x_0))$, $\varphi^{(n-1)}(x)$ exists and is continuous. We know that $(x - x_0)$ is infinitely differentiable, thus $f(x) = \varphi(x)(x - x_0) + f(x_0)$ is at least $(n-1)$-th differentiable in the neighborhood $U(x_0)$, and by chain rule, we can see

$$\begin{aligned}
f'(x) &= \varphi'(x)(x - x_0) + \varphi(x) \\
f'(x) &= \varphi''(x)(x - x_0) + 2\varphi'(x) \\
&\vdots \\
f^{(n-1)}(x) &= \varphi^{(n-1)}(x)(x - x_0) + (n-1)\varphi^{(n-2)}(x)
\end{aligned}$$

We consider $f^{(n)}(x_0)$ by definition,

$$\begin{aligned}
f^{(n)}(x_0) &= \lim_{x\to x_0} \frac{f^{(n-1)}(x) - f^{(n-1)}(x_0)}{x - x_0} \\
&= \lim_{x\to x_0} \frac{\varphi^{(n-1)}(x)(x - x_0) + (n-1)\varphi^{(n-2)}(x) - (n-1)\varphi^{(n-2)}(x_0)}{x - x_0} \\
&= \varphi^{(n-1)}(x_0) + \lim_{x\to x_0} \frac{(n-1)\varphi^{(n-2)}(x) - (n-1)\varphi^{(n-2)}(x_0)}{x - x_0} \\
&= \varphi^{(n-1)}(x_0) + (n-1) \lim_{x\to x_0} \frac{\varphi^{(n-2)}(x) - \varphi^{(n-2)}(x_0)}{x - x_0} \\
&= \varphi^{(n-1)}(x_0) + (n-1)\varphi^{(n-1)}(x_0) \\
&= n\varphi^{(n-1)}(x_0)
\end{aligned}$$

Therefore, $f(x)$ has a derivative $f^{(n)}(x_0)$ of order n at x_0, which equal to $n\varphi^{(n-1)}(x_0)$.

Question 5.2-6. Give an example showing that the assumption that f^{-1} be continuous at the point y_0 cannot be omitted from Theorem 3 (The derivative of an inverse function).

We can take $f : [0,1) \cup [2,3] \mapsto [0,2]$ which is defined as

$$f(x) = \begin{cases} x & x \in [0,1) \\ x - 1 & x \in [2,3] \end{cases}$$

Notice that $f(x)$ is continuous on $[0,1) \cup [2,3]$, and since we define the derivative at endpoint $x = 0, 2, 3$ as the left or right hand side derivative, $f(x)$ is differentiable at any point in its domain.

We can easily see that $f^{-1}: [0,2] \mapsto [0,1) \cup [2,3]$ is defined as

$$f^{-1}(x) = \begin{cases} x & x \in [0,1) \\ x+1 & x \in [1,2] \end{cases}$$

It's obvious that $f^{-1}(x)$ is discontinuous at $x = 1$, because $f(1-) = 1$ and $f(1+) = 2$. Since $f^{-1}(x)$ is not continuous, its derivative does not exist at $x = 1$, let alone $(f^{-1})'(1)$ will not coincide with $(f'(2))^{-1}$.

Question 5.3-1. Choose numbers a and b so that the function $f(x) = \cos x - \dfrac{1+ax^2}{1+bx^2}$ is an infinitesimal of highest possible order as $x \to 0$.

Using Taylor series, we have

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + o(x^4)$$

Function $f(x)$ can be rewritten as

$$f(x) = \frac{-1 - ax^2 + (1+bx^2)[1 - \frac{x^2}{2!} + \frac{x^4}{4!} + o(x^4)]}{1+bx^2} = \frac{(b-a-1/2)x^2 + (1/24 - b/2)x^4 + o(x^4)}{1+bx^2}$$

Since we need to have

$$\lim_{x\to 0} \frac{f(x)}{x^n} = 0$$

To make n larger, we need to make the lowest order term in numerator as high as possible. Since we only have two degree of freedom, we can only make $(b-a-1/2) = 0$ and $(1/24 - b/2) = 0$. This yields $b = 1/12$ and $a = -5/12$. Although the exact highest order of infinitesimal is not easy to see, such a, b indeed ensure you the highest order.

Question 5.3-2. Find $\lim_{x\to\infty} x\left[\frac{1}{e} - \left(\frac{x}{x+1}\right)^x\right]$.

Take $x = \frac{1}{t}$, we conclude that

$$\begin{aligned}
\lim_{x\to\infty} x\left[\frac{1}{e} - \left(\frac{x}{x+1}\right)^x\right] &= \lim_{t\to 0} \frac{1}{t}\left[\frac{1}{e} - \left(\frac{1}{1+t}\right)^{1/t}\right] \\
&= \lim_{t\to 0} \frac{(1+t)^{1/t} - e}{e(1+t)^{1/t}t} \\
&= \lim_{t\to 0} \frac{e^{t^{-1}\ln(1+t)} - e}{e(1+t)^{1/t}t} \\
&= \lim_{t\to 0} \frac{e^{t^{-1}\ln(1+t)-1} - 1}{(1+t)^{1/t}t} \\
&= \lim_{t\to 0} \frac{t^{-1}\ln(1+t) - 1}{(1+t)^{1/t}t} \\
&= \lim_{t\to 0} \frac{\ln(1+t) - t}{(1+t)^{1/t}t^2} \\
&= \lim_{t\to 0} \frac{1}{(1+t)^{1/t}} \frac{\ln(1+t) - t}{t^2}
\end{aligned}$$

Now we consider the following limit

$$\lim_{t\to 0}\frac{\ln(1+t)-t}{t^2}=\lim_{t\to 0}\frac{(1+t)^{-1}-1}{2t}$$
$$=\lim_{t\to 0}\frac{-1}{2(1+t)}=-\frac{1}{2}$$

Consider what we have already known

$$\lim_{t\to 0}(1+t)^{1/t}=e$$

Therefore, we conclude that

$$\lim_{x\to\infty}x\left[\frac{1}{e}-\left(\frac{x}{x+1}\right)^x\right]=\lim_{t\to 0}\frac{1}{(1+t)^{1/t}}\frac{\ln(1+t)-t}{t^2}$$
$$=\lim_{t\to 0}\frac{1}{(1+t)^{1/t}}\lim_{t\to 0}\frac{\ln(1+t)-t}{t^2}$$
$$=-\frac{1}{2e}$$

Question 5.3-3. Write a Taylor polynomial of e^x at zero that makes it possible to compute the values of e^x on the closed interval $-1\le x\le 2$ within 10^{-3}.

We expand e^x at $x=0$, and we can first figure out that $e<3$ by using the definition of e. Hence,

$$e^x=1+\frac{1}{1!}x+\frac{1}{2!}x^2+\cdots+\frac{1}{n!}x^n+\frac{e^\xi}{(n+1)!}x^{n+1}$$

For $x,\xi\in[-1,2]$, we have

$$\left|\frac{e^\xi}{(n+1)!}x^{n+1}\right|\le\frac{e^2}{(n+1)!}|x|^{n+1}\le\frac{9\times 2^{n+1}}{(n+1)!}$$

Solve the inequality

$$\frac{9\times 2^{n+1}}{(n+1)!}\le 10^{-3}$$

We get $n\ge 10$, hence the required polynomial is

$$e^x\approx 1+\frac{1}{1!}x+\frac{1}{2!}x^2+\cdots+\frac{1}{10!}x^{10}$$

Question 5.3-4. Let f be a function that is infinitely differentiable at 0. Show that

a) if f is even, then its Taylor series at 0 contains only even powers of x;

We have proved in **Question 5.2-4** that if f is even, then f' will be odd, and if f' is odd, $f''=(f')'$ will be even. Thus, by induction we can see that f^n will be odd if n is odd, f^n will be even if n is even. The Taylor series of f at 0 is

$$f(x)=f(0)+\frac{f'(0)}{1!}x+\frac{f''(0)}{2!}x^2+\cdots+\frac{f^n(0)}{n!}x^n+\cdots$$

Since f^n will be odd if n is odd, so for all odd n, $f^n(0) = 0$. Hence, for $k \in \mathbb{N}$,

$$f(x) = f(0) + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^{2k}(0)}{(2k)!}x^{2k} + \cdots$$

Therefore, the Taylor series of f at 0 contains only even power (if $f \equiv c$, where c is constant, then it contains 0 power of x, which is also even.)

b) if f is odd, then its Taylor series at 0 contains only odd powers of x.

Similarly, if f is odd, then $f(0) = 0$, f' is even, and $f'' = (f')'$ is odd. Thus, by induction, we can see that f^n will be even if n is odd, and odd if n is even. The Taylor series of f at 0 is

$$f(x) = 0 + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^n(0)}{n!}x^n + \cdots$$

Since f^n will be odd if n is even, so for all even n, $f^n(0) = 0$. Hence, for $k \in \mathbb{N}$,

$$f(x) = \frac{f'(0)}{1!}x + \cdots + \frac{f^{2k+1}(0)}{(2k+1)!}x^{2k+1} + \cdots$$

Therefore, the Taylor series of f at 0 contains only odd power (if $f \equiv c = 0$, then it contains 0 power of x, which is even, but such case should not be included.)

Question 5.3-5. Show that if $f \in \mathcal{C}^{(\infty)}[-1,1]$ and $f^{(n)}(0) = 0$ for $n = 0, 1, 2, \ldots$, and there exists a number C such that $\sup\limits_{-1\le x\le 1} |f^{(n)}(x)| \le n!C$ for $n \in \mathbb{N}$, then $f \equiv 0$ on $[-1,1]$.

Expand $f(x)$ at x_0 using Taylor's expansion, we have

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}$$

Take $x_0 = 0$, for some $\xi \in [0, 1]$ we have

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \cdots + \frac{f^{(n)}(0)}{n!}x^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}x^{n+1}$$

Since $f^{(n)}(0) = 0$ for all n, we have

$$|f(x)| = \left| \frac{f^{(n+1)}(\xi)}{(n+1)!}x^{n+1} \right| \le |c||x^{n+1}|$$

If $-1 \le x < 1$, for each fixed x, take limit $(n \to \infty)$ on both sides, we have $|f(x)| \le 0$, hence $f \equiv 0$. Since $f(x)$ must be continuous on $[-1, 1]$, thus $f(1) = f(1-) = 0$ and $f(-1) = f((-1)+) = 0$. Therefore $f \equiv 0$ on $[-1, 1]$.

Question 5.3-6. Let $f \in \mathcal{C}^{(n)}(-1,1)$ and $\sup\limits_{-1\le x\le 1} |f(x)| \le 1$. Let $m_k(I) = \inf\limits_{x\in I} |f^{(k)}(x)|$, where I is an interval contained in $(-1, 1)$. Show that

a) if I is partitioned into three successive intervals I_1, I_2, and I_3 and μ is the length of I_2, then

$$m_k(I) \le \frac{1}{\mu}\Big(m_{k-1}(I_1) + m_{k-1}(I_3)\Big)$$

Apply MVT on interval I_2, there exists $\xi \in I_2$, such that

$$f^{(k)}(\xi) = \frac{f^{(k-1)}(x_3) - f^{(k-1)}(x_1)}{x_3 - x_1}$$

Since $f^{(k)}(\xi) \geq \inf_{x \in I} |f^{(k)}(x)| = m_k(I)$, we have

$$\begin{aligned} m_k(I) &\leq \frac{f^{(k-1)}(x_3) - f^{(k-1)}(x_1)}{x_3 - x_1} \\ &\leq \frac{1}{\mu}\Big(f^{(k-1)}(x_3) - f^{(k-1)}(x_1) \Big) \\ &\leq \frac{1}{\mu}\Big(|f^{(k-1)}(x_3)| + |f^{(k-1)}(x_1)| \Big) \end{aligned}$$

Since for all x_1, x_3, we have the above relation, i.e., the right hand side is an upper bound of $m_k(I)$, the least upper bound also satisfies the above relation. Thus, we have

$$m_k(I) \leq \frac{1}{\mu}\Big(m_{k-1}(I_3) + m_{k-1}(I_1) \Big)$$

b) if I has length λ, then

$$m_k(I) \leq \frac{2^{k(k+1)/2} k^k}{\lambda^k}$$

We prove it by induction. When $k = 1$, this is obviously true, because $\sup_{-1 \leq x \leq 1} |f(x)| \leq 1$. Suppose it is true for $k = n$, then for $k = n+1$, we have (Denote the length of I_i as $|I_i|$)

$$\begin{aligned} m_{n+1}(I) &\leq \frac{1}{|I_2|}[m_n(I_3) + m_n(I_1)] \\ &\leq \frac{1}{|I_2|}\left[\frac{2^{n(n+1)/2} n^n}{|I_1|^n} + \frac{2^{n(n+1)/2} n^n}{|I_3|^n} \right] \\ &= 2^{n(n+1)/2} n^n \frac{1}{|I_2|}\left[\frac{1}{|I_1|^n} + \frac{1}{|I_3|^n} \right] \end{aligned}$$

Notice that the above inequality holds for any partition of I into I_1, I_2, I_3, so we can take $|I_2| = \frac{\lambda}{n+1}$ and $|I_1| = |I_3| = \frac{n\lambda}{2(n+1)}$. Therefore, we have

$$\begin{aligned} m_{n+1}(I) &\leq 2^{n(n+1)/2} n^n \frac{1}{|I_2|}\left[\frac{1}{|I_1|^n} + \frac{1}{|I_3|^n} \right] \\ &= 2^{n(n+1)/2} n^n \frac{n+1}{\lambda} \frac{2 \cdot 2^n (n+1)^n}{n^n \lambda^n} \\ &= \frac{2^{(n+1)(n+2)/2} (n+1)^{(n+1)}}{\lambda^{n+1}} \end{aligned}$$

Hence, we verify that for $k = n+1$, our assumption still holds, meaning that our assumption is correct. Hence the proof is finished.

c) there exists a number α_n depending only on n such that if $|f'(0)| \geq \alpha_n$, then the equation $f^{(n)}(x) = 0$ has at least $n-1$ distinct roots in $(-1, 1)$.

Question 5.3-7. Show that if a function f is defined and differentiable on an open interval I and $[a, b] \subset I$, then

a) the function $f'(x)$ (even if it is not continuous!) assumes on $[a, b]$ all the values between $f'(a)$ and $f'(b)$;

We only consider the case that $f'(a) < \lambda < f'(b)$. Let $g(x) = f(x) - \lambda x$ on $[a, b]$, we have $g'(x) = f'(x) - \lambda$. It's easy to see that $g'(a) = f'(a) - \lambda < 0$, which means $g(a)$ is not the maximum value of $g(x)$ in $[a, b]$. Similarly, $g'(b) > 0$ means that $g(b)$ is not the maximum value of $g(x)$ in $[a, b]$. However, since $g(x)$ is continuous function on closed interval, so it must assume its maximum value in $[a, b]$. Hence, there exist $\xi \in (a, b)$, such that $g(\xi)$ assume the maximum value of g. Since $g(x)$ is differentiable at ξ, $g'(\xi) = 0$. Thus, $f'(\xi) = \lambda$. We choose λ arbitrarily, so $f'(x)$ assumes all value between $f'(a)$ and $f'(b)$ on $[a, b]$.

The case that $f'(a) > \lambda > f'(b)$ is left as exercise so that you can check whether you really understand such proof. In conclu

b) if $f''(x)$ also exists in (a, b), then there is a point $\xi \in (a, b)$ such that $f'(b) - f'(a) = f''(\xi)(b - a)$.

Since $f'(x)$ may not be continuous on $[a, b]$, we cannot apply MVT directly. Instead, suppose such ξ doesn't exist. Then let $m = \frac{f'(b)-f'(a)}{b-a}$, and $f''(x) \neq m$ for all $x \in (a, b)$. Although $f'(x)$ may be discontinuous at a or b, $f''(x)$ satisfies intermediate value property on any closed interval contained in $[a, b]$. This implies that either $f''(x) > m$ for all $x \in (a, b)$ or $f''(x) < m$ for all $x \in (a, b)$, because if $f''(x_1) > m$ and $f''(x_2) < m$, then there must exist $p \in [x_1, x_2]$ such that $f''(p) = m$, which is a contradiction.

