Mathematics 2

Yoshitsugu Yamamoto*

September 2004

1 Maximum and Minimum Values

Definition 1.1 (CRITICAL NUMBERS).
A number x_0 in the domain of f such that either $f'(x_0) = 0$ or $f'(x_0)$ is not defined is called a *critical number* of f.

Proposition 1.2 (SECOND DERIVATIVE TEST FOR RELATIVE EXTREMA).
Assume that $f'(x_0) = 0$ and that $f''(x_0)$ exists. Then:

(i) if $f''(x_0) < 0$, then f has a relative maximum at x_0;

(ii) if $f''(x_0) > 0$, then f has a relative minimum at x_0;

(iii) if $f''(x_0) = 0$, we do not know what is happening at x_0.

Proposition 1.3 (FIRST DERIVATIVE TEST).
Assume $f'(x_0) = 0$.

Case $\{+, -\}$: *If f' is positive in an open interval immediately to the left of x_0, and negative in an open interval immediately to the right of x_0, then f has a relative maximum at x_0.*

Case $\{-, +\}$: *If f' is negative in an open interval immediately to the left of x_0, and positive in an open interval immediately to the right of x_0, then f has a relative minimum at x_0.*

Case $\{+, +\}$ and $\{-, -\}$: *If f' has the same sign in open intervals immediately to the left and to the right of x_0, then f has neither a relative maximum nor a relative minimum at x_0.*

Proposition 1.4 (ABSOLUTE MAXIMUM AND MINIMUM).
An absolute maximum of a function f on a set S occurs at x_0 in S if $f(x) \leq f(x_0)$ for all x in S. An absolute minimum of a function f on a set S occurs at x_0 in S if $f(x) \geq f(x_0)$ for all x in S.

2 Curve Sketching, Concavity, Symmetry

Definition 2.1 (CONCAVITY).
From an intuitive standpoint, an arc of a curve is said to be *concave upward* if it has the shape of a cup and is said to be *concave downward* if it has the shape of a cap. An arc is concave upward if, for each x_0, the arc lies above the tangent line at x_0 in some open interval around x_0. Similarly, an arc is concave downward if, for each x_0, the arc lies below the tangent line at x_0 in some open interval around x_0.

Theorem 2.2.
(a) If $f''(x) > 0$ for x in (a, b), then the graph of f is concave upward for $a < x < b$.
(b) If $f''(x) < 0$ for x in (a, b), then the graph of f is concave downward for $a < x < b$.

*3F1007, ex5001, yamamoto@sk.tsukuba.ac.jp, http://infoshako.sk.tsukuba.ac.jp/~yamamoto/

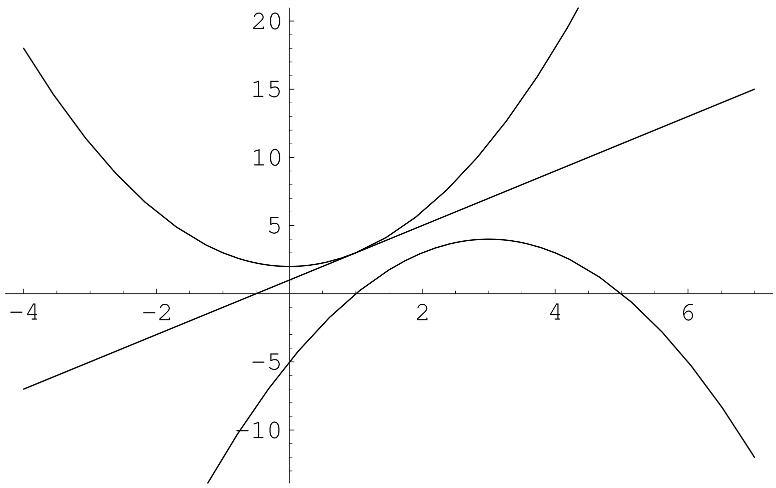

Figure 1: Concave upward and downward

Proof. We show (a). Let x_0 be a point in the interval (a, b). The assumption in (a) implies that $f'(x)$ is increasing on an interval (a, b). Let x be a point in (x_0, b). Then by the law of mean, we have a point $x^* \in (x_0, x)$ such that

$$f(x) - f(x_0) = f'(x^*)(x - x_0).$$

Since f' is increasing we also have $f'(x_0) < f'(x^*)$. Thus we obtain

$$f(x) - f(x_0) = f'(x^*)(x - x_0) > f'(x_0)(x - x_0)$$

or

$$f(x) > f(x_0) + f'(x_0)(x - x_0),$$

the right hand side is the tangent line of f at x_0. We can obtain the same inequality when x is in (a, x_0), where $x - x_0 < 0$ and $f'(x^*) < f'(x_0)$. □