If $f''(x) > m$ for all $x \in (a, b)$, then $g(x) = f'(x) - mx$ is strictly increasing and differentiable on (a, b) and defined on $[a, b]$. This implies that $\lim_{x \to a+} g(x)$ exists and can be a finite number or negative infinity. If it is negative infinity, then there exists a right half-neighborhood $N_\delta^+(a)$ of a such that for all $x \in N_\delta^+(a)$, $g(x) < g(a) - 1$. Then on closed interval $[a, \delta/2]$, $g(x)$ cannot attain all values between $g(a)$ and $g(\delta/2)$, therefore the intermediate value property fails, which is a contradiction to the conclusion of part (a), thus $\lim_{x \to a+} g(x)$ must be a finite number. Then $f'(x)$ must have right-hand-side limit at a. Similarly, $f'(x)$ must also have left-hand-side limit at b. Therefore $f'(x)$ is continuous on $[a, b]$, and we can apply MVT to obtain ξ such that $f''(\xi) = m$, which contradicts to our assumption that such ξ doesn't exist.

The only possibility now is that $f''(x) < m$. However, using similar argument we can show that in this case $f'(x)$ is still continuous on $[a, b]$, so MVT implies the existence of such ξ, which means our assumption is wrong, i.e., there exists ξ such that $f''(\xi) = m$, and the proof is finished.

Question 5.3-8. A function $f(x)$ may be differentiable on the entire real line, without having a continuous derivative $f'(x)$.

a) Show that $f'(x)$ can have only discontinuities of second kind.

Suppose $f'(x)$ has discontinuity of first kind at a, i.e., both one-side limits exist, but at least one of them are not equal to the function value at that point. W.O.L.G., we can assume this one-side limit is $f'(a+)$, and $f'(a+) \neq f'(a)$. Here we only consider the case that $f'(a+) < f'(a)$.

Since $f'(a+) < f'(a)$, there exists ξ, such that $f'(a+) < \xi < f'(a)$. Since $f'(a+)$ is the limit of $f'(x)$ as $x \to a+$, by definition, for all $\varepsilon > 0$, there exists $\delta > 0$, such that $|f'(x) - f'(a+)| < \varepsilon$. for all $x \in (a, a+\delta)$. Take $\varepsilon = (\xi - f'(a+))/2$, so we have a δ satisfies the above relation, and

$$f'(x) < f'(a+) + \varepsilon = \frac{\xi + f'(a+)}{2} < \xi \quad \text{for all } x \in (a, a+\delta)$$

However, since $f'(a+\delta/2) < \xi < f'(a)$, by what we proved in **Question 5.3-7(a)**, there exists $\lambda \in (a, a+\delta/2)$, such that $f'(\lambda) = \xi$. This is a contradiction, because we have just proved that for any $x \in (a, a+\delta)$, $f'(x) < \xi$. Hence $f'(a+) \geq f'(a)$.

The contradiction of $f'(a+) > f'(a)$ is similar, and you can prove it to check whether you understand this. After that, we yield that $f'(a+) = f'(a)$, which contradicts our assumption that $f'(a+) \neq f'(a)$. Hence, no discontinuity of first kind can appear in the derivative of some differentiable function on the entire real line.

b) Find the flaw in the following "proof" that $f'(x)$ is continuous.

> *Proof.* Let x_0 be an arbitrary point on $\mathbb{R}$ and $f'(x_0)$ the derivative of f at the point x_0. By definition of the derivative and Lagrange's theorem
>
> $$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to x_0} f'(\xi) = \lim_{\xi \to x_0} f'(\xi)$$
>
> where ξ is a point between x_0 and x and therefore tends to x_0 as $x \to x_0$. □

The flaw lies in the **third** equality above. Actually the third equality fails if the function is not continuously differentiable, i.e., $f'(x)$ is not continuous. This is mainly because $f'(\xi(x))$ is continuous with respect to x, but not continuous with respect to ξ. Combine with what we proved in part a), $f'(\xi)$ can only have second kind of discontinuity with respect to ξ, so $\lim_{\xi \to x_0} f'(\xi)$ may not exist. Hence, even if as $x \to x_0$, ξ also tends to x_0. we can only take the limit as $x \to x_0$ but not $\xi \to x_0$, and

$$\lim_{x \to x_0} f'(\xi(x)) \neq \lim_{\xi \to x_0} f'(\xi(x))$$

Question 5.3-9. Let f be twice differentiable on an interval I. Let $M_0 = \sup_{x \in I} |f(x)|$, $M_1 = \sup_{x \in I} |f'(x)|$ and $M_2 = \sup_{x \in I} |f''(x)|$. Show that

a) if $I = [-a, a]$, then

$$|f'(x)| \leq \frac{M_0}{a} + \frac{x^2 + a^2}{2a} M_2$$

We consider Taylor's expansion, when $h_1 \neq h_2$ and $x + h_1, x + h_2 \in [-a, a]$,

$$f(x + h_1) = f(x) + f'(x) h_1 + \frac{f''(x + \theta_1 h_1)}{2!} h_1^2 \qquad 0 < \theta_1 < 1 \tag{1}$$

$$f(x+h_2) = f(x) + f'(x)h_2 + \frac{f''(x+\theta_2 h_2)}{2!}h_2^2 \qquad 0 < \theta_2 < 1 \tag{2}$$

Consider $(1) - (2)$, we have

$$f'(x) = \frac{f(x+h_1) - f(x+h_2)}{h_1 - h_2} + \frac{f''(x+\theta_1 h_1)}{2(h_1-h_2)}h_1^2 - \frac{f''(x+\theta_2 h_2)}{2(h_1-h_2)}h_2^2$$

Thus, by triangular inequality,

$$\begin{aligned}
|f'(x)| &\le \frac{|f(x+h_1) - f(x+h_2)|}{|h_1 - h_2|} + \frac{|f''(x+\theta_1 h_1)|}{2|h_1-h_2|}h_1^2 + \frac{|f''(x+\theta_2 h_2)|}{2|h_1-h_2|}h_2^2 \\
&\le \frac{2}{|h_1-h_2|}M_0 + \frac{(h_1^2+h_2^2)}{2|h_1-h_2|}M_2
\end{aligned}$$

Let

$$\frac{2}{|h_1-h_2|} = \frac{1}{a} \qquad \text{and} \qquad \frac{(h_1^2+h_2^2)}{2|h_1-h_2|} = \frac{x^2+a^2}{2a}$$

Then we have

$$\begin{cases} |h_1 - h_2| = 2a \\ h_1^2 + h_2^2 = 2a^2 + 2x^2 \end{cases}$$

W.L.O.G., we can assume $h_1 > h_2$, because if not, we can regard h_1 as h_2 and h_2 as h_1. Thus, we can solve

$$\begin{cases} h_1 = a + |x| \\ h_2 = -a + |x| \end{cases} \quad \text{if } x \in [-a, 0], \qquad \begin{cases} h_1 = a - |x| \\ h_2 = -a - |x| \end{cases} \quad \text{if } x \in [0, a]$$

One can check that such solutions truly satisfy $x + h_1, x + h_2 \in [-a, a]$. Therefore, we have finished our proof.

b) $\begin{cases} M_1 \le 2\sqrt{M_0 M_2}, \text{ if the length of } I \text{ is not less than } 2\sqrt{M_0/M_2} \\ M_1 \le \sqrt{2M_0M_2}, \text{ if } I = \mathbb{R} \end{cases}$

If the length of I is not less than $2\sqrt{M_0/M_2}$, let $I = [a, a+l]$, then we take $h_1 = a + l - x$, $h_2 = a - x$, where $x \in [a, a+l]$. Then from part a), we have

$$|f'(x)| \le \frac{2}{|h_1-h_2|}M_0 + \frac{(h_1^2+h_2^2)}{2|h_1-h_2|}M_2$$

Thus, the above inequality means for all l, we have

$$|f'(x)| \le \frac{2}{l}M_0 + \frac{l^2 + 2h_1h_2}{2l}M_2 \le \frac{2}{l}M_0 + \frac{l^2}{2l}M_2 = \frac{2}{l}M_0 + \frac{l}{2}M_2$$

since $h_1 h_2 < 0$. Take $l = 2\sqrt{M_0/M_2}$, we have

$$M_1 = \sup_{x \in I} |f'(x)| \le 2\sqrt{M_0 M_2}$$

Note that if I is not closed, then we just take $[a+\epsilon, a+l-\epsilon]$, and then take limit as $\epsilon \to 0$, we will obtain exactly the same answer.

If $I = \mathbb{R}$, for any h_1, take $h_2 = -h_1$, then we have

$$|f'(x)| \le \frac{1}{h_1}M_0 + \frac{h_1}{2}M_2$$

Let $h_1 = \sqrt{2M_0/M_2}$, we have

$$M_1 = \sup_{x\in I} |f'(x)| \le \sqrt{2M_0M_2}$$

c) the numbers 2 and $\sqrt{2}$ in part b) cannot be replaced by smaller numbers;

First, consider $f(x) = 2x^2 - 1$ on $I = [-1,1]$. We can check that $M_0 = 1$, $M_1 = M_2 = 4$ on $[-1,1]$, and the length of I is 2, which is no less than $2\sqrt{1/4} = 1$. And $4 = M_1 = 2\sqrt{M_0M_2} = 2\sqrt{1\cdot 4}$. Since the equality can be obtained, this 2 cannot be replaced by smaller number.

Second, consider the following function,

$$f(x) = \begin{cases} -\dfrac{1-(x+1/2)^2}{1+(x+1/2)^2} & \text{if } x \le -\dfrac{1}{2} \\ 4(x+1/2)^2 - 1 & \text{if } -\dfrac{1}{2} \le x \le 0 \\ -4(x-1/2)^2 + 1 & \text{if } 0 \le x \le \dfrac{1}{2} \\ \dfrac{1-(x-1/2)^2}{1+(x-1/2)^2} & \text{if } x \ge \dfrac{1}{2} \end{cases}$$

For this function, it's easy to see that its range is $[-1,1]$ on $\mathbb{R}$, so $M_0 = 1$. Take the first order derivative, one can verify $M_1 = 4$ and also $M_2 = 8$ by taking the second derivative (A little bit tedious for rigorous proof, but quite intuitive if you draw the graph of the function). This shows that

$$4 = M_1 = \sqrt{2M_0M_2} = \sqrt{2\cdot 1\cdot 8} = 4$$

Since the equality can be obtained, this $\sqrt{2}$ cannot be replaced by smaller number.

d) if f is differentiable p times on $\mathbb{R}$ and the quantities M_0 and $M_p = \sup_{x\in\mathbb{R}} |f^{(p)}(x)|$ are finite, then the quantities $M_k = \sup_{x\in\mathbb{R}} |f^{(k)}(x)|$, $1 \le k < p$, are also finite and

$$M_k \le 2^{k(p-k)/2} M_0^{1-k/p} M_p^{k/p}$$

The case when $p = 2$ is proved in part b), now we assume what we need to prove is true when $p = m$, i.e., for $1 \le k \le m-1$,

$$M_k \le 2^{k(m-k)/2} M_0^{1-k/m} M_n^{k/m}$$

Also, you should know how to prove $M_m \le \sqrt{2M_{m-1}M_{m+1}}$ $(m \ge 1)$ in general, by using similar method in part a), and setting the same value to h_1, h_2 as part b) (This is a good exercise for you to check whether you really understand such method.) Suppose you have proved it, then we have

$$M_m \le \sqrt{2M_{m-1}M_{m+1}} \le \sqrt{2\cdot 2^{(m-1)/2} M_0^{1/m} M_m^{1-1/m} M_{m+1}}$$

Solve M_m (Notice that there is a $M_m^{1-1/m}$ term on right hand side!), we have

$$M_m \le 2^{m/2} M_0^{1/(m+1)} M_{m+1}^{m/(m+1)}$$

Substitute it into our assumption, we have

$$\begin{aligned}M_k &\le 2^{k(m-k)/2} M_0^{1-k/m} M_n^{k/m} \\ &\le 2^{k(m-k)/2} M_0^{1-k/m} 2^{k/2} M_0^{k/[m(m+1)]} M_{m+1}^{k/(m+1)} \\ &\le 2^{k(m+1-k)/2} M_0^{1-k/(m+1)} M_{m+1}^{k/(m+1)}\end{aligned}$$

which shows for $p = m+1$, our assumption still holds. By induction, the proof is finished.

Question 5.3-10. Show that if a function f has derivatives up to order $n+1$ inclusive at a point x_0 and $f^{(n+1)}(x_0) \neq 0$, then in the Lagrange form of the remainder in Taylor's formula

$$r_n(x_0; x) = \frac{1}{n!} f^{(n)}\left(x_0 + \theta(x - x_0)\right)(x - x_0)^n$$

where $0 < \theta < 1$ and the quantity $\theta = \theta(x)$ tends to $\frac{1}{n+1}$ as $x \to x_0$.

Apply Taylor's expansion to order $n-1$ and n at $x = x_0$, we have

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \cdots + \frac{f^{(n-1)}(x_0)}{(n-1)!}(x - x_0)^{n-1} + \frac{f^{(n)}(x_0 + \theta(x - x_0))}{n!}(x - x_0)^n$$

and

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \cdots + \frac{f^{(n+1)}(x_0)}{(n+1)!}(x - x_0)^{n+1} + o\big((x - x_0)^{n+1}\big)$$

Use the second one to subtract the first one, we have

$$f^{(n)}(x_0 + \theta(x - x_0)) - f^{(n)}(x_0) = \frac{f^{(n+1)}(x_0)}{n+1}(x - x_0) + o\big((x - x_0)\big)$$

Divide both sides by $\theta(x - x_0)$, and rearrange the equation as follows

$$\theta = \frac{\dfrac{f^{(n+1)}(x_0)}{n+1} + o(1)}{\dfrac{f^{(n)}(x_0 + \theta(x - x_0)) - f^{(n)}(x_0)}{\theta(x - x_0)}}$$

Take $x \to x_0$, $\theta(x - x_0) \to 0$, we have

$$\theta \longrightarrow \frac{\dfrac{f^{(n+1)}(x_0)}{n+1}}{f^{(n+1)}(x_0)} = \frac{1}{n+1}$$

Question 5.3-11. Let f be a function that is differentiable n times on an interval I. Prove the following statements.

a) If f vanishes at $(n+1)$ points of I, there exists a point $\xi \in I$ such that $f^{(n)}(\xi) = 0$.

Suppose f vanishes at $x_1 < x_2 < \cdots < x_n < x_{n+1}$, and all $x_i \in I$. Apply Rolle's theorem to each interval (x_i, x_{i+1}), we can find $z_1, \ldots, z_n$ such that $x_1 < z_1 < x_2 \cdots < z_n < x_{n+1}$, and $f'(z_i) = 0$. In this case, all z_i will lie in I. Now apply Rolle's theorem to $f'(x)$ on each interval (z_i, z_{i+1}), we can find $w_1, \ldots, w_{n-1}$ such that $z_1 < w_1 < z_2 \cdots < w_{n-1} < z_n$, and $f''(w_i) = 0$. By induction, we can continue this process until we find $\xi_1 < \xi_2$, $\xi_i \in I$ such that $f^{(n-1)}(\xi_i) = 0$. Again, using Rolle's theorem, we can obtain $\xi \in (\xi_1, \xi_2)$, such that $f^{(n)}(\xi) = 0$. Thus, we proved the Generalized Rolle's Theorem.

b) If $x_1, x_2, \ldots, x_n$ are points of the interval I, there exists a unique polynomial $L(x)$ (the Lagrange interpolation polynomial) of degree at most $(n-1)$ such that $f(x_i) = L(x_i)$, $i = 1, \ldots, n$. In addition, for $x \in I$ there exists a point $\xi \in I$ such that

$$f(x) - L(x) = \frac{(x - x_1)\cdots(x - x_n)}{n!} f^{(n)}(\xi)$$

The existence is easy, since we can find one polynomial as follows

$$P(x) = \sum_{k=1}^{n} f(x_k) L_{n,k}(x)$$

where, for each $k = 1, \ldots, n$,

$$L_{n,k}(x) = \frac{(x - x_1)(x - x_2)\cdots(x - x_{k-1})(x - x_k)\cdots(x - x_n)}{(x_k - x_1)(x_k - x_2)\cdots(x_k - x_{k-1})(x_k - x_{k+1})\cdots(x_k - x_n)} = \prod_{\substack{i=1 \\ i \neq k}}^{n} \frac{(x - x_i)}{(x_k - x_i)}$$

It's obvious that

$$L_{n,k}(x_i) = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{if } i \neq k \end{cases}$$

Hence, we can easily verify that $P(x_i) = f(x_i)$ for $i = 1, \ldots, n$. Also, this polynomial is of degree $n-1$, because each $L_{n,k}$ is of degree $n-1$.

To prove the uniqueness, assume that $Q(x)$ is another polynomial of degree $n-1$ agreeing with f at $x_1, \ldots, x_n$. Consider the polynomial $D = P - Q$, we have

$$D(x_k) = P(x_k) - Q(x_k) = f(x_k) - f(x_k) = 0 \quad \text{for } k = 1, \ldots, n$$

Thus each x_k is a root of $D(x)$ with at least multiplicity one, which means

$$D(x) = (x - x_1)\cdots(x - x_n)R(x)$$

where $R(x)$ is another polynomial. However, this is impossible, because $D(x)$ is at most of degree $n-1$, and now you need degree of at least n to obtain n roots with multiplicity one. The only possible is that $R(x) \equiv 0$, which means $D(x) \equiv 0$. Hence $Q(x) = P(x)$, showing that $P(x)$ is unique.

Therefore, such $P(x)$ is exactly the $L(x)$ we need to find. Finally, we need to prove $L(x)$ satisfies

$$f(x) - L(x) = \frac{(x - x_1)\cdots(x - x_n)}{n!} f^{(n)}(\xi)$$

Note that if $x = x_k$, for any $k = 1, \ldots, n$, then $f(x_k) = L(x_k)$, and choosing any ξ yields the above result.

If $x \neq x_k$, for all $k = 1, \ldots, n$, define $g(t)$ on I, such that

$$g(t) = f(t) - L(t) - [f(x) - L(x)] \prod_{i=1}^{n} \frac{(t - x_i)}{(x - x_i)}$$

It's easy to see $g(t) \in \mathcal{C}^{(n)}(I)$, and for $t = x_k$, $g(x_k) = 0$. Moreover,

$$g(x) = f(x) - L(x) - [f(x) - L(x)] \prod_{i=1}^{n} \frac{(x - x_i)}{(x - x_i)} = 0$$

Hence, g vanishes at $n+1$ distinct numbers $x, x_1, \dots, x_n$. By part a), there exists a number $\xi \in I$, such that $g^{(n)}(\xi) = 0$. Thus,

$$0 = g^{(n)}(\xi) = f^{(n)}(\xi) - L^{(n)}(\xi) - [f(x) - L(x)] \frac{d^{n+1}}{dt^{n+1}} \left[\prod_{i=1}^{n} \frac{(t - x_i)}{(x - x_i)} \right]_{t=\xi}$$

Since L is of degree $n-1$, so $L^{(n)}(t) = 0$. Also, the last term is a polynomial of degree n, so its derivative is just the product of its leading coefficient and $n!$, i.e.,

$$\frac{d^{n+1}}{dt^{n+1}} \left[\prod_{i=1}^{n} \frac{(t - x_i)}{(x - x_i)} \right]_{t=\xi} = \frac{n!}{\prod_{i=1}^{n}(x - x_i)}$$

Therefore, we could yield

$$0 = f^{(n)}(\xi) - [f(x) - L(x)] \frac{n!}{\prod_{i=1}^{n}(x - x_i)}$$

which is equivalent to

$$f(x) - L(x) = \frac{(x - x_1) \cdots (x - x_n)}{n!} f^{(n)}(\xi)$$

c) If $x_1 < x_2 < \cdots < x_p$ are points of I and n_i, $1 \le i \le p$, are natural numbers such that $n_1 + n_2 + \cdots + n_p = n$ and $f^{(k)}(x_i) = 0$ for $0 \le k \le n_i - 1$, then there exists a point ξ in the closed interval $[x_1, x_p]$ at which $f^{(n-1)}(\xi) = 0$.

Since $n_1 + n_2 + \cdots + n_p = n$, for all n_i, we have $0 \le n_i \le n$. Denote m_i as the number of j such that $n_j = i$. Therefore,

$$0 \cdot m_0 + 1 \cdot m_1 + \cdots + n \cdot m_n = n$$

For simplicity, we define $S_k = \sum_{i=k}^{n} m_i$, then the above equation is equivalent to

$$0 + S_1 + S_2 + \cdots = \sum_{k=1}^{n} S_k = n$$

In this case, we can choose S_1 points $z_1^0 < \cdots < z_{S_1}^0$ out of x_i (hence these points are distinct), such that $f(z_i^0) = 0$. By Rolle's theorem, there exists $S_1 - 1$ points $z_1^1 < \cdots < z_{S_1-1}^1$, such that $f'(z_i^1) = 0$, and each z_i^1 lies in the open interval between z_i^0 and z_{i+1}^0. Notice that except for the $S_1 - 1$ points we find, originally there exists S_2 points which are the roots of $f'(x)$. These S_2 points are all distinct from the previous $S_1 - 1$ ones (This is essential, you should consider the reason). In total, we have $S_1 + S_2 - 1$ distinct roots of $f'(x)$, meaning that we can find $S_1 + S_2 - 2$ points $z_1^2 < \cdots < z_{S_1+S_2-2}^2$ such that $f''(z_i^2) = 0$, where z_i^2 lies in open interval between z_i^1 and z_{i+1}^1. By induction, you can proceed this process until you find $S_1 + \cdots + S_{n-1} - (n-1)$ points such that $f^{(n-1)}(x_i^{n-1}) = 0$. Again, don't forget originally you have S_n points satisfying the same condition. Hence, the total number of points that satisfies $f^{(n-1)}(x_i^{n-1}) = 0$ is

$$S_1 + \cdots + S_{n-1} + S_n - (n-1) = n - (n-1) = 1$$

Obviously, such $\xi = x_1^{n-1}$ lies in $[x_1, x_p]$.

d) There exists a unique polynomial $H(x)$ of degree $(n-1)$ such that $f^{(k)}(x_i) = H^{(k)}(x_i)$ for $0 \leq k \leq n_i - 1$. Moreover, inside the smallest interval containing the points x and x_i, $i = 1, \ldots, p$, there is a point ξ such that

$$f(x) = H(x) + \frac{(x-x_1)^{n_1} \cdots (x-x_n)^{n_p}}{n!} f^{(n)}(\xi)$$

Suppose there is $H(x)$ such that

$$H(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{n-1} x^{n-1}$$

Then for each i and k, $f^{(k)}(x_i) = H^{(k)}(x_i)$ is one equation, and we have n such equations in total, so we can form an $n \times n$ linear system $A\vec{x} = \vec{b}$, where $\vec{x}$ is the coefficients of $H(x)$, $\vec{b}$ is all $f^{(k)}(x_i)$. If we can prove A is nonsingular, then $\vec{x}$ is unique, and $H(x)$ must exist and is unique.

To prove A is nonsingular, consider its null space, $A\vec{x} = \vec{0}$. This shows $H^{(k)}(x_i) = 0$ for all i, k. Thus, $H(x) = C(x) \sum_{i=1}^{p} (x - x_i)^{n_i}$. Since $H(x)$ is only of order $n-1$, $C(x)$ can only equal to zero, meaning that $H(x) \equiv 0$. Thus, $\vec{x} = \vec{0}$. Since the null space of A only contains the trival element, its rank is zero, meaning that the column space of A is of rank n. Thus, A is full rank and full rank matrix must be nonsingular.

Note that if $x = x_i$, where $n_i \geq 1$, then for any i, $f(x_i) = H(x_i)$, and choosing any ξ yields the above result.

If $x \neq x_i$, where $n_i \geq 1$, then we define $g(t)$ on I, such that

$$g(t) = f(t) - H(t) - [f(x) - H(x)] \prod_{i=1}^{p} \frac{(t-x_i)^{n_i}}{(x-x_i)^{n_i}}$$

Similar to part b), you can verify that $g^{(k)}(x_i) = 0$ for any $i = 1, \cdots, p$ and $0 \leq k \leq n_i - 1$. Notice that $g(x) = 0$, so we can denote x as x_{p+1}, with $n_{p+1} = 1$, then $n_1 + \cdots + n_{p+1} = n+1$. By part c), there exists a $\xi \in I^*$ where I^* is the smallest interval containing x_i, $i = 1, \cdots, p+1$, such that $g^{(n)}(\xi) = 0$, which means

$$0 = g^{(n)}(\xi) = f^{(n)}(\xi) - 0 - [f(x) - H(x)] \frac{d^n}{dt^n} \left[\prod_{i=1}^{p} \frac{(t-x_i)^{n_i}}{(x-x_i)^{n_i}} \right]_{t=\xi}$$

which is equivalent to say

$$0 = f^{(n)}(\xi) - [f(x) - H(x)] \prod_{i=1}^{p} \frac{n!}{(x-x_i)^{n_i}}$$

Therefore, we have

$$f(x) = H(x) + \frac{(x-x_1)^{n_1}\cdots(x-x_n)^{n_p}}{n!} f^{(n)}(\xi)$$

Question 5.3-12. Show that

a) between two real roots of a polynomial $P(x)$ with real coefficients there is a root of its derivative $P'(x)$;

We only consider when two real roots are different, for the case that they are equal, see part b). If x_1, x_2 are two real roots of $P(x)$, we have $P(x_1) = P(x_2)$. Since $P(x)$ is a polynomial, it is continuous and differentiable on $\mathbb{R}$, so we can apply Rolle's Theorem, which means there exists $\xi \in (x_1, x_2)$, such that $P'(\xi) = 0$. Hence, between two real roots of $P(x)$ there is a root of $P'(x)$.

b) if the polynomial $P(x)$ has a multiple root, the polynomial $P'(x)$ has the same root, but its multiplicity as a root of $P'(x)$ is one less than its multiplicity as a root of $P(x)$;

Suppose this root has multiplicity of order $m \geq 2$, then it can be written as

$$P(x) = (x-x_0)^m \widetilde{P}(x)$$

where $\widetilde{P}(x_0) \neq 0$. Take the derivative, we have

$$P'(x) = m(x-x_0)^{m-1}\widetilde{P}(x) + (x-x_0)^m \widetilde{P}'(x)$$

Thus, we could yield $P'(x_0) = 0 + 0 = 0$. Hence, the polynomial $P'(x)$ has the same root as $P(x)$.

Since

$$P'(x) = (x-x_0)^{m-1}(m\widetilde{P}(x) + (x-x_0)\widetilde{P}'(x))$$

we consider a new polynomial $R(x)$, where

$$R(x) = m\widetilde{P}(x) + (x-x_0)\widetilde{P}'(x)$$

One could easily see that $R(x_0) \neq 0$ because

$$R(x_0) = m\widetilde{P}(x_0) + (x_0 - x_0)\widetilde{P}'(x) = m\widetilde{P}(x_0) + 0 \neq 0$$

Thus $P'(x)$ only has root x_0 of order $m-1$, which is one less than that of $P(x)$.

c) if $Q(x)$ is the greatest common divisor of the polynomials $P(x)$ and $P'(x)$, where $P'(x)$ is the derivative of $P(x)$, then the polynomial $\frac{P(x)}{Q(x)}$ has the roots of $P(x)$ as its roots, all of them being roots of multiplicity 1.

From part b), we know that if $P(x)$ has a real root x_0, then we have $P(x) = (x-x_0)^m\widetilde{P}(x)$, where $\widetilde{P}'(x_0) \neq 0$. Also, we can write $P'(x) = (x-x_0)^{m-1}R(x)$, where $R(x_0) \neq 0$. This shows that their great common divisor $Q(x) = (x-x_0)^{m-1}\widetilde{Q}(x)$, where $\widetilde{Q}(x_0) \neq 0$.

Consider

$$\frac{P(x)}{Q(x)} = \frac{(x-x_0)^m \widetilde{P}(x)}{(x-x_0)^{m-1}\widetilde{Q}(x)} = \frac{(x-x_0)\widetilde{P}(x)}{\widetilde{Q}(x)}$$

We can obtain that $P(x_0)/Q(x_0) = 0/\widetilde{Q}(x_0) = 0$ because $\widetilde{Q}(x_0) \neq 0$. Hence, the polynomial $P(x)/Q(x)$ has the same roots of $P(x)$. Also, we can write

$$\frac{P(x)}{Q(x)} = (x-x_0)\frac{\widetilde{P}(x)}{\widetilde{Q}(x)}$$

where $\widetilde{P}(x_0) \neq 0$ and $\widetilde{Q}(x_0) \neq 0$, thus $\widetilde{P}(x_0)/\widetilde{Q}(x_0) \neq 0$. This proves that the roots of $\frac{P(x)}{Q(x)}$ only have multiplicity 1.

Question 5.3-13. Show that

a) any polynomial $P(x)$ admits a representation in the form $c_0 + c_1(x-x_0) + \cdots + c_n(x-x_0)^n$;

Apply the Taylor's expansion to any polynomial $P_n(x)$ with Lagrange form of remainder, we have

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \cdots + \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \frac{f^{(k+1)}(\xi)}{(k+1)!}(x-x_0)^{k+1}$$

However, here $f = P_n(x)$ and if $k \geq n$, $f^{(k+1)}(x) \equiv 0$. Hence,

$$P_n(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$$

Denote $c_0 = f(x_0)$, $c_1 = f'(x_0)$, and so on, we can conclude that for any $P_n(x)$, we have

$$P_n(x) = c_0 + c_1(x-x_0) + c_2(x-x_0)^2 + \cdots + c_n(x-x_0)^n$$

b) there exists a unique polynomial of degree n for which $f(x) - P(x) = o\big((x-x_0)^n\big)$ as $E \ni x \to x_0$. Here f is a function defined on a set E and x_0 is a limit point of E.

(Here $f(x)$ must be at least $\mathcal{C}^n$ function. Hence the existence can be ensured by Taylor expansion.)

Suppose there exist two polynomials satisfies the condition, denote them as $P_(x)$ and $Q(x)$. Then we have

$$D_n(x) = P_(x) - Q(x) = o\big((x-x_0)^n\big) \quad \text{as } x \to x_0$$

Since $D_n(x)$ is a polynomial of degree n, we can write

$$D_n(x) = a_n(x-x_0)^n + a_{n-1}(x-x_0)^{n-1} + \cdots + a_0$$

Since $D_n(x) = o\big((x-x_0)^n\big)$ as $x \to x_0$, we have

$$\lim_{x\to x_0} \frac{a_n(x-x_0)^n + a_{n-1}(x-x_0)^{n-1} + \cdots + a_0}{(x-x_0)^n} = 0$$

However, this is true only if all coefficients are zero, because if there exists some nonzero coefficient(s), then the one with the highest order of n in denominator will be the dominant term, and since all terms above tend to infinity as $x \to x_0$, it will never vanish. Hence, all coefficients are zero, $D_n(x) \equiv 0$. This shows $P(x) = Q(x)$, which means there exists a unique polynomial.

Question 5.3-15.

a) Applying Lagrange's theorem to the function $\frac{1}{x^\alpha}$, where $\alpha > 0$, show that the inequality

$$\frac{1}{n^{1+\alpha}} < \frac{1}{\alpha}\left(\frac{1}{(n-1)^\alpha} - \frac{1}{n^\alpha}\right)$$

holds for $n \in \mathbb{N}$ and $\alpha > 0$.