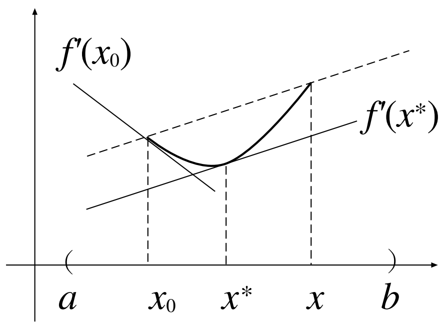

Figure 2: Illustration of proof

Definition 2.3 (POINTS OF INFLECTION).
A point of inflection on a curve $y = f(x)$ is a point at which the concavity changes, that is, the curve is concave upward on one side and concave downward on the other side of the point.

Theorem 2.4.
If the graph of f has an inflection point at x_0 and f'' exists in an open interval containing x_0 and f'' is continuous at x_0, then $f''(x_0) = 0$.

Definition 2.5 (ASYMPTOTES).

1. The vertical line $x = x_0$ is a *vertical asymptote* of f when either $\lim_{x \to x_0+} f(x)$ or $\lim_{x \to x_0-} f(x)$ is $+\infty$ or $-\infty$.
2. The horizontal line $y = y_0$ is a *horizontal asymptote* of f when either $\lim_{x \to +\infty} f(x) = y_0$ or $\lim_{x \to -\infty} f(x) = y_0$.

Remark 2.6 *(Limit).*

1. $\alpha = \lim_{x \to a+} f(x) \quad \rightleftharpoons \quad \forall \epsilon > 0 \; \exists \delta > 0$ such that $0 < x - a < \delta$ implies $|f(x) - \alpha| < \epsilon$
2. $\alpha = \lim_{x \to a-} f(x) \quad \rightleftharpoons \quad \forall \epsilon > 0 \; \exists \delta > 0$ such that $0 < a - x < \delta$ implies $|f(x) - \alpha| < \epsilon$
3. $\alpha = \lim_{x \to a} f(x) \quad \rightleftharpoons \quad \forall \epsilon > 0 \; \exists \delta > 0$ such that $0 < |x - a| < \delta$ implies $|f(x) - \alpha| < \epsilon$

where δ can depend on the given ϵ.

Definition 2.7 (SYMMETRY).

1. The graph $F(x, y) = 0$ is symmetric w.r.t. the y axis if and only if $F(x, y) = 0$ implies $F(-x, y) = 0$.
2. The graph $F(x, y) = 0$ is symmetric w.r.t. the x axis if and only if $F(x, y) = 0$ implies $F(x, -y) = 0$.
3. The graph $F(x, y) = 0$ is symmetric w.r.t. the origin if and only if $F(x, y) = 0$ implies $F(-x, -y) = 0$.

Definition 2.8 (EVEN AND ODD).
A function f is *even* if $f(x) = f(-x)$ for all x, and *odd* if $f(x) = -f(-x)$ for all x.

Example 2.9.
The function $f(x) = \frac{x^2}{1+x^2}$ and its derivatives.