Applying Lagrange's theorem to the function $\frac{1}{x^\alpha}$, for $n \geq 2$, we have

$$f(n) - f(n-1) = f'(\xi)[n-(n-1)] \Longrightarrow \left(\frac{1}{n^\alpha} - \frac{1}{(n-1)^\alpha}\right) = -\alpha\xi^{\alpha-1}$$

Since $\alpha > 0$, we have $-\alpha\xi^{\alpha-1}$ is increasing for $\xi \in [n-1, n]$, which shows

$$\left(\frac{1}{n^\alpha} - \frac{1}{(n-1)^\alpha}\right) \leq -\frac{\alpha}{n^{\alpha+1}}$$

Slightly rearrange the terms we will obtain

$$\frac{1}{n^{1+\alpha}} < \frac{1}{\alpha}\left(\frac{1}{(n-1)^\alpha} - \frac{1}{n^\alpha}\right)$$

b) Use the result of a) to show that the series $\sum\limits_{n=1}^{\infty} \frac{1}{n^\sigma}$ converges for $\sigma > 1$.

Using what we prove in part a), let $\sigma = \alpha + 1$ where $\alpha > 0$, we have

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{1}{n^{\alpha+1}} &< \sum_{n=1}^{\infty} \frac{1}{\alpha}\left(\frac{1}{(n-1)^\alpha} - \frac{1}{n^\alpha}\right) \\ &= 1 + \lim_{n\to\infty} \frac{1}{\alpha}\left(1 - \frac{1}{n^\alpha}\right) \\ &= 1 + \frac{1}{\alpha}\end{aligned}$$

Hence positive series $\sum\limits_{n=1}^{\infty} \frac{1}{n^\sigma}$ converges when $\sigma > 1$.

Question 5.4-1. Let $x = (x_1, \ldots, x_n)$ and $\alpha = (\alpha_1, \ldots, \alpha_n)$, where $x_i \geq 0$, $\alpha_i > 0$ for $i = 1, \ldots, n$ and $\sum\limits_{i=1}^{n} \alpha_i = 1$. For any number $t \neq 0$ we consider the *mean of order t of the numbers $x_1, \ldots, x_n$ with weights α_i*:

$$M_t(x, \alpha) = \left(\sum_{i=1}^{n} \alpha_i x_i^t\right)^{1/t}$$

In particular, when $\alpha_1 = \cdots = \alpha_n = \frac{1}{n}$, we obtain the harmonic, arithmetic, and quadratic means for $t = -1, 1, 2$ respectively.

Show that

a) $\lim_{t\to 0} M_t(x,\alpha) = x_1^{\alpha_1}\cdots x_n^{\alpha_n}$, that is, in the limit one can obtain the geometric mean;

If one of x_i, say x_k, is zero, then we need to consider the left and right limit separately. If $t > 0$, then $0^t = 0$; if $t < 0$, then $0^t = \infty$. We can see

$$M_t(x,\alpha) = \left(\sum_{\substack{i=1\\ i\neq k}}^{n} \alpha_i x_i^t + \alpha_k 0^t\right)^{1/t}$$

Thus,

$$\lim_{t\to 0+}\left(\sum_{\substack{i=1\\ i\neq k}}^{n} \alpha_i x_i^t + \alpha_k 0^t\right)^{1/t} = \lim_{t\to 0+}\left(\sum_{\substack{i=1\\ i\neq k}}^{n} \alpha_i x_i^t\right)^{1/t} = 0$$

because

$$\lim_{t\to 0+}\sum_{\substack{i=1\\ i\neq k}}^{n} \alpha_i x_i^t = 1-\alpha_k < 1$$

and $1/t \to +\infty$. However, for the left limit, we need to be careful,

$$\lim_{t\to 0-}\left(\sum_{\substack{i=1\\ i\neq k}}^{n} \alpha_i x_i^t + \alpha_k 0^t\right)^{1/t} = \lim_{t\to 0-}\left(\alpha_k 0^t\right)^{1/t} = 0$$

Although 0^t is not defined when $t \leq 0$, but we can regard $f(t) = (0^t)^{1/t} = 0$ as a function defined on $t \neq 0$. Hence, we conclude that when there exists $x_i = 0$,

$$\lim_{t\to 0} M_t(x,\alpha) = 0 = x_1^{\alpha_1}\cdots x_{k-1}^{\alpha_{k-1}}\cdot 0^{\alpha_k}\cdot x_{k+1}^{\alpha_{k+1}}\cdots x_n^{\alpha_n}$$

If all x_i is nonzero, it would be a standard problem, we first take logarithm and then apply L'Hôpital's rule, then

$$\begin{aligned}
\lim_{t\to 0}\left(\sum_{i=1}^{n}\alpha_i x_i^t\right)^{1/t} &= \exp\left\{\lim_{t\to 0}\frac{1}{t}\ln\left(\sum_{i=1}^{n}\alpha_i x_i^t\right)\right\}\\
&= \exp\left\{\lim_{t\to 0}\frac{\sum_{i=1}^{n}\alpha_i x_i^t \ln x_i}{\sum_{i=1}^{n}\alpha_i x_i^t}\right\}\\
&= \exp\left\{\frac{\sum_{i=1}^{n}\alpha_i \ln x_i}{\sum_{i=1}^{n}\alpha_i}\right\}\\
&= \exp\left\{\sum_{i=1}^{n}\alpha_i \ln x_i\right\}\\
&= x_1^{\alpha_1}\cdots x_n^{\alpha_n}
\end{aligned}$$

b) $\lim_{t\to +\infty} M_t(x,\alpha) = \max_{1\leq i\leq n} x_i$;

This is also standard, suppose there $x_k \neq 0$ is the largest one (not necessarily unique), then

$$\begin{aligned}\lim_{t\to+\infty}\left(\sum_{i=1}^{n}\alpha_i x_i^t\right)^{1/t} &= \lim_{t\to+\infty}\left(x_k^t\sum_{i=1}^{n}\alpha_i\left(\frac{x_i}{x_k}\right)^t\right)^{1/t}\\ &= x_k\lim_{t\to+\infty}\left(\sum_{i=1}^{n}\alpha_i\left(\frac{x_i}{x_k}\right)^t\right)^{1/t}\\ &= x_k = \max_{1\le i\le n} x_i\end{aligned}$$

This is because $x_i/x_k \le 1$, and

$$\sum_{i=1}^{n}\alpha_i\left(\frac{x_i}{x_k}\right)^t = a_k + \sum_{\substack{i=1\\ i\neq k}}^{n}\alpha_i\left(\frac{x_i}{x_k}\right)^t \le 1$$

Hence, we have (notice that $a_k > 0$)

$$1 = \lim_{t\to+\infty}\alpha_k^{1/t} \le \lim_{t\to+\infty}\left(\sum_{i=1}^{n}\alpha_i\left(\frac{x_i}{x_k}\right)^t\right)^{1/t} \le \lim_{t\to+\infty} 1^{1/t} = 1$$

If $x_k = 0$, then all $x_i = 0$, and $M_t(x,\alpha) \equiv 0$, which obviously satisfies what we need to prove.

c) $\lim_{t\to-\infty} M_t(x,\alpha) = \min_{1\le i\le n} x_i$;

If none of x_i is zero, then it is similar to part b), suppose x_k is the smallest one (not necessarily unique)

$$\begin{aligned}\lim_{t\to-\infty}\left(\sum_{i=1}^{n}\alpha_i x_i^t\right)^{1/t} &= \lim_{t\to+\infty}\left(\sum_{i=1}^{n}\alpha_i x_i^{-t}\right)^{-1/t}\\ &= \lim_{t\to+\infty}\left(x_k^{-t}\sum_{i=1}^{n}\alpha_i\left(\frac{x_i}{x_k}\right)^{-t}\right)^{-1/t}\\ &= x_k\lim_{t\to+\infty}\left(\sum_{i=1}^{n}\alpha_i\left(\frac{x_k}{x_i}\right)^t\right)^{-1/t}\\ &= x_k = \min_{1\le i\le n} x_i\end{aligned}$$

This is because $x_k/x_i \le 1$, and

$$1 = \lim_{t\to+\infty}\alpha_k^{-1/t} \ge \lim_{t\to+\infty}\left(\sum_{i=1}^{n}\alpha_i\left(\frac{x_i}{x_k}\right)^t\right)^{-1/t} \ge \lim_{t\to+\infty} 1^{-1/t} = 1$$

When some $x_i = 0$, then $x_k = 0$ must be the smallest one. When $t < 0$, we assume that $0^t > x_i^t$ for any nonzero x_i. Then

$$a_k 0^t \le \sum_{i=1}^{n}\alpha_i x_i^t \le \sum_{i=1}^{n}\alpha_i 0^t = 0^t$$

which implies

$$0 = \lim_{t\to-\infty}\left(a_k 0^t\right)^{1/t} \ge \lim_{t\to-\infty}\left(\sum_{i=1}^{n}\alpha_i x_i^t\right)^{1/t} \ge \lim_{t\to-\infty}\left(0^t\right)^{1/t} = 0$$

Hence, we have

$$\lim_{t\to-\infty}\left(\sum_{i=1}^{n}\alpha_i x_i^t\right)^{1/t}=0=x_k=\min_{1\le i\le n} x_i$$

d) $M_t(x,\alpha)$ is a nondecreasing function of t on $\mathbb{R}$ and is strictly increasing if $n>1$ and the numbers x_i are all nonzero.

If $n=1$, $M_t(x,\alpha)=x_1$, which is constant function, hence it is a nondecreasing function. If some x_i is zero, then $M_t(x,\alpha)=0$ for all $t<0$. The reason is the same as part c), for all $t<0$,

$$0=a_k^{1/t}\cdot 0\ge \left(a_k 0^t\right)^{1/t}\ge\left(\sum_{i=1}^{n}\alpha_i x_i^t\right)^{1/t}\ge\left(0^t\right)^{1/t}=0$$

If all x_i is nonzero, then we can apply standard procedure. Consider

$$g(t)=\frac{1}{t}\ln\left(\sum_{i=1}^{n}\alpha_i x_i^t\right)$$

We can show that

$$g'(t)=\frac{1}{t^2}\left[t\cdot\frac{\sum_{i=1}^{n}\alpha_i x_i^t\ln x_i}{\sum_{i=1}^{n}\alpha_i x_i^t}-\ln\left(\sum_{i=1}^{n}\alpha_i x_i^t\right)\right]$$

Denote

$$h(t)=t\cdot\frac{\sum_{i=1}^{n}\alpha_i x_i^t\ln x_i}{\sum_{i=1}^{n}\alpha_i x_i^t}-\ln\left(\sum_{i=1}^{n}\alpha_i x_i^t\right)$$

We can show that

$$h'(t)=\frac{t}{\sum_{i=1}^{n}\alpha_i x_i^t}\left[\left(\sum_{i=1}^{n}\alpha_i x_i^t\ln^2 x_i\right)\left(\sum_{i=1}^{n}\alpha_i x_i^t\right)-\left(\sum_{i=1}^{n}\alpha_i x_i^t\ln x_i\right)^2\right]$$

By Cauchy-Schwarz inequality (let $u_i=\sqrt{a_i x_i^t}\ln x_i$ and $v_i=\sqrt{a_i x_i^t}$), we can conclude that $h'(t)>0$ when $t>0$, $h'(t)<0$ when $t<0$. Hence, $h(t)$ decreasing when $t<0$, increasing when $t>0$. Since $h(0)=0$, $h(t)>0$ for $t\neq 0$. Thus $g'(t)>0$ for $t\neq 0$. Hence, $g(t)$ is strictly increasing when $t<0$ and $t>0$. In conclusion, $M_t(x,\alpha)$ is strictly increasing when $n>1$ and all $x_i>0$.

Similar procedure can be applied to the case that some x_i is zero for $t>0$, and the result shows $M_t(x,\alpha)$ is still strictly increasing when $t>0$ (You can just delete those terms that has zero x_i, because $0^t=0$ for $t>0$). However, previously we show that it is constant zero when $t<0$, we can only say it is a nondecreasing function for t on $\mathbb{R}$.

Question 5.4-2. Show that $|1+x|^p\ge 1+px+c_p\varphi_p(x)$, where c_p is a constant depending only on p,

$$\varphi_p(x)=\begin{cases}|x|^2 & \text{for } |x|\le 1\\ |x|^p & \text{for } |x|>1\end{cases}\qquad \text{if } 1<p\le 2$$

and $\varphi_p(x)=|x|^p$ on $\mathbb{R}$ if $2<p$.

Since when $x = 0$, $\phi_p(0) = 0$, the inequality always holds, so we don't consider that trivial case. First we consider when $1 < p \leq 2$. For $|x| \leq 1$, let

$$\phi(x) = \frac{(1+x)^p - 1 - px}{x^2}$$

We want to find the minimum value of it. Take the derivative, we have

$$\phi'(x) = \frac{g(x)}{x^3}, \text{ where } g(x) = px(1+x)^{p-1} + px - 2(1+x)^p + 2$$

Also,

$$g'(x) = p(p-1)x(1+x)^{p-2} + p - p(1+x)^{p-1}, \quad \text{and} \quad g''(x) = p(p-1)(p-2)x(1+x)^{p-3}$$

Since $1+x \geq 0$, we know that when $x > 0$, $g''(x) < 0$; $x < 0$, $g''(x) > 0$. This means $g'(x)$ increasing on $[-1, 0)$, and decreasing on $(0, 1]$. The maximum value of $g'(x)$ is $g'(0) = 0$. Hence $g'(x) \leq 0$, meaning that $g(x)$ is decreasing on $[-1, 1]$. But $g(0) = 0$, meaning that $g(x) > 0$ on $[-1, 0)$; $g(x) < 0$ on $(0, 1]$. Hence, $\phi'(x) < 0$, and $\phi(x)$ is decreasing on $[-1, 0)$ and $(0, 1]$. Apply L'Hôpital's rule, we can verify that $\phi(x)$ only has removable discontinuity at $x = 0$, so $\phi(x)$ is decreasing on $[-1, 1]$. The minimum value is $\phi(1) = 2^p - p - 1$.

Although it is tedious, but we can do the same thing for $x > 1$ and $x < -1$. For $x > 1$, we let

$$\phi(x) = \frac{(1+x)^p - 1 - px}{x^p}$$

Similarly, we have

$$\phi'(x) = \frac{pg(x)}{x^{p+1}}, \text{ where } g(x) = -(1+x)^{p-1} - x + 1 + px$$

Also,

$$g'(x) = (p-1)[1 - (1+x)^{p-2}] > 0$$

But since $g(0) = 0$, so $g(x) > 0$ when $x > 1$, meaning that $\phi'(x) > 0$. Hence $\phi(x)$ is increasing on $(1, \infty)$. Let's consider when $x \to \infty$,

$$\lim_{x\to\infty} \phi(x) = \lim_{x\to\infty} \left(1 + \frac{1}{x}\right)^p - \lim_{x\to\infty} \frac{1}{x^p} - \lim_{x\to\infty} \frac{p}{x^{p-1}} = 1$$

This means on $(1, \infty)$, $1 > \phi(x) > \phi(1) = 2^p - p - 1$.

For $x < -1$, by exactly the same method, set

$$\phi(x) = \frac{(-1-x)^p - 1 - px}{(-x)^p}$$

We can show that there exists a unique x_p (only depends on p) such that $\phi(x)$ increasing on $(-\infty, x_p)$ and decreasing on $(x_p, -1)$. Check the limit when $x \to -\infty$, we have

$$\lim_{x\to-\infty} \phi(x) = 1$$

Hence $\phi(x) > 1$ on $(-\infty, -1)$, and since $\phi(-1) > \phi(1) = 2^p - p - 1$, we conclude that when $1 < p \leq 2$, the minimum value of $\phi(x)$ for $x \in \mathbb{R}$ is $2^p - p - 1$. Thus, the maximum we can take is $c_p = 2^p - p - 1$, which only depends on p.

Then we consider $p > 2$. When $x > 0$, define

$$\phi(x) = \frac{(1+x)^p - 1 - px}{x^p}$$

Similarly, we have

$$\phi'(x) = \frac{pg(x)}{x^{p+1}}, \text{ where } g(x) = -(1+x)^{p-1} - x + 1 + px$$

Also,

$$g'(x) = (p-1)[1 - (1+x)^{p-2}] < 0$$

Since $g(0) = 0$, so $g(x) < 0$ when $x > 0$, meaning that $\phi'(x) < 0$. Hence $\phi(x)$ is decreasing on $(0, \infty)$. Now you need to be careful, because the discontinuity of $\phi(x)$ at $x = 0$ is not removable, actually when $x \to 0+$, by L'Hôpital's rule, we have

$$\lim_{x \to 0+} \phi(x) = \lim_{x \to 0+} \left(1 + \frac{1}{x}\right)^{p-2} = +\infty$$

The limit of $\phi(x)$ as $x \to \infty$ is the same as before, which is just 1. Hence, $\phi(x) > 1$ for $x > 0$.