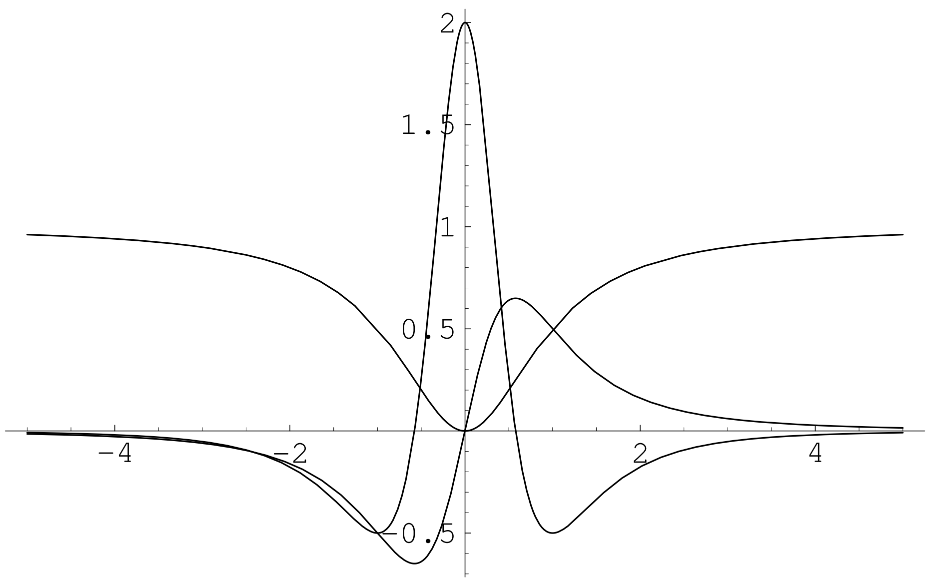

Figure 3: $f(x) = \frac{x^2}{1+x^2}$ and its derivatives

3 Review of Trigonometry

Definition 3.1 (ANGLE MEASURE).
Consider a circle of radius 1 and with center at a point C. Let CA and CB be two radii for which the arc $\overset{\frown}{AB}$ of the circle has length 1. Then one radian is taken to be the measure of the central angle ACB.

Definition 3.2 (SINE AND COSINE FUNCTIONS).
Consider a coordinate system with origin at O and point A at $(1, 0)$. Rotate the arrow OA through an angle of θ radian to a new position OB. Then

1. $\cos\theta$ is defined to be the x coordinate of the point B.
2. $\sin\theta$ is defined to be the y coordinate of the point B.

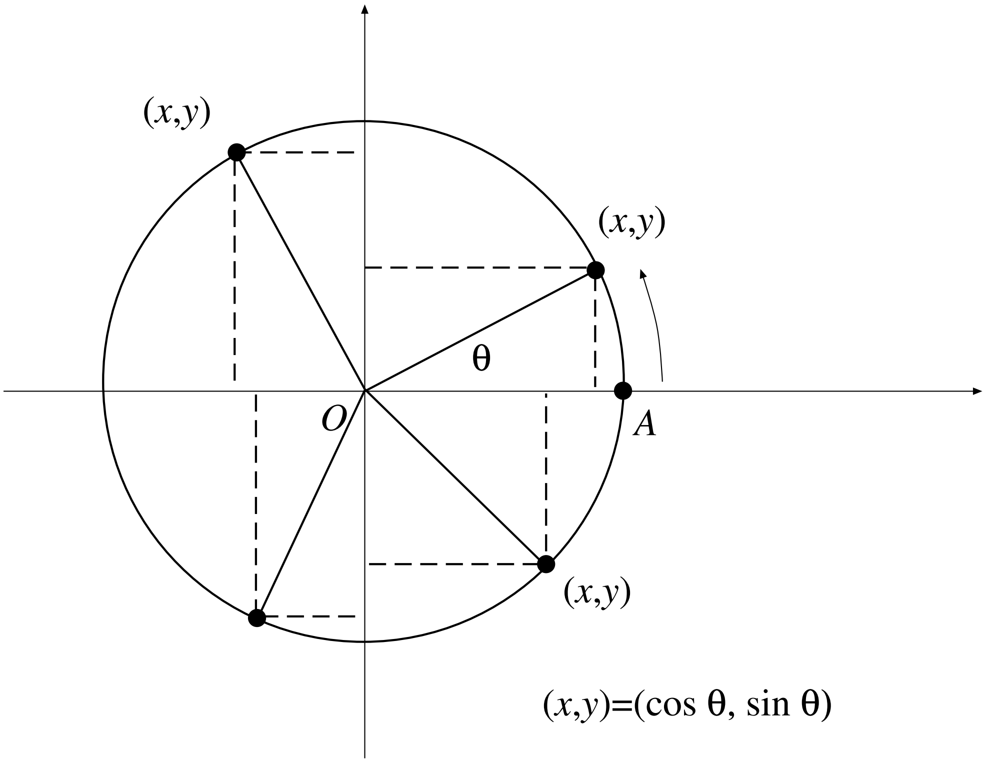

Figure 4: Definition of sine and cosine

Proposition 3.3.