Finally, when $x < 0$, we need to further separately consider when $x < -1$ and $x \in [-1, 0)$. For the case $x \in [-1, 0)$, we define

$$\phi(x) = \frac{(1+x)^p - 1 - px}{(-x)^p}$$

You can easily check that $\phi(x)$ is increasing on $[-1, 0)$, and the limit as $x \to 0-$ is also $+\infty$. Hence, the minimum is obtained by $\phi(-1) = p - 1 > 1$. Thus, the global minimum is not in $[-1, 0)$.

Consider if $x < -1$, we define

$$\phi(x) = \frac{(-1-x)^p - 1 - px}{(-x)^p}$$

We have

$$\phi'(x) = \frac{pg(x)}{x^{p+1}}, \text{ where } g(x) = -(-1-x)^{p-1} + x - 1 - px$$

If we consider

$$g'(x) = (p-1)[(-1-x)^{p-2} - 1] > 0$$

Then $\phi'(x)$ is increasing on $(-\infty, -1)$, and $\phi'(-2) = 2p - 4 > 0$, $\phi'(-6) = 6p - 7 - 5^{p-1} < 0$. For the second inequality, consider $h(p) = 6p - 7 - 5^{p-1}$, $h'(p) = 6 - 5^{p-1}\ln 5 < h'(2) = 6 - 5\ln 5 < 0$, so $h(p)$ decreasing when $p > 2$. But $h(2) = 0$, so $h(p) < 0$. Thus, $\phi'(x) = 0$ has a unique solution x_p in $(-6, -2)$ for any $p > 2$, where x_p only depends on p. We also know that $\phi(x)$ is decreasing on $(-\infty, x_p)$ and increasing on $(x_p, -1)$. Since we know that $\phi(x_p) < \phi(-2) = p/2^{p-1} < 1$, so x_p is the global minimum. (Also, you can show that the limit of $\phi(x)$ as $x \to -\infty$ is 1.) In this way, the largest c_p we can take is $c_p = \phi(x_p)$.

Actually, we can find that $\phi(x_p) < \phi(-2) = p/2^{p-1}$, taking $p \to \infty$, we have $\phi(x_p) \le 0$. But $\phi(x)$ is always nonnegative, so the limit of $\phi(x_p)$ is zero. Thus, the largest c_p we can take will tend to zero as p grows, and x_p will tend to -2.

Question 5.4-3. Verify that $\cos x < \left(\frac{\sin x}{x}\right)^3$ for $0 < |x| < \frac{\pi}{2}$.

Consider the function $f(x)$ for $|x| < \frac{\pi}{2}$ defined by

$$f(x) = \frac{\sin x}{(\cos x)^{1/3}} - x$$

Take the first derivative, we have

$$\begin{aligned}
f'(x) &= \frac{(\cos x)^{4/3} + (1/3)\sin^2 x(\cos x)^{-2/3}}{(\cos x)^{2/3}} - 1 \\
&= \frac{(\cos x)^{4/3} + (1/3)(1-\cos^2 x)(\cos x)^{-2/3} - (\cos x)^{2/3}}{(\cos x)^{2/3}} \qquad \text{Take } t = (\cos x)^{2/3} \\
&= \frac{3t^3 + (1-t^3) - 3t^2}{3t^2} \\
&= \frac{(2t+1)(t-1)^2}{3t^2} \qquad \text{Since } t \in (0,1) \\
&> 0
\end{aligned}$$

Therefore, $f'(x) > 0$ for $0 < |x| < \frac{\pi}{2}$, meaning $f(x)$ is increasing on $(-\frac{\pi}{2}, \frac{\pi}{2})$. Notice that $f(0) = 0$, hence $f(x) > 0$ on $(0, \frac{\pi}{2})$ and $f(x) < 0$ on $(-\frac{\pi}{2}, 0)$.

When $x \in (0, \frac{\pi}{2})$, we have $x^3 > 0$ and $\cos x > 0$, so

$$\frac{\sin x}{(\cos x)^{1/3}} - x > 0 \Longrightarrow \frac{\sin x}{(\cos x)^{1/3}} > x \Longrightarrow \frac{\sin^3 x}{\cos x} > x^3 \Longrightarrow \frac{\sin^3 x}{x^3} > \cos x$$

When $x \in (-\frac{\pi}{2}, 0)$, we have $x^3 < 0$ and $\cos x > 0$, so

$$\frac{\sin x}{(\cos x)^{1/3}} - x < 0 \Longrightarrow \frac{\sin x}{(\cos x)^{1/3}} < x \Longrightarrow \frac{\sin^3 x}{\cos x} < x^3 \Longrightarrow \frac{\sin^3 x}{x^3} > \cos x$$

Therefore, for $0 < |x| < \frac{\pi}{2}$, we have

$$\cos x < \left(\frac{\sin x}{x}\right)^3$$

Question 5.4-4. Study the function $f(x)$ and construct its graph if

a) $f(x) = \arctan \log_2 \cos\left(\pi x + \frac{\pi}{4}\right)$;

b) $f(x) = \arccos\left(\frac{3}{2} - \sin x\right)$;

c) $f(x) = \sqrt[3]{x(x+3)^2}$.

d) Construct the curve defined in polar coordinates by the equation $\varphi = \frac{\rho}{\rho^2+1}$, $\rho \geq 0$, and exhibit its asymptotics.

e) Show how, knowing the graph of the function $y = f(x)$, one can obtain the graph of the following functions $f(x) + B$, $Af(x)$, $f(x+b)$, $f(ax)$, and, in particular $-f(x)$ and $f(-x)$.

Question 5.4-5. Show that if $f \in \mathcal{C}(a,b)$ and the inequality

$$f\left(\frac{x_1 + x_2}{2}\right) \leq \frac{f(x_1) + f(x_2)}{2}$$

holds for any points $x_1, x_2 \in (a, b)$, then the function f is convex on (a, b).

We need to prove for all $x \in (a, b)$, the inequality

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$$

holds for all $\lambda \in [0, 1]$. First we only consider if $\lambda = \frac{k}{2^n}$, where $n = 0, 1, \ldots$ and $k \in \mathbb{N}$, $k \leq 2^n$.

When $n = 0$, $\lambda = 0, 1$, the inequality is obvious correct. When $n = 1$, $\lambda = 0, \frac{1}{2}, 1$, the inequality is also true because of the hypothesis in the question. Hence, we suppose it is true for n, and we attempt to verify it is also true for $n+1$. For any $\lambda = \frac{k}{2^{n+1}}$, if k is even, then write $k = 2m$, we can reduce λ to $\frac{m}{2^n}$. This must hold because of our assumption.

If k is odd, then write $k = (k-1)/2 + (k+1)/2$, denote $s = (k-1)/2$, $t = (k+1)/2$, and we have integers $s, t \in [0, 2^n]$.

$$\lambda = \frac{k}{2^{n+1}} = \frac{\frac{s}{2^n} + \frac{t}{2^n}}{2} = \frac{p+q}{2}$$

where $p = \frac{s}{2^n}$, $q = \frac{t}{2^n}$. Thus, we have

$$\begin{aligned}
f(\lambda x + (1-\lambda)y) &= f\left(\frac{[px + (1-p)y] + [qx + (1-q)y]}{2}\right) \\
&\leq \frac{f(px + (1-p)y) + f(qx + (1-q)y)}{2} \\
&\leq \frac{[pf(x) + (1-p)f(y)] + [qf(x) + (1-q)f(y)]}{2} \\
&= \frac{p+q}{2} f(x) + \left(1 - \frac{p+q}{2}\right) f(y) \\
&= \lambda f(x) + (1-\lambda)f(y)
\end{aligned}$$

Hence, by induction, we proved that for any n, the Jensen's inequality holds for $\lambda = \frac{k}{2^n}$, where $k \in \mathbb{N}$, $k \leq 2^n$.

Now consider function $g(\lambda) : D \mapsto \mathbb{R}$, where D is unknown and

$$g(\lambda) = \lambda f(x) + (1-\lambda)f(y) - f(\lambda x + (1-\lambda)y)$$

Since f is continuous, it is easy to check g is also continuous. The pre-image of $[0, +\infty)$ (closed set) under g must be closed. Also, the pre-image of $[0, +\infty)$ contains $\lambda = \frac{k}{2^n}$, hence it must contain the closure of $\{\frac{k}{2^n}\}$, which is exactly the interval $[0, 1]$. Hence for any $\lambda \in [0, 1]$, $g(\lambda) \geq 0$, which shows

$$\lambda f(x) + (1-\lambda)f(y) \geq f(\lambda x + (1-\lambda)y)$$

Therefore, $f(x)$ is convex by definition.

Question 5.4-6. Show that

a) if a convex function $f : \mathbb{R} \mapsto \mathbb{R}$ is bounded, it is constant;

If a convex function is bounded and not a constant, there exists $x_1 < x_2$, such that $f(x_1) \neq f(x_2)$. W.O.L.G., we can assume $f(x_1) < f(x_2)$. The secant line determined by $(x_1, f(x_1))$ and $(x_2, f(x_2))$ is

$$y_1(x) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1) + f(x_1)$$

Then we claim that for all $x > x_2$, $f(x) \geq y_1(x)$. If not, there exists at least one point $x_0 > x_2 > x_1$, such that $f(x_0) < y_1(x_0)$, which is

$$f(x_0) < \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x_0 - x_1) + f(x_1)$$

Slightly change the form of the above equation, we have

$$\frac{f(x_0) - f(x_1)}{x_0 - x_1} < \frac{f(x_2) - f(x_1)}{x_2 - x_1} \tag{1}$$

However, since $x_2 \in (x_1, x_0)$, and open interval is convex on $\mathbb{R}$, we can find $t \in (0, 1)$, such that $x_2 = tx_1 + (1-t)x_0$, and if we denote the secant going through $(x_1, f(x_1))$ and $(x_0, f(x_0))$ as

$$y_2(x) = \frac{f(x_0) - f(x_1)}{x_0 - x_1}(x - x_1) + f(x_1)$$

we have $f(x_2) = f(tx_1 + (1-t)x_0) \leq tf(x_1) + (1-t)f(x_0) = y_2(x_2)$. This shows that

$$f(x_2) \leq \frac{f(x_0) - f(x_1)}{x_0 - x_1}(x_2 - x_1) + f(x_1)$$

Slightly change the form, we will obtain

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_0) - f(x_1)}{x_0 - x_1} \tag{2}$$

Combine (1) and (2), we can prove our claim that for all $x > x_2$, $f(x) \geq y_1(x)$ by contradiction.

Since $y_1(x)$ is a linear function, and the slope of it is positive, if $x \to +\infty$, $y_1(x)$ will tend to infinity. However, when $x > x_2$, $f(x) \geq y_1(x)$, so $f(x)$ also tends to infinity as $x \to +\infty$, which contradicts the fact that $f(x)$ is bounded.

b) if

$$\lim_{x \to -\infty} \frac{f(x)}{x} = \lim_{x \to +\infty} \frac{f(x)}{x} = 0$$

for a convex function $f : \mathbb{R} \mapsto \mathbb{R}$, then f is constant.

Similar to part a), suppose it is not constant, there exists $x < y \in \mathbb{R}$, such that $f(x) \neq f(y)$. W.O.L.G., we suppose $f(x) < f(y)$. From part a), for any $z > y > x$,

$$f(z) \geq g(z) = \frac{f(y) - f(x)}{y - x}(z - y) + f(y)$$

when $z > 0$,

$$\frac{f(z)}{z} \geq \frac{f(y) - f(x)}{y - x}\frac{z - y}{z} + \frac{f(y)}{z}$$

Let $z \to +\infty$, we have

$$\lim_{z \to +\infty} \frac{f(z)}{z} \geq \lim_{z \to +\infty} \left[\frac{f(y) - f(x)}{y - x}\frac{z - y}{z} + \frac{f(y)}{z}\right] = \frac{f(y) - f(x)}{y - x} > 0$$

which contradicts the assumption that $\lim_{z \to +\infty}[f(x)/x] = 0$. Hence, $f(x)$ is a constant.

c) for any convex function f defined on an open interval $a < x < +\infty$ (or $-\infty < x < a$), the ratio $\frac{f(x)}{x}$ tends to a finite limit or to infinity as x tends to infinity in the domain of definition of the function.

Since the open interval $(a, +\infty)$ and $(-\infty, a)$ are symmetric, we only consider one case, $(-\infty, a)$. We need to prove the function defined on $(-\infty, a-1]$

$$g(x) = \frac{f(x) - f(a-1)}{x - (a-1)}$$

is an increasing function. If not, there exists $x < y < a - 1$, such that

$$\frac{f(x) - f(a-1)}{x - (a-1)} > \frac{f(y) - f(a-1)}{y - (a-1)}$$

Slightly change the form of the above equation, we have

$$\frac{f(x) - f(a-1)}{x - (a-1)}[y - (a-1)] + f(a-1) < f(y)$$

This means $(y, f(y))$ is above the secant line going through $(x, f(x))$ and $(a-1, f(a-1))$. However, since y lies between x and $a-1$, y is a convex combination of them, meaning that $(y, f(y))$ should be under the secant line of $(x, f(x))$ and $(a-1, f(a-1))$. This gives a contradiction, so the function $g(x)$ is increasing.

Since $g(x)$ is increasing, $h(x) = g(-x)$ must be decreasing. If $h(x)$ is bounded, then it will converge to a finite value as $x \to +\infty$; if not, it will diverge to negative infinity as $x \to +\infty$. This means $g(x)$ will converge to a finite value or diverge to infinity as $x \to -\infty$.

Question 5.4-7. Show that if $f : (a, b) \mapsto \mathbb{R}$ is a convex function, then

a) at any point $x \in (a, b)$ it has a left-hand derivative f'_- and a right-hand derivative f'_+, defined as

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

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

and $f'_-(x) \leq f'_+(x)$;

We first prove the existence of one-side derivative of arbitrary convex function. If we denote

$$F(h) = \frac{f(x+h) - f(x)}{h} \qquad h \in (a-x, 0) \cup (0, b-x)$$

Then it suffices to show that one-side limit of $F(h)$ exists at $h = 0$. To show that, we can show a sufficient condition for that, that is, $F(h)$ is nondecreasing with respect to h. If $F(h)$ is nondecreasing on $(a-x, 0) \cup (0, b-x)$, $F(x)$ on $(a-x, 0)$ is bounded above by $f(y), y \in (0, b-x)$, so $F(x)$ has left-hand limit at $x = 0$; $F(x)$ on $(0, b-x)$ is bounded below by $f(y), y \in (a-x, 0)$, so $F(x)$ has right-hand limit at $x = 0$.

There are three cases we need to consider, namely, $h_1 < h_2 < 0$, $0 < h_1 < h_2$, and $h_1 < 0 < h_2$. Here we only prove the most complicated case, i.e., $h_1 < 0 < h_2$. First we write the secant line passing through $(x + h_1, f(x + h_1))$ and $(x + h_2, f(x + h_2))$,

$$y(\xi) = \frac{f(x+h_2) - f(x+h_1)}{h_2 - h_1}(\xi - x - h_1) + f(x + h_1)$$

Since $f(x)$ is convex, we know that $f(x) \leq y(x)$, i.e.,

$$f(x) \leq \frac{f(x+h_2) - f(x+h_1)}{h_2 - h_1}(-h_1) + f(x+h_1)$$

which yields that

$$\frac{f(x+h_2) - f(x)}{h_2} \geq \frac{f(x+h_1) - f(x)}{h_1} \Longleftrightarrow F(h_2) \geq F(h_1)$$

Hence $F(x)$ is nondecreasing for $h_1 < 0 < h_2$. The other two cases are similar, so we conclude that $F(x)$ is nondecreasing on the whole domain $(a-x, 0) \cup (0, b-x)$. According to what we analyze just now, f'_- and $f'_+(x)$ exist.

Now we prove $f'_-(x) \leq f'_+(x)$. For small $h > 0$, we have $F(-h) \leq F(h)$, i.e.,

$$\frac{f(x-h) - f(x)}{-h} \leq \frac{f(x+h) - f(x)}{h}$$

Let $h \to 0+$, the above inequality yields $f'_-(x) \leq f'_+(x)$.

b) the inequality $f'_+(x_1) \leq f'_-(x_2)$ holds for $x_1, x_2 \in (a, b)$ and $x_1 < x_2$;

Let $x = x_1$ in $F(h)$, we have

$$F(h) = \frac{f(x_1+h) - f(x_1)}{h}$$

For $h_1 = \xi - x_1$ and $h_2 = x_2 - x_1$, $\forall\, \xi \in (x_1, x_2)$, we have $h_1 < h_2$, thus $F(h_1) \leq F(h_2)$, i.e.,

$$\frac{f(\xi) - f(x_1)}{\xi - x_1} \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

Similarly, let $x = x_2$ in $F(h)$ and consider $h_1 = x_2 - \xi$, $h_2 = x_2 - x_1$, $F(-h_2) \leq F(-h_1)$, we can obtain

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_2) - f(\xi)}{x_2 - \xi}$$

For the first inequality, let $\xi \to x_1+$, we have

$$f'_+(x_1) \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

Similarly, let $\xi \to x_2-$ in the second inequality, we have

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq f'_-(x_2)$$

Hence $f'_+(x_1) \leq f'_-(x_2)$.

c) the set of cusps of the graph of $f(x)$ (for which $f'_-(x) \neq f'_+(x)$) is at most countable.

For the cusp x_0 of the graph, $f'_-(x_0) < f'_+(x_0)$. Thus we can consider the intervals $(f'_-(x_0), f'_+(x_0))$ for all cusps x_0. We need to prove for distinct x_0, the intervals they form are mutually disjoint. For all x_0, consider any y, z, $y < x_0 < z$, we have

$$f'_-(y) \leq f'_+(x_0) \leq f'_-(x_0) < f'_+(x_0) \leq f'_-(z) < f'_+(z)$$

Hence, as long as x_0 are different, the open interval it forms will be disjoint. In this way, if we pick one $r_0 \in \mathbb{Q}$ in each different interval, they are also distinct. Assign each interval with a rational number in that interval, collect all assigned rational numbers in a set A, and denote the set of all cusps-generated intervals as B, then there exists a bijective mapping from $A \to B$. Since A is at most countable, B is also at most countable. Hence, the set of cusps of the graph of $f(x)$ is at most countable.

Question 5.4-8. The *Legendre transform* of a function $f : I \mapsto \mathbb{R}$ defined on an interval $I \subset \mathbb{R}$ is the function

$$f^*(t) = \sup_{x \in I} \big(tx - f(x)\big)$$

Show that

a) The set I^* of values of $t \in \mathbb{R}$ for which $f^*(t) \in \mathbb{R}$ (that is, $f^*(t) \neq \infty$) is either empty or consists of a single point, or is an interval of the line, and in this last case the function $f^*(t)$ is convex on I^*.

Denote $g_x(t) = tx - f(x)$ for each fixed $x \in I$, then $g_x(t)$ is a linear function with effective domain $\mathbb{R}$, so it is convex on $\mathbb{R}$. This is equivalent to say the epigraph of $g_x(t)$ is a convex set in $\mathbb{R}^2$. Therefore, if we take the supremum of all $g_x(t)$, the epigraph of the resulting function $f^*(t)$ should be the intersection of the epigraphs of all $g_x(t)$. Since the intersection of convex sets is always convex, the epigraph (notice that epigraph excludes all points where the function value is $+\infty$) of $f^*(t)$ is convex, i.e., $f^*(t)$ is convex on $\mathbb{R}$. Then it is easy to see the effective domain of $f^*(t)$ must be a convex set in $\mathbb{R}$. However, the convex set in $\mathbb{R}$ can only be empty set, singleton, or interval.

Now we only need to show all three cases exist for some specific example. If $f(x) = -x^2$ with $I = \mathbb{R}$, then $f^*(t) = \sup_{x\in\mathbb{R}}(tx + x^2) = +\infty$ for all $t \in \mathbb{R}$, so in this case $I^* = \varnothing$. If $f(x) = x$ with $I = \mathbb{R}$, then

$$f^*(t) = \sup_{x\in\mathbb{R}}(tx - x) = \begin{cases} 0 & \text{if } t = 1 \\ +\infty & \text{if } \neq 1 \end{cases}$$

which shows that in this case $I^* = \{1\}$ is a single point. If $f(x) = x^2$ with $I = \mathbb{R}$, then $f^*(t) = \sup_{x\in\mathbb{R}}(tx - x^2) = \frac{t^2}{4}$. In this case, $I^* = \mathbb{R}$. Therefore, all three cases are possible. In particular, if $f^*(t)$ is finite over an interval, since the epigraph of $f^*(t)$ is convex on $\mathbb{R}$, it is convex on I^*, so $f^*(t)$ is a convex function restricted on I^*.

b) If f is a convex function, then $I^* \neq \varnothing$, and for $f^* \in \mathcal{C}(I^*)$

$$(f^*)^* = \sup_{t \in I^*} \big(xt - f^*(t)\big) = f(x)$$

for any $x \in I$. Thus the Legendre transform of a convex function is *involutive*, (its square is the identity transform).

If we do not assume f is a closed function on I, then this statement is wrong. Consider the following function

$$f(x) = \begin{cases} x^2 & x \in [-1, 1) \\ 2 & x = 1 \end{cases}$$

where $I = [-1, 1]$ and $f(x)$ is obviously convex but not closed. By simple calculation, we can obtain $I^* = \mathbb{R}$, and

$$f^*(t) = \begin{cases} t - 1 & t \geq 2 \\ t^2/4 & t \in (-2, 2) \\ -t - 1 & t \leq -2 \end{cases}$$

which is a continuous function on I^*. However, if we consider $x = 1$, then

$$f^{**}(1) = \sup_{t \in \mathbb{R}} (t - f^*(t)) = 1 \neq 2 = f(1)$$

Therefore, we need to add an assumption that f is a closed function on I. Under such a condition, we can prove the desired result. For any fixed $x \in I$, $f(x) \geq xt - f^*(t)$ for all $t \in I^*$, which is true by definition of $f^*(t)$. Therefore, by taking supremum on both sides, it is trivial that $f(x) \geq f^{**}(x)$.

Now we need to prove $f(x) \leq f^{**}(x)$ for all $x \in I$. Suppose not, if we denote the epigraph of f as $\text{epi}(f)$, combined with the fact that $f(x) \geq f^{**}(x)$, $\text{epi}(f) \subsetneq \text{epi}(f^{**})$. Note that $(x_0, t_0) \in \text{epi}(f^{**})$ but $(x_0, t_0) \notin \text{epi}(f)$. Since f is convex and closed, $\text{epi}(f)$ is a convex closed set, so by separating hyperplane theorem, there exists $(a, b) \neq (0, 0)$ and a constant c such that

$$ax + bt < c < ax_0 + bf^{**}(x_0), \quad \forall (x, t) \in \text{epi}(f)$$

where $b \leq 0$ because if $b > 0$, as $t \to \infty$, the LHS is positive infinity but the RHS is a finite value, which is a contradiction.

If $b = 0$, then $ax < c < ax_0$ for all $(x, t) \in \text{epi}(f)$. In this case, choose a $\hat{y} \in I^*$, and we want to find a small enough $\epsilon > 0$ such that

$$(a + \hat{y}\epsilon)x - \epsilon t < c < (a + \hat{y}\epsilon)x_0 - \epsilon f^{**}(x_0)$$

Since $f^*(\hat{y}) \geq x\hat{y} - f(x) \geq x\hat{y} - t$ for all $(x, t) \in \text{epi}(f)$, it suffices to find ϵ such that

$$ax < c < ax_0 - \epsilon(f^{**}(x_0) - \hat{y}x_0 + f^*(\hat{y}))$$

This $\epsilon > 0$ exists because $f^{**}(x_0) - \hat{y}x_0 + f^*(\hat{y})$ is a finite constant, which means as long as ϵ is small enough, the above inequality will hold. Therefore, we find a new pair of $a' = a + \hat{y}\epsilon$ and $b' = -\epsilon < 0$ such that the resulting hyperplane strictly separates the point (x_0, t_0) and $\text{epi}(f)$. Then by normalization, we can always take $b = -1$ and find a corresponding a and c such that

$$ax - t < c < ax_0 - f^{**}(x_0), \quad \forall (x, t) \in \text{epi}(f)$$

Since $(x, f(x)) \in \text{epi}(f)$, we can obtain $ax - f(x) < c < ax_0 - f^{**}(x_0)$ for all $x \in I$. Take the supremum on x over I on both sides, we have $f^*(a) \leq c < ax_0 - f^{**}(x_0)$. This contradicts to the fact that $f^*(t) \geq xt - f^{**}(x)$ for all $t \in I^*$ and $x \in I$. Therefore, $f(x) \leq f^{**}(x)$ for all $x \in I$, and we are done.

c) The following inequality holds:

$$xt \leq f(x) + f^*(t) \text{ for } x \in I \text{ and } t \in I^*$$

We only consider the case when I^* is nonempty. In this case, since for all $x \in I$, we have

$$f^*(t) = \sup_{x \in I} \left(tx - f(x)\right) \geq tx - f(x)$$

It is easy to see $xt \leq f(x) + f^*(t)$ for $x \in I$ and $t \in I^*$.

d) When f is a convex differentiable function, $f^*(t) = tx_t - f(x_t)$, where x_t is determined from the equation $t = f'(x)$. Use this relation to obtain a geometric interpretation of the Legendre transform f^* and its argument t, showing that the Legendre transform is a function defined on the set of tangents to the graph of f.

Since f is convex differentiable, $h_t(x) = tx - f(x)$ on I is a concave differentiable function. It is obvious that if $h_t(x)$ has a stationary point x_t in the interior of I, then this point must be a global maximizer of $h_t(x)$ over interval I. In this case $f^*(t) = tx_t - f(x_t)$ where $t = f'(x_t)$. This shows Legendre transform is a function defined on the set of tangents to the graph of f.

However, if for some $t \in I^*$, $t = f'(x)$ has no solution for $x \in I$ (e.g., consider $f(x) = x^2$ for $x \in I$, where $I = [-1, 1]$ and $I^* = \mathbb{R}$), then we have several cases to consider. Notice that we only assume $f(x)$ to be differentiable on the interior of I no matter I is open, closed or half-open half-closed. This means $f(x)$ may be discontinuous at boundary of I. Also, it is reasonable to assume I has nonempty interior. Denote I_c^d as interval with end ponts c, d ($c < d$ can be finite or infinite), then

- When $|f(c+)| = |f(d-)| = \infty$ (we assume $f(-\infty) = f(-\infty+)$ and $f(\infty) = f(\infty-)$), first note that it is impossible that $f(c+) = f(d-) = -\infty$, because $f(x)$ is convex, so any point $x \in (c, d)$ satisfies $f(x) \leq -\infty$, then $f(x)$ is not a real-valued function on I, which is a contradiction.

  Second, if $f(c+) = f(d-) = \infty$, then we need to consider whether the end points c, d are fintie or not. If $c > -\infty$ and $d < \infty$, then $f'(x) \to \infty$ as $x \to d-$ and $f'(x) \to -\infty$ as $x \to c+$. Imagine if $f'(x)$ is bounded above, fixed $x_0 \in (c, d)$, for $z > x_0$, by MVT, $f(z) - f(x_0) = f'(\xi)(z - x_0)$. As $z \to d-$, $f(z) - f(x_0) \to \infty$, $z - x_0 \to d - x_0 > 0$, but $d - x_0$ is bounded, so $f'(\xi) \to \infty$, which is a contradiction. Similarly, if $f'(x)$ is bounded below, we can also obtain such contradiction. By intermediate value property of $f'(x)$ (even if $f'(x)$ is not continuous, this property holds), $f'(x)$ can attain any value in $\mathbb{R}$ when $x \in (c, d)$, so $t = f'(x)$ always has a solution. If one and only one of c, d is infinite, WLOG, assume $c = -\infty$ and $d < \infty$. In this case, $f'(x) \to \infty$ as $x \to d-$, so when t is large enough, $t = f'(x)$ always has a solution. It has no solution only when $f'(x)$ is bounded below. However, $f'(x)$ is decreasing as $x \to -\infty$, so $f'(x)$ converges to some finite number r as $x \to -\infty$. If $t < r$, then $tx - f(x)$ tends to ∞ as $x \to -\infty$, so such t makes $f^*(t) = \infty$ and $t \notin I^*$; if $t > r$, $t = f'(x)$ has a solution by intermediate value property of $f'(x)$; if $t = r$, then there are two possibilities, either $tx - f(x) \to \infty$ as $x \to -\infty$ or $tx - f(x) \to C$ where C is some constant number. If $tx - f(x) \to \infty$, then such $t \notin I^*$; if $tx - f(x) \to C$, then $t = f'(x)$ has no solution for real-valued x, but since $f^*(t)$ takes supremum over x, $f^*(t) = C = (tx - f(x))\Big|_{x=-\infty}$, so we can regarded it as $t = f'(x)$ has a solution $x_d = -\infty$. The case when $c > -\infty$ and $d = \infty$ is similar to the above one. The last case is when $c = -\infty$ and $d = \infty$. In this case, to make $t = f'(x)$ has no solution, it is possible that $f'(x)$ are bounded either above, below or both. The argument is quite similar to the $c = -\infty$ and $d < \infty$ case, but rather compliacted, so we omit it here.

Finally, if one of $f(c+)$ and $f(d-)$ is positive infinity and the other is negative infinity, WLOG, assume $f(c+) = \infty$ and $f(d-) = -\infty$. In this case $f(x)$ must be decreasing on (c, d). Also, we can see I can only be of the form (c, ∞). If $c > -\infty$, then $f'(x) \to -\infty$ as $x \to c+$, and $f'(x) \to r \le 0$ as $x \to \infty$ where r is a finite constant. Then if $t < r$, $t = f'(x)$ always has a solution $x \in (c, \infty)$; if $t > r$, $tx - f(x) \to \infty$ as $x \to \infty$, so such $t \notin I^*$; if $t = r$, then either $tx - f(x) \to \infty$ or $tx - f(x) \to C$ as $x \to \infty$. If $tx - f(x) \to \infty$, then $t \notin I^*$; if $tx - f(x) \to C$, then $f^*(t) = C = (tx - f(x))\Big|_{x=\infty}$. The other case is $c = -\infty$, in this situation, $f'(x)$ can converge to some finite number s or $f'(x) \to -\infty$ as $x \to -\infty$. The $-\infty$ case is similar as previous case, so we only consider $f'(x) \to s$ as $x \to -\infty$. If $t < s$, then $tx - f(x) \to \infty$ as $x \to -\infty$, so $t \notin I^*$; if $t = s$, then similarly we have either $tx - f(x) \to \infty$ or $tx - f(x) \to C$. When $t \in I^*$, $f^*(t) = C = (tx - f(x))\Big|_{x=-\infty}$. If $s < t < r$, then $t = f'(x)$ has a solution on $(-\infty, \infty)$; if $t = r$ and $t > r$, we can use the same argument as the case when $c > -\infty$.