(1) $\cos(\theta + 2\pi) = \cos\theta$ *and* $\sin(\theta + 2\pi) = \sin\theta$

(2) $\cos(-\theta) = \cos(\theta)$ *and* $\sin(-\theta) = -\sin\theta$

(3) $\sin^2\theta + \cos^2\theta = 1$, *where* $sin^2\theta$ *denotes* $(\sin\theta)^2$. *A direct consequence of the Pythagorean theorem.*

(4) For any point $A(x, y)$ *different from the origin* O, *let* r *be its distance from the origin, and let* θ *be the radian measure of the angle from the positive* x *axis to the arrow* OA. *The pair* (r, θ) *are called polar coordinates of* A. *Then* $x = r\cos\theta$ *and* $y = r\sin\theta$.

(5) **(a)** *(Law of Cosines). In any triangle* $\triangle ABC$
$c^2 = a^2 + b^2 - 2ab\cos\theta$

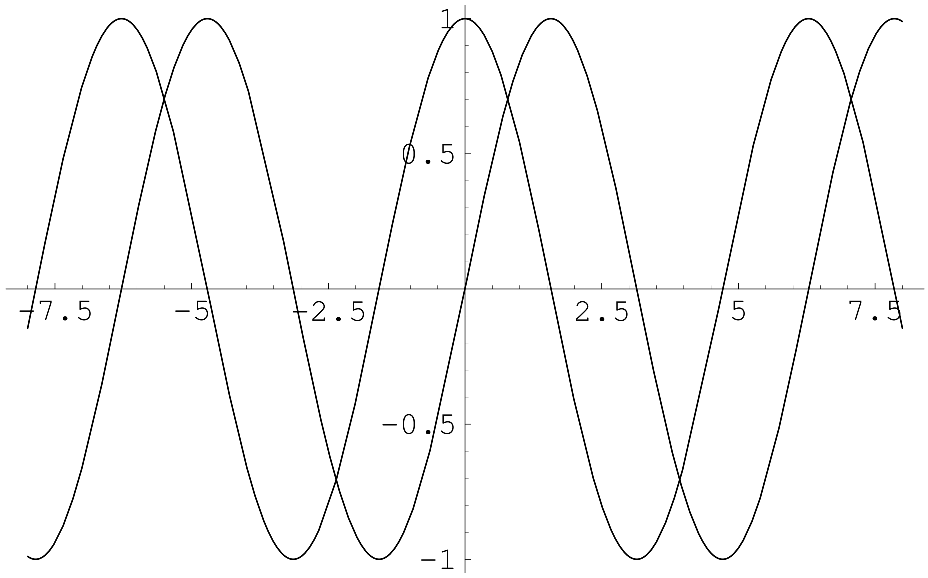

Figure 5: Functions of sine and cosine

Proof. Consider the triangle in Fig. 6, where $x = b\cos\theta$, $y = b\sin\theta$, and $c^2 = (a-x)^2 + y^2$. Then

$$\begin{aligned} c^2 &= (b\cos\theta - a)^2 + (b\sin\theta)^2 \\ &= a^2 + b^2(\cos^2\theta + \sin^2\theta) - 2ab\cos\theta \\ &= a^2 + b^2 - 2ab\cos\theta. \end{aligned}$$

□

(b) *(Law of Sines)*

$$\frac{\sin A}{a} = \frac{\sin B}{b} = \frac{\sin C}{c}$$

where $\sin A$ *is* $\sin(\angle BAC)$,*and similarly for* $\sin B$ *and* $\sin C$.

Proof. Consider the triangle in Fig. 7, where $\sin B = h/c$, $\triangle ABC = (1/2)ah = (1/2)ac\sin B$, $\triangle ABC = (1/2)bc\sin A$, $\triangle ABC = (1/2)ab\sin C$. Then

$$bc\sin A = ac\sin B = ab\sin C,$$

or

$$\frac{\sin A}{a} = \frac{\sin B}{b} = \frac{\sin C}{c}.$$

□

(6) $\cos(u - v) = \cos u\cos v + \sin u\sin v$

Proof. Consider the triangle $\triangle OBC$ in Fig. **??** with $O(0,0)$, $C(\cos u, \sin u)$, $B(\cos v, \sin v)$. By the

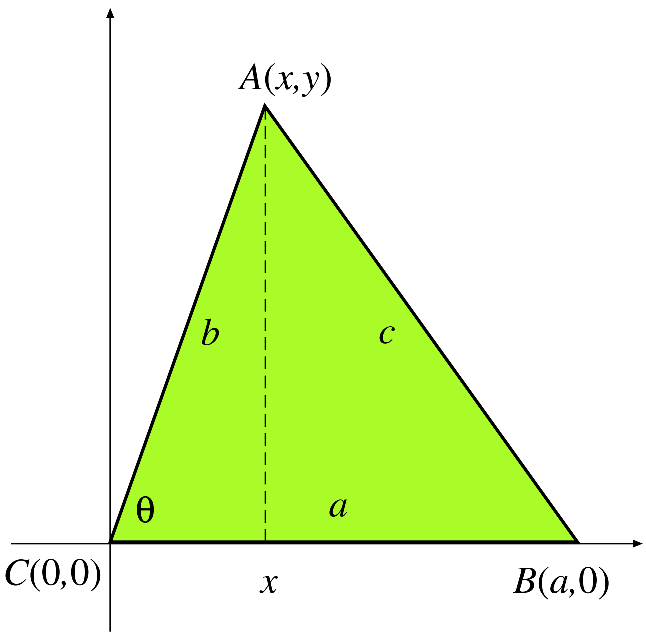

Figure 6: Law of Cosines

Law of Cosine we have two equations

$$BC^2 = 1^2 + 1^2 - 2 \times 1 \times 1 \times \cos(u - v) = 2 - 2\cos(u - v)$$
$$\begin{aligned} BC^2 &= (\cos u - \cos v)^2 + (\sin u - \sin v)^2 \\ &= \cos^2 u + \cos^2 v - 2\cos u \cos v + \sin^2 u + \sin^2 v - 2 \sin u \sin v. \end{aligned}$$
(by the Pythagorean theorem)

These give the desired equation. □

(7) $\cos(u + v) = \cos u \cos v - \sin u \sin v$

(8) $\cos(\frac{\pi}{2} - v) = \sin v$ *and* $\sin(\frac{\pi}{2} - v) = \cos v$

(9) $\sin(u + v) = \sin u \cos v + \cos u \sin v$

Proof.

$$\begin{aligned} \sin(u + v) &= \cos(\pi/2 - (u + v)) = \cos((\pi/2 - u) - v) \\ &= \cos(\pi/2 - u)\cos v + \sin(\pi/2 - u)\sin v \\ &= \sin u \cos v + \cos u \sin v. \end{aligned}$$

□

(10) $\sin(u - v) = \sin u \cos v - \cos u \sin v$

(11) $\cos 2u = \cos^2 u - \sin^2 u = 2\cos^2 u - 1 = 1 - 2\sin^2 u$

(12) $\sin 2u = 2 \sin u \cos u$

(13) $\cos^2(\frac{u}{2}) = \frac{1+\cos u}{2}$

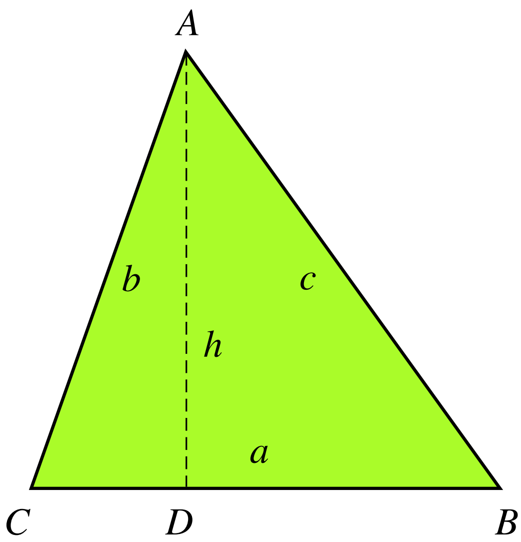

Figure 7: Law of Sines

Proof.

$$\cos u = \cos(2 \times \frac{u}{2}) = 2\cos^2(\frac{u}{2}) - 1.$$

□

(14) $\sin^2(\frac{u}{2}) = \frac{1-\cos u}{2}$

Proof.

$$\sin^2(\frac{u}{2}) = 1 - \cos^2(\frac{u}{2}) = 1 - \frac{1+\cos u}{2} = \frac{1-\cos u}{2}.$$

□

4 Differentiation of Trigonometric Functions

Proposition 4.1.