- When one and only one of $f(c+)$ and $f(d-)$ is infinite, then WLOG, assume $|f(c+)| = \infty$ and $|f(d-)| < \infty$. If $f(c+) = -\infty$ and $|f(d-)| < \infty$, then $f(x)$ must be strictly increasing on (c, d). Furthermore, I can only be $(-\infty, d)$ or $(-\infty, d]$ with $d < \infty$. Note that $f'(-\infty)$ converges to a finite number $r \ge 0$ and $f'(d-) \le \infty$. If $t \ge f'(d-)$, no matter $t = f'(x)$ has solution or not, it is always true that $f^*(t) = h_t(d-)$; if $f'(-\infty) < t < f'(d-)$, then obviously $f'(x) = t$ has a solution in $(-\infty, d)$; if $t < f'(-\infty)$, then $tx - f(x) \to \infty$ as $x \to -\infty$, so such $t \notin I^*$; if $t = f'(-\infty)$, then there are two possibilities, one is that $tx - f(x) \to \infty$ as $x \to -\infty$, but this case shows $t \notin I^*$; the other case is that $tx - f(x)$ converges to fintie number C, then $f^*(t) = C = (tx - f(x))\Big|_{x=-\infty}$. In conclusion, if one and only of $f(c+)$ and $f(d-)$ is infinite, $f^*(t)$ is $h_t(x)$ evaluated at boundary or the limit as x approaches to boundary.
- If $f(c+)$ and $f(d-)$ are both finite, then when $c = -\infty$ and $d = \infty$, $f(x)$ is a constant function, so $I^*\{0\}$ and $f^*(0) = -f(x)$. If one and only one of c, d is infinite, say $c = -\infty$ and $d < \infty$, then $f(x)$ is increasing, and $f(x) \to s$ where s is finite as $x \to -\infty$. This also shows $f'(x) \to 0$ as $x \to -\infty$. Thus, if $t < 0$, then $tx - f(x) \to \infty$ as $x \to -\infty$, so $t \notin I^*$; if $t = 0$, then $f^*(t) = -s$; if $0 < t < f'(d-)$, then $t = f'(x)$ has a solution; if $t \ge f'(d-)$, then $f^*(t) = td - f(d-)$ since d is finite. If $c > -\infty$ and $d < \infty$, then if $t \le f'(c+)$, then $f^*(t) = tc - f(c+)$; if $f'(c+) < t < f'(d-)$, then $t = f'(x)$ has a solution in (c, d); if $t \ge f'(d-)$, then $f^*(t) = td - f(d-)$.

e) The Legendre transform of the function $f(x) = \frac{1}{\alpha}x^\alpha$ for $\alpha > 1$ and $x \geq 0$ is the function $f^*(t) = \frac{1}{\beta}t^\beta$, where $t \geq 0$ and $\frac{1}{\alpha} + \frac{1}{\beta} = 1$. Taking account of c), use this fact to obtain Young's inequality, which we already know:

$$xt \leq \frac{1}{\alpha}x^\alpha + \frac{1}{\beta}t^\beta$$

Notice that $I = \mathbb{R}^+ = \{x \,|\, x \geq 0\}$. By part d), since on I, $f(x)$ is always differentiable and convex, we can let $t = f'(x) = x^{\alpha-1}$ and obtain that $x_t = t^{\frac{1}{\alpha-1}}$. Also note that $t \geq 0$, so $f^*(t) = (1 - \frac{1}{\alpha})t^{\frac{\alpha}{\alpha-1}}$. By part c), we have

$$xt \leq \frac{1}{\alpha}x^\alpha + \left(1 - \frac{1}{\alpha}\right)t^{\frac{\alpha}{\alpha-1}} = \frac{1}{\alpha}x^\alpha + \frac{1}{\beta}t^\beta$$

if we let $\beta = \frac{\alpha}{\alpha-1}$, i.e., $\frac{1}{\alpha} + \frac{1}{\beta} = 1$.

f) The Legendre transform of the function $f(x) = e^x$ is the function $f^*(t) = t\ln\frac{t}{e}$, $t > 0$, and the inequality

$$xt \leq e^x + t\ln\frac{t}{e}$$

holds for $x \in \mathbb{R}$ and $t > 0$.

Notice that here $I = \mathbb{R}$. By part d), let $t = f'(x) = e^x$, we can solve $x_t = \ln t$ where $t > 0$. Therefore, $f^*(t) = t\ln t - t$. By part c), we conclude immediately that $xt \leq e^x + t\ln(t/e)$ for all $x \in \mathbb{R}$ and $t > 0$.

Question 5.5-1. Using the geometric interpretation of complex numbers

a) explain the inequalities $|z_1 + z_2| \leq |z_1| + |z_2|$ and $|z_1 + \cdots + z_n| \leq |z_1| + \cdots + |z_n|$;

The first inequality $|z_1 + z_2| \leq |z_1| + |z_2|$ is just a special case of the second one. There are several explanation of them. One is that between two points the shortest distance is assumed by line segment between these two points. Since the addition and norm of complex number is similar to vector addition and norm, if we treat each complex number as a vector, then we can translate all vectors such that the head of each vector is connected by the tail of another vector. Thus, $|z_1 + \cdots + z_n|$ means the length of line segment between the tail of the first vector and the head of the last vector. $|z_1| + \cdots + |z_n|$ means the length of path from the tail of the first vector to the head of the last vector along vector $z_1, \ldots, z_n$.

You can also interpret it in this way: the total length of any $n-1$ edges in a n-polygon is larger than the length of the remaining edge. They are equal if and only if all edges lie in the same line.

b) exhibit the locus of points in the plane $\mathbb{C}$ satisfying the relation $|z-1| + |z+1| \leq 3$;

All points satisfying the relation above will lie in or on the ellipse

$$\frac{4x^2}{9} + \frac{4y^2}{5} = 1$$

where $(1,0)$ and $(-1,0)$ are two foci of this ellipse. We first notice that the total distance between z and fixed points $(-1,0)$ and $(1,0)$ is less than or equal to 3. If it is equal to three, then it lie on the ellipse with foci $(1,0)$ and $(-1,0)$; also, we know $2a=3$, so $a=3/2$, and $b=\sqrt{5}/2$. By the property of ellipse, any points with total distance between two foci that is less than $2a$ will lie in the interior of ellipse.

c) describe all the nth roots of unity and find their sum;

All the nth roots of unity will satisfy $z_k^n=1$, which is given by

$$z_k = \exp\left(\frac{2k\pi i}{n}\right), \quad k=0,1,\dots,n-1$$

Their sum is given by

$$\sum_{k=0}^{n-1} z_k = \sum_{k=1}^{n-1} \exp\left(\frac{2k\pi i}{n}\right) = \frac{1(1-z_1^n)}{1-z_1} = 0$$

d) explain the action of the transformation of the plane $\mathbb{C}$ defined by the formula $z \mapsto \bar{z}$.

This transformation just maps z into its symmetric point in plane $\mathbb{C}$ with respect to x axis. This is because $\bar{z}=x-yi$ while $z=x+yi$.

Question 5.5-2. Find the following sums:

a) $1+q+\cdots+q^n$;

This is trivial, because when $q \neq 1$,

$$1+q+\cdots+q^n = \frac{1-q^{n+1}}{1-q}$$

when $q=1$, the summation is just $n+1$.

b) $1+q+\cdots+q^n+\cdots$ for $|q|<1$;

Since $|q|<1$, we have

$$1+q+\cdots+q^n+\cdots = \lim_{n\to\infty}\frac{1-q^{n+1}}{1-q} = \frac{1}{1-q}$$

This is because for any complex number $|q|<1$, we have

$$\lim_{n\to\infty}|q^n| = \lim_{n\to\infty}|r^n||e^{in\theta}| = 0, \quad \text{and } \lim_{n\to\infty}|q^n|=0 \Longrightarrow \lim_{n\to\infty}q^n=0$$

c) $1+e^{i\varphi}+\cdots+e^{in\varphi}$;

Let $q=e^{i\varphi}$, from part a), when $e^{i\varphi}\neq 1$, i.e. $\varphi \neq 2k\pi$, $k\in\mathbb{Z}$, we have

$$1+e^{i\varphi}+\cdots+e^{in\varphi} = \frac{1-e^{i(n+1)\varphi}}{1-e^{i\varphi}}$$

when $\varphi=2k\pi$, $k\in\mathbb{Z}$, the summation is just $n+1$.

d) $1 + re^{i\varphi} + \cdots + r^n e^{in\varphi}$;

Let $q = re^{i\varphi}$, from part a), when $re^{i\varphi} \neq 1$, we have

$$1 + re^{i\varphi} + \cdots + re^{in\varphi} = \frac{1 - r^{n+1}e^{i(n+1)\varphi}}{1 - re^{i\varphi}}$$

when $re^{i\varphi} = 1$, the summation is just $n + 1$.

e) $1 + re^{i\varphi} + \cdots + r^n e^{in\varphi} + \cdots$ for $|r| < 1$;

Since $|r| < 1$, we have

$$1 + re^{i\varphi} + \cdots + r^n e^{in\varphi} + \cdots = \lim_{n\to\infty} \frac{1 - r^{n+1}e^{i(n+1)\varphi}}{1 - re^{i\varphi}} = \frac{1}{1 - re^{i\varphi}}$$

f) $1 + r\cos\varphi + \cdots + r^n \cos n\varphi$;

Recall part d), we have

$$1 + r\cos\varphi + \cdots + r^n \cos n\varphi = \text{Re}\left\{\frac{1 - r^{n+1}e^{i(n+1)\varphi}}{1 - re^{i\varphi}}\right\}$$

which is

$$1 + r\cos\varphi + \cdots + r^n \cos n\varphi = \frac{1 - r\cos\varphi - r^{n+1}\cos(n+1)\varphi + r^{n+2}\cos n\varphi}{1 - 2r\cos\varphi + r^2}$$

If $r = 1$, $\varphi = 2k\pi$ or if $r = -1$, $\varphi = (2k+1)\pi$, where $k \in \mathbb{Z}$, we have

$$1 + r\cos\varphi + \cdots + r^n \cos n\varphi = n + 1$$

g) $1 + r\cos\varphi + \cdots + r^n \cos n\varphi + \cdots$ for $|r| < 1$;

Since $|r| < 1$, cosine function is bounded by 1, we have

$$\begin{aligned} 1 + r\cos\varphi + \cdots + r^n \cos n\varphi + \cdots &= \lim_{n\to\infty} \frac{1 - r\cos\varphi - r^{n+1}\cos(n+1)\varphi + r^{n+2}\cos n\varphi}{1 - 2r\cos\varphi + r^2} \\ &= \frac{1 - r\cos\varphi}{1 - 2r\cos\varphi + r^2} \end{aligned}$$

h) $1 + r\sin\varphi + \cdots + r^n \sin n\varphi$;

Recall part d), we have

$$1 + r\sin\varphi + \cdots + r^n \sin n\varphi = 1 + \text{Im}\left\{\frac{1 - r^{n+1}e^{i(n+1)\varphi}}{1 - re^{i\varphi}}\right\}$$

which is

$$1 + r\sin\varphi + \cdots + r^n \sin n\varphi = 1 + \frac{r\sin\varphi - r^{n+1}\sin(n+1)\varphi + r^{n+2}\sin n\varphi}{1 - 2r\cos\varphi + r^2}$$

If $r = 1$, $\varphi = 2k\pi$ or if $r = -1$, $\varphi = (2k+1)\pi$, where $k \in \mathbb{Z}$, we have

$$1 + r\sin\varphi + \cdots + r^n \sin n\varphi = 1$$

i) $1 + r\sin\varphi + \cdots + r^n \sin n\varphi + \cdots$ for $|r| < 1$.

Since $|r| < 1$, sine function is bounded by 1, we have

$$\begin{aligned} 1 + r\sin\varphi + \cdots + r^n \sin n\varphi + \cdots &= \lim_{n\to\infty}\left(1 + \frac{r\sin\varphi - r^{n+1}\sin(n+1)\varphi + r^{n+2}\sin n\varphi}{1 - 2r\cos\varphi + r^2}\right) \\ &= 1 + \frac{r\sin\varphi}{1 - 2r\cos\varphi + r^2} \end{aligned}$$

Question 5.5-3. Find the modulus and argument of the complex number $\lim_{n\to\infty}\left(1 + \frac{z}{n}\right)^n$ and verify that this number is e^z.

The modulus (which is a continuous function of z) of it is found by

$$\begin{aligned} \lim_{n\to\infty}\left|\left(1 + \frac{x}{n} + \frac{y}{n}i\right)^n\right| &= \lim_{n\to\infty}\left|1 + \frac{x}{n} + \frac{y}{n}i\right|^n \\ &= \lim_{n\to\infty}\left(1 + \frac{2x}{n} + \frac{x^2+y^2}{n^2}\right)^{n/2} \\ &= \exp\left\{\lim_{n\to\infty}\frac{n}{2}\ln\left(1 + \frac{2x}{n} + \frac{x^2+y^2}{n^2}\right)\right\} \\ &= \exp\left\{\lim_{n\to\infty}\frac{n}{2}\left(\frac{2x}{n} + O\left(\frac{1}{n^2}\right)\right)\right\} \\ &= e^x \end{aligned}$$

The argument (which is also a continuous function of z) of it is found by

$$\begin{aligned} \lim_{n\to\infty}\arg\left(1 + \frac{x}{n} + \frac{y}{n}i\right)^n &= \lim_{n\to\infty} n\arg\left(1 + \frac{x}{n} + \frac{y}{n}i\right) \\ &= \lim_{n\to\infty} n\arctan\left(\frac{y}{n+x}\right) \\ &= \lim_{n\to\infty} n\left(\frac{y}{n+x} + O\left(\frac{1}{n^3}\right)\right) \\ &= y \end{aligned}$$

Hence, we have

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

Question 5.5-4.

a) Show that the equation $e^w = z$ in w has the solution $w = \ln|z| + i\text{Arg } z$. It is natural to regard w as the *natural logarithm* of z. Thus $w = \text{Ln } z$ is not a functional relation, since Arg z is multi-valued.

Let $w = x + yi$ and $z = a + bi$, then we have

$$e^{x+yi} = e^x(\cos y + i \sin y) = e^x \cos y + ie^x \sin y = a + bi$$

Thus, we have $e^x \cos y = a$ and $e^x \sin y = b$. Solve x, y in terms of a, b, we have

$$e^x = \sqrt{x^2 + y^2} = |z|, \qquad \tan y = \frac{b}{a}$$

Hence, $x = \ln |z|$, and $y = \text{Arg } z$, so $w = \ln |z| + i\text{Arg } z$. Here Arg z is multi-valued.

b) Find Ln 1 and Ln i.

By part a), we have

$$\text{Ln } 1 = \ln |1| + i\text{Arg } 1 = 0 + 2k\pi i = 2k\pi i, \quad k \in \mathbb{Z}$$
$$\text{Ln } i = \ln |i| + i\text{Arg } i = 0 + \left(\frac{\pi}{2} + 2k\pi\right) i = \frac{4k+1}{2}\pi i, \quad k \in \mathbb{Z}$$

c) Set $z^\alpha = e^{\alpha \text{Ln } z}$. Find 1^π and i^i.

According to the formula and what we calculate in part b), we have

$$1^\pi = e^{\pi \text{Ln } 1} = e^{2k\pi^2 i}, \quad k \in \mathbb{Z}$$

which means 1^π is multi-valued (actually since π is irrational, it has infinitely many value; recall what we proved in Question 3.1-3, we can further show that all of its value are dense on unit circle).

$$i^i = e^{i\text{Ln } i} = e^{-\frac{4k+1}{2}\pi}, \quad k \in \mathbb{Z}$$

Here i^i also has infinitely many value, but they are not dense on real line and have one accumulation point 0.

d) Using the representation $w = \sin z = \frac{1}{2i}\left(e^{iz} - e^{-iz}\right)$, obtain an expression for $z = \arcsin w$.

Since $w = \frac{1}{2i}\left(e^{iz} - e^{-iz}\right)$, we can solve e^{iz} in terms of w,

$$e^{iz} = iw + (1 - w^2)^{1/2}$$

Hence, we have

$$z = \arcsin w = -i \,\text{Ln}\, \left[iw + (1 - w^2)^{1/2}\right]$$

e) Are there points in $\mathbb{C}$ where $|\sin z| = 2$?

We can just find some points z such that $\sin z = 2$, then its modulus must be 2. Use formula in part d), we have

$$z = \arcsin 2 = -i \,\text{Ln}\, \left[2i + (-3)^{1/2}\right]$$

Since $(-3)^{1/2} = \pm\sqrt{3}i$, we have

$$z = -i \,\text{Ln}\, \left[\left(2 \pm \sqrt{3}\right) i\right] = \frac{\pi}{2} + 2k\pi - i \ln\left(2 \pm \sqrt{3}\right), \quad k \in \mathbb{Z}$$

Hence there exist points in $\mathbb{C}$ such that $|\sin z| = 2$.