(1) $\lim_{\theta \to 0} \frac{\sin\theta}{\theta} = 1$

Proof. Since $\sin\theta/\theta$ is an even function, we have only to consider $\theta > 0$. See Fig. 9 and consider two sectors COD and AOB, and a triangle COB. We see

$$\begin{aligned} \text{Area of sector } COD &= (1/2)\theta\cos^2\theta \\ \text{Area of triangle } COB &= (1/2)\sin\theta\cos\theta \\ \text{Area of sector } AOB &= (1/2)\theta. \end{aligned}$$

Then

$$(1/2)\theta\cos^2\theta \le (1/2)\sin\theta\cos\theta \le (1/2)\theta, \quad \text{i.e. } \cos\theta \le \frac{\sin\theta}{\theta} \le \frac{1}{\cos\theta}.$$

Taking the limit, we have

$$1 \le \lim_{\theta \to 0} \frac{\sin\theta}{\theta} \le 1.$$

□

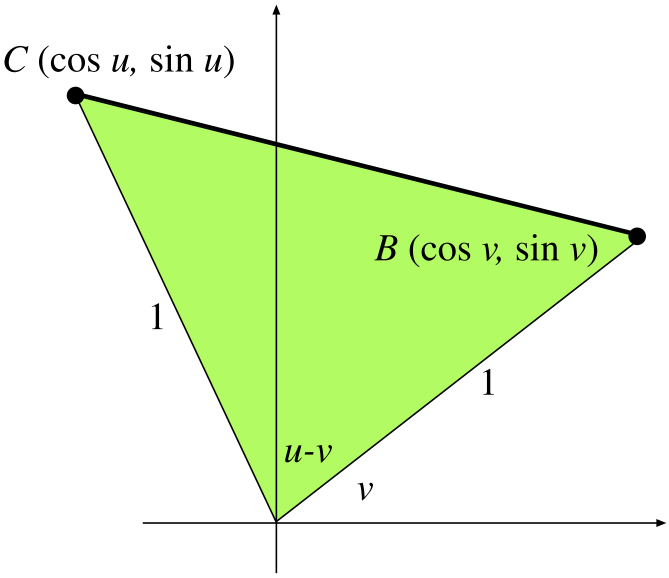

Figure 8: $\cos(u - v) = \cos u \cos v + \sin u \sin v$

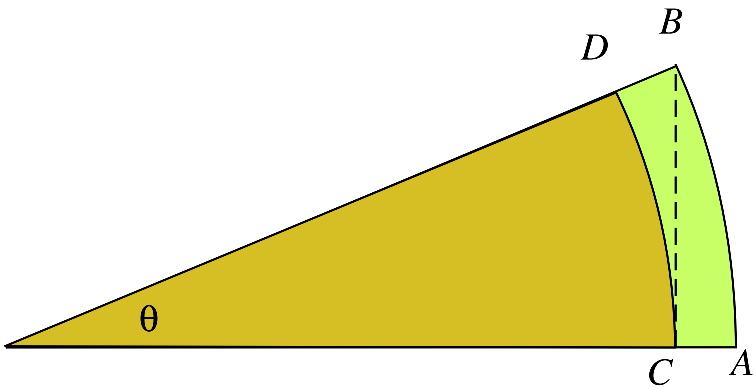

Figure 9: $\lim_{\theta\to 0} \sin\theta/\theta = 1$

(2) $\lim_{\theta\to 0} \dfrac{1-\cos\theta}{\theta} = 0$

Proof.

$$\begin{aligned}
\frac{1-\cos\theta}{\theta} &= \frac{1-\cos\theta}{\theta}\frac{1+\cos\theta}{1+\cos\theta} \\
&= \frac{1-\cos^2\theta}{\theta(1+\cos\theta)} \\
&= \frac{\sin^2\theta}{\theta(1+\cos\theta)} \\
&= \frac{\sin\theta}{\theta}\frac{\sin\theta}{1+\cos\theta}.
\end{aligned}$$

□

(3) $D_x(\sin x) = \cos x$

Proof. Note first that

$$\begin{aligned}\sin(x+\Delta x)-\sin x &= \cos x\sin\Delta x+\sin x\cos\Delta x-\sin x\\ &= \cos x\sin\Delta x+\sin x(\cos\Delta x-1).\end{aligned}$$

Then

$$\begin{aligned}\frac{d}{dx}\sin x &= \lim_{\Delta x\to 0}\frac{\sin(x+\Delta x)-\sin x}{\Delta x}\\ &= \lim_{\Delta x\to 0}\frac{\cos x\sin\Delta x+\sin x(\cos\Delta x-1)}{\Delta x}.\end{aligned}$$

Since

$$\lim_{\Delta x\to 0}\frac{\sin\Delta x}{\Delta x}=1 \quad \text{and} \quad \lim_{\Delta x\to 0}\frac{\cos\Delta x-1}{\Delta x}=0$$

we have

$$\begin{aligned}\lim_{\Delta x\to 0}\frac{\cos x\sin\Delta x+\sin x(\cos\Delta x-1)}{\Delta x} &= \cos x\lim_{\Delta x\to 0}\frac{\sin\Delta x}{\Delta x}+\sin x\lim_{\Delta x\to 0}\frac{\cos\Delta x-1}{\Delta x}\\ &= \cos x.\end{aligned}$$

□

(4) $D_x(\cos x)=-\sin x$

Proof.

$$D_x(\cos x)=D_x(\sin(\pi/2-x))=\cos(\pi/2-x)\times(-1)=-\sin x.$$

□

Definition 4.2 (OTHER TRIGONOMETRIC FUNCTIONS).

1. Tangent $\tan x=\dfrac{\sin x}{\cos x}$
2. Cotangent $\cot x=\dfrac{\cos x}{\sin x}=\dfrac{1}{\tan x}$
3. Secant $\sec x=\dfrac{1}{\cos x}$
4. Cosecant $\csc x=\dfrac{1}{\sin x}$

Proposition 4.3 (DERIVATIVES).

(1) $D_x(\tan x)=\sec^2 x$

(2) $D_x(\cot x)=-\csc^2 x$

(3) $D_x(\sec x)=\tan x\sec x$

(4) $D_x(\csc x)=-\cot x\csc x$

Proposition 4.4 (OTHER RELATIONSHIPS).

(1) $\tan^2 x+1=\sec^2 x$

Proof.

$$\tan^2 x+1=\frac{\sin^2 x}{\cos^2 x}+1=\frac{\sin^2 x+\cos^2 x}{\cos^2 x}=\frac{1}{\cos^2 x}=\sec^2 x.$$

□

(2) $\tan(x+\pi)=\tan x$ *and* $\cot(x+\pi)=\cot x$

(3) $\tan(-x)=-\tan x$ *and* $\cot(-x)=-\cot x$

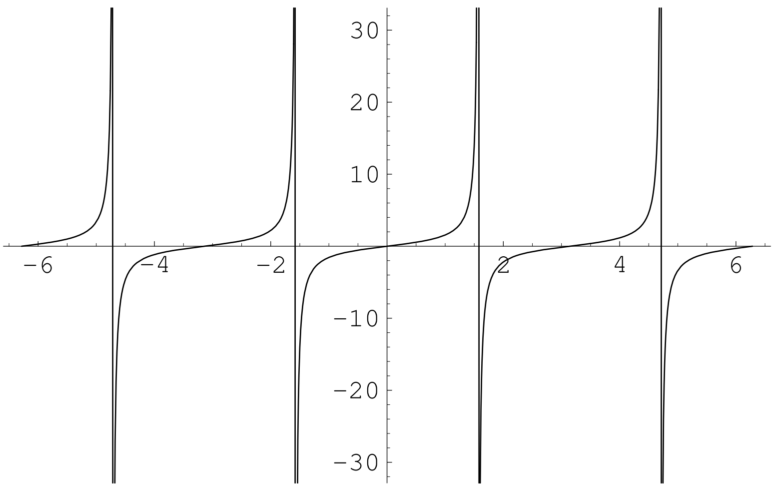

Figure 10: Tangent

5 Inverse Trigonometric Functions

Definition 5.1.
Looking at the graph of $y = \sin x$, we note that on the interval $-\frac{\pi}{2} \leq x \leq \frac{\pi}{2}$ the restriction of $\sin x$ is one-to-one. We then define $\sin^{-1} x$ to be the corresponding inverse function. The domain of this function is $[-1, 1]$, which is the range of $\sin x$. Thus,

1. $\sin^{-1}(x) = y$ if and only if $\sin y = x$.
2. The domain of $\sin^{-1} x$ is $[-1, 1]$.
3. The range of $\sin^{-1} x$ is $[-\frac{\pi}{2}, \frac{\pi}{2}]$.