Question 5.5-5.

a) Investigate whether the function $f(z) = \frac{1}{1+z^2}$ is continuous at all points of the plane $\mathbb{C}$.

Since the function $\frac{1}{w}$ has only one discontinuous point $w = 0$, $f(z)$ can be discontinuous only at $w = 1 + z^2 = 0$, which yields $z = \pm i$. Thus, $f(z)$ is continuous at every point in $\mathbb{C}\backslash\{\pm i\}$.

b) Expand the function $f(z) = \frac{1}{1+z^2}$ in a power series around $z_0 = 0$ and find its radius of convergence.

The Taylor expansion at $z_0 = 0$ is

$$\frac{1}{1+z^2} = \frac{1}{1-(-z^2)} = 1 - z^2 + z^4 - z^6 + \cdots = \sum_{i=0}^{\infty}(-1)^i z^{2i}$$

Apply root test, the convergence radius is

$$R = \frac{1}{\overline{\lim_{n\to\infty}} |c_n|^{1/n}} = 1, \qquad \text{where } c_n \text{ is } 1, 0, -1, 0, 1, 0, -1, \ldots$$

Hence, the radius of convergence of $f(z) = \frac{1}{1+z^2}$ is 1, it converges when $|z| < 1$.

c) Solve parts a) and b) for the function $\frac{1}{1+\lambda^2 z^2}$, where $\lambda \in \mathbb{R}$ is a parameter.

Similar to part a), since the function $\frac{1}{w}$ has only one discontinuous point $w = 0$, $f(z)$ can be discontinuous only at $w = 1 + \lambda^2 z^2 = 0$, which yields $z = \pm\lambda^{-1} i$. Thus, $f(z)$ is continuous at every point in $\mathbb{C}\backslash\{\pm\lambda^{-1} i\}$. If $\lambda = 0$, then $f(z)$ is continuous everywhere in $\mathbb{C}$.

Similar to part b), the Taylor expansion at $z_0 = 0$ is

$$\frac{1}{1+\lambda^2 z^2} = \frac{1}{1-(-\lambda^2 z^2)} = 1 - \lambda^2 z^2 + \lambda^4 z^4 - \lambda^6 z^6 + \cdots = \sum_{i=0}^{\infty}(-1)^i (\lambda z)^{2i}$$

Apply root test, the convergence radius (if $\lambda \neq 0$) is

$$R = \frac{1}{\overline{\lim_{n\to\infty}} |c_n|^{1/n}} = \frac{1}{|\lambda|}, \qquad \text{where } c_n \text{ is } 1, 0, -\lambda^2, 0, \lambda^4, 0, -\lambda^6, 0, \ldots$$

Hence, the radius of convergence of $f(z) = \frac{1}{1+\lambda z^2}$ is $|\lambda|^{-1}$, it converges when $|z| < \frac{1}{|\lambda|}$. If $\lambda = 0$, then the radius of convergence of $f(z)$ is $+\infty$.

Can you make a conjecture as to how the radius of convergence is determined by the relative location of certain points in the plane $\mathbb{C}$? Could this relation have been understood on the basis of the real line alone, that is, by expanding the function $\frac{1}{1+\lambda^2 x^2}$, where $\lambda \in \mathbb{R}$ and $x \in \mathbb{R}$?

The radius of convergence is equal to the distance between point you expand the function and the closest singular point (singularity). If we restrict it in real line, this may not be true, because $\frac{1}{1+x^2}$ is differentiable in $\mathbb{R}$ and has no singularity, but it still has a radius of convergence, that is, 1. Thus, if we want to use the distance to the closest singularity to determine the radius of convergence, we need to consider the singularity of the function in $\mathbb{C}$.

Question 5.5-6.

a) Investigate whether the Cauchy function

$$f(z) = \begin{cases} e^{-1/z^2} & z \neq 0 \\ 0 & z = 0 \end{cases}$$

is continuous at $z = 0$.

If $z = x + iy$, and we let $y = 0$, then as $z \to 0$, we have

$$\lim_{z\to 0} e^{-1/z^2} = \lim_{x\to 0} e^{-1/x^2} = 0 = f(0)$$

But if we let $x = 0$, and as $z \to 0$, we have

$$\lim_{z\to 0} e^{-1/z^2} = \lim_{y\to 0} e^{1/y^2} = +\infty \neq f(0)$$

Hence $f(z)$ is not continuous at $z = 0$.

b) Is the restriction $f\Big|_{\mathbb{R}}$ of the function f in a) to the real line continuous?

It is continuous, because as we proved in a), the function

$$f(x) = \begin{cases} e^{-1/x^2} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

has limit

$$\lim_{x\to 0} e^{-1/x^2} = 0 = f(0)$$

So it is continuous at $x = 0$, and the continuity of other points is easy to see.

c) Does the Taylor series of the function f in a) exist at the point $z_0 = 0$?

No, because $z_0 = 0$ is essential singularity point, or you can say it is not continuous, so not differentiable. Then its Taylor series of course does not exist.

d) Are there functions analytic at a point $z_0 \in \mathbb{C}$ whose Taylor series converge only at the point z_0?

If a function is analytic at a point z_0, then it must be analytic in a neighborhood of z_0, but analytic in a neighborhood means its Taylor series converges in this neighborhood. Thus, the Taylor series of such function is impossible to converge at only one point.

e) Invent a power series $\sum\limits_{n=0}^{\infty} c_n(z - z_0)^n$ that converges only at the one point z_0.

Consider the power series

$$\sum_{n=0}^{\infty} n!(z - z_0)^n$$

Check its radius of convergence by root test,

$$R = \frac{1}{\overline{\lim\limits_{n\to\infty}} |c_n|^{1/n}} = \frac{1}{\overline{\lim\limits_{n\to\infty}} (n!)^{1/n}}$$

To compute the limit above, we consider

$$\begin{aligned}\lim_{n\to\infty}(n!)^{1/n} &= \lim_{n\to\infty}\exp\left\{\frac{1}{n}\sum_{k=1}^{n}\ln k\right\}\\ &= \lim_{n\to\infty}\exp\left\{\ln n+\frac{1}{n}\sum_{k=1}^{n}\ln\frac{k}{n}\right\}\\ &= \lim_{n\to\infty} n\exp\left\{\frac{1}{n}\sum_{k=1}^{n}\ln\frac{k}{n}\right\}\\ &= \lim_{n\to\infty} n\exp\left\{\int_0^1 \ln x\,dx\right\}\\ &= \lim_{n\to\infty} ne^{-1}\to+\infty\end{aligned}$$

Hence,

$$R=\frac{1}{\overline{\lim_{n\to\infty}}(n!)^{1/n}}=0$$

But when $z=z_0$, the series is constant zero, so it is convergent. Thus, z_0 is the only point where this series converges, because if another convergent point exists, convergence radius should be no less than the distance between these two points.

Question 5.5-7.

a) Making the formal substitution $z-a=(z-z_0)+(z_0-a)$ in the power series $\sum_{n=0}^{\infty}A_n(z-a)^n$ and gathering like terms, obtain a series $\sum_{n=0}^{\infty}C_n(z-z_0)^n$ and expressions for its coefficients in terms of A_k and $(z_0-a)^k$, $k=0,1,\ldots$.

Substitute $z-a$ with $(z-z_0)+(z_0-a)$ and gather like terms, we obtain a new series (not necessarily equal to the original one, and not even converges; this is because rearrange an infinite series may change the convergence of the original one). Using binomial expansion, the result is as follows

$$\sum_{n=0}^{\infty}C_n(z-z_0)^n=\sum_{n=0}^{\infty}\left[\sum_{k=n}^{\infty}A_k\binom{k}{n}(z_0-a)^{k-n}\right](z-z_0)^n$$

where

$$C_n=\sum_{k=n}^{\infty}A_k\binom{k}{n}(z_0-a)^{k-n}$$

b) Verify that if the original series converges in the disk $|z-a|<R$ and $|z_0-a|=r<R$, then the series defining C_n, $n=0,1,\ldots$, converge absolutely and the series $\sum_{n=0}^{\infty}C_n(z-z_0)^n$ converges for $|z-z_0|<R-r$.

Since the radius of convergence of the original series is R, we have

$$\frac{1}{\overline{\lim_{n\to\infty}}|A_n|^{1/n}}=R$$

If we can prove

$$\frac{1}{\overline{\lim}_{k\to\infty} \left|A_k\binom{k}{n}\right|^{1/k}} = R$$

Then the series

$$\sum_{k=n}^{\infty} A_k \binom{k}{n} (z-a)^{k-n}$$

will be convergent for $|z-a| < R$, but since $|z_0 - a| = r < R$, C_n will be convergent.

Notice that

$$\lim_{k\to\infty} \binom{k}{n}^{1/k} = \lim_{k\to\infty} \left[\frac{k(k-1)\cdots(k-n+1)}{n!}\right]^{1/k}$$

Since for any fixed number a, the limit of $(k-a)^{1/k}$ is 1, and there are only finitely many terms in numerator, thus

$$\lim_{k\to\infty} \binom{k}{n}^{1/k} = \lim_{k\to\infty} \frac{k^{1/k}(k-1)^{1/k}\cdots(k-n+1)^{1/k}}{(n!)^{1/k}} = 1$$

Also, you could verify (by definition) that

$$\overline{\lim}_{k\to\infty} \left|A_k\binom{k}{n}\right|^{1/k} = \lim_{k\to\infty} \binom{k}{n}^{1/k} \overline{\lim}_{k\to\infty} |A_k|^{1/k}$$

Hence, we proved that

$$\frac{1}{\overline{\lim}_{k\to\infty} \left|A_k\binom{k}{n}\right|^{1/k}} = R$$

and the convergence (absolutely, because the original series converges absolutely within radius R) of C_n follows immediately.

To prove the second part, we denote

$$\sum_{n=0}^{\infty}\left[\sum_{k=n}^{\infty} A_k \binom{k}{n}(z_0-a)^{k-n}\right](z-z_0)^n = \sum_{n=0}^{\infty}\sum_{k=n}^{\infty} B_{kn}$$

$$\sum_{k=0}^{\infty}\sum_{n=0}^{k} A_k \binom{k}{n}(z_0-a)^{k-n}(z-z_0)^n = \sum_{k=0}^{\infty}\sum_{n=0}^{k} B'_{kn}$$

However, we have

$$\sum_{k=0}^{\infty}\sum_{n=0}^{k} B'_{kn} = \sum_{k=0}^{\infty} A_k[(z-z_0)+(z_0-a)]^n$$

The right hand side converges absolutely if $|z-z_0|+|z_0-a| < R$, i.e., $|z-z_0| < R-r$. Hence, if $|z-z_0| < R-r$, we can interchange the order of the double summation of $\sum_{k=0}^{\infty}\sum_{n=0}^{k} B'_{kn}$, which is (Theorem 8.3 in Rudin's book)

$$\sum_{k=0}^{\infty}\sum_{n=0}^{k} A_k \binom{k}{n}(z_0-a)^{k-n}(z-z_0)^n = \sum_{n=0}^{\infty}\left[\sum_{k=n}^{\infty} A_k \binom{k}{n}(z_0-a)^{k-n}\right](z-z_0)^n$$

This means if $|z-z_0| < R-r$, the right hand side is equal to left hand side, but left hand side is convergent, so the right hand side is also convergent, and the proof is finished.

c) Show that if $f(z) = \sum\limits_{n=0}^{\infty} A_n(z-a)^n$ in the disk $|z-a| < R$ and $|z_0 - a| < R$, then in the disk $|z - z_0| < R - |z_0 - a|$ the function f admits the representation $f(z) = \sum\limits_{n=0}^{\infty} C_n(z-z_0)^n$.

In part b), we have proved this new series $\sum_{n=0}^{\infty}\sum_{k=n}^{\infty} B_{kn}$ converges and will converge to the same value as $\sum_{k=0}^{\infty}\sum_{n=0}^{k} B'_{kn}$, but $\sum_{k=0}^{\infty}\sum_{n=0}^{k} B'_{kn}$ is just $\sum_{n=0}^{\infty} A_n(z-a)^n$ and $\sum_{n=0}^{\infty}\sum_{k=n}^{\infty} B_{kn}$ is just $\sum_{n=0}^{\infty} C_n(z-z_0)^n$. In the disk $|z - z_0| < R - |z_0 - a|$, we must have

$$f(z) = \sum_{n=0}^{\infty} A_n(z-a)^n = \sum_{n=0}^{\infty} C_n(z-z_0)^n$$

Question 5.5-8. Verify that

a) as the point $z \in \mathbb{C}$ traverses the circle $|z| = r > 1$ the point $w = z + z^{-1}$ traverses an ellipse with center at zero and foci at ± 2;

Since $|z| = r > 1$, we set $z = re^{i\theta}$ with $r > 1$. Then we have

$$w = re^{i\theta} + \frac{1}{r}e^{-i\theta} = \left(r + \frac{1}{r}\right)\cos\theta + \left(r - \frac{1}{r}\right) i\sin\theta$$

Hence point w has locus (u, v) with

$$u = \left(r + \frac{1}{r}\right)\cos\theta, \qquad v = \left(r - \frac{1}{r}\right)\sin\theta$$

Since $r > 1$ is fixed, we have

$$\frac{u^2}{(r+1/r)^2} + \frac{v^2}{(r-1/r)^2} = 1$$

which is an ellipse with center at zero and foci at ± 2, because

$$c^2 = a^2 - b^2 = \left(r + \frac{1}{r}\right)^2 - \left(r - \frac{1}{r}\right)^2 = 4 \Longrightarrow c = 2$$

b) when a complex number is squared (more precisely, under the mapping $w \mapsto w^2$), such an ellipse maps to an ellipse with a focus at 0, traversed twice.

From part a), we have

$$w^2 = z^2 + \frac{1}{z^2} + 2 = \left(r^2 + \frac{1}{r^2}\right)\cos 2\theta + 2 + \left(r^2 - \frac{1}{r^2}\right) i\sin 2\theta$$

Hence point w has locus (u, v) with

$$u = \left(r^2 + \frac{1}{r^2}\right)\cos 2\theta + 2, \qquad v = \left(r^2 - \frac{1}{r^2}\right)\sin 2\theta$$

Since $r > 1$ is fixed, we have

$$\frac{(u-2)^2}{(r^2+1/r^2)^2} + \frac{v^2}{(r^2-1/r^2)^2} = 1$$

This is also an ellipse, but translated along positive direction of x-axis by 2 units. Thus, the original focus $(-2,0)$ moves to $(0,0)$ after the translation. It traversed twice because for z, its argument is $\theta \in [0,2\pi)$, but here w^2 have argument $2\theta \in [0,4\pi)$.

c) under squaring of complex numbers, any ellipse with center at zero maps to an ellipse with a focus at 0.

Let $z = a\cos\theta + ib\sin\theta$, then $w = z^2 = a^2\cos^2\theta - b^2\sin^2\theta + 2abi\sin\theta\cos\theta$. Hence point w has locus (u,v) with

$$u = a^2\cos^2\theta - b^2\sin^2\theta = \left(a^2 - \frac{1}{2}c^2\right)\cos 2\theta + \frac{1}{2}c^2, \qquad v = ab\sin 2\theta$$

Thus we have

$$\frac{(x-(1/2)c^2)^2}{(a^2-(1/2)c^2)^2} + \frac{y^2}{a^2b^2} = 1$$

We can regard this ellipse as a translation of another ellipse E_1 along x-axis by $(1/2)c^2$ in the positive direction, i.e.,

$$\frac{x^2}{(a^2-(1/2)c^2)^2} + \frac{y^2}{a^2b^2} = 1$$

Note that one of the focus of E_1 is given by

$$(c^*)^2 = (a^2-(1/2)c^2)^2 - a^2b^2 = (a^2-(1/2)c^2)^2 - a^2(a^2-c^2) = \frac{1}{4}c^4 \Longrightarrow c^* = \frac{1}{2}c^2$$

Hence, after translation, this focus $(-(1/2)c^2, 0)$ will move to $(0,0)$.