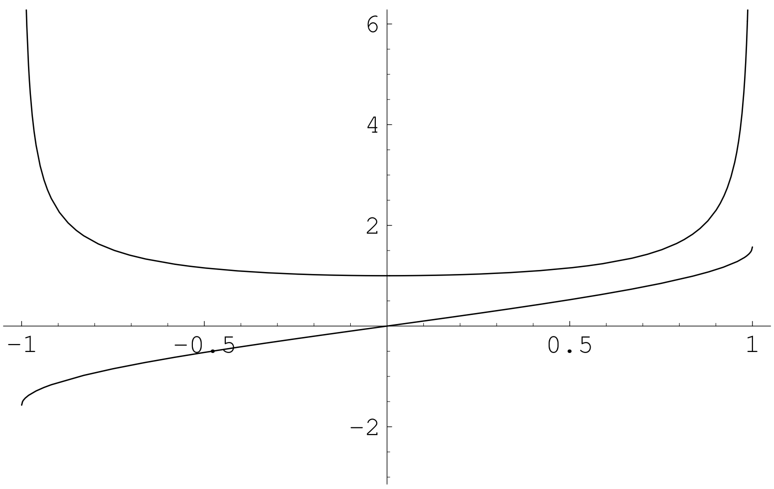

Figure 11: Inverse of Sine $\sin^{-1}(x)$ and its derivative $1/\sqrt{1-x^2}$

Proposition 5.2.
$D_x(\sin^{-1} x) = \dfrac{1}{\sqrt{1-x^2}}$

Proof. By the following remark

$$D_x(\sin^{-1} x) = \frac{1}{D_y(\sin y)}$$
$$= \frac{1}{\cos y} \quad \text{evaluated at } y \text{ with } \sin y = x.$$

Since the range of $\sin^{-1}$ is $[-\frac{\pi}{2}, \frac{\pi}{2}]$, where $\cos y$ is nonnegative, we have $\cos y = \sqrt{1-\sin^2 y} = \sqrt{1-x^2}$. □

Remark 5.3. For a one-to-one function $f : R \to R$ the inverse f^{-1} is defined by

$$y = f^{-1}(x) \text{ if and only if } x = f(y).$$

Its derivative $D_x(f^{-1}(x))$ is given by

$$D_x(f^{-1}(x)) = 1/D_y(f(y))|_{y=f^{-1}(x)} = 1/D_y(f(y))|_{x=f(y)}.$$

Definition 5.4 (THE INVERSE COSINE FUNCTION).
If we restrict the domain of $\cos x$ to $[0, \pi]$, we obtain a one-to-one function (with range $[-1, 1]$). So we can define $\cos^{-1} x$ to be the inverse of that restriction.

1. $\cos^{-1}(x) = y$ if and only if $\cos y = x$.
2. The domain of $\cos^{-1} x$ is $[-1, 1]$.
3. The range of $\cos^{-1} x$ is $[0, \pi]$.

Proposition 5.5.
$D_x(\cos^{-1} x) = -\frac{1}{\sqrt{1-x^2}}$

Proof. Left for the readers. □

Definition 5.6 (THE INVERSE TANGENT FUNCTION).

1. $\tan^{-1}(x) = y$ if and only if $\tan y = x$.
2. The domain of $\tan^{-1} x$ is $(-\infty, +\infty)$.
3. The range of $\tan^{-1} x$ is $(-\frac{\pi}{2}, \frac{\pi}{2})$.

Proposition 5.7.

(1) $D_x(\tan^{-1} x) = \frac{1}{1+x^2}$

(2) $D_x(\cot^{-1} x) = -\frac{1}{1+x^2}$

(3) $D_x(\sec^{-1} x) = \frac{1}{x\sqrt{x^2-1}}$

(4) $D_x(\csc^{-1} x) = -\frac{1}{x\sqrt{x^2-1}}$

Proof. Left for the readers. □

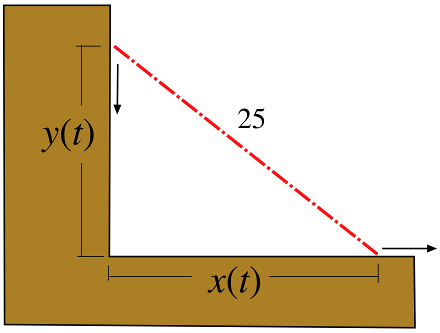

Figure 12: Sliding ladder

6 Related Rates

Example 6.1.
A 25-foot ladder rests against a vertical wall. If the bottom of the ladder is sliding away from the base of wall at the rate of 3 ft/sec, how fast is the top of the ladder moving down the wall when the bottom of the ladder is 7 feet from the base?

By the Pythagorean Theorem,

$$x^2(t) + y^2(t) = (25)^2 = 625.$$

This is the relation that holds between $x(t)$ and $y(t)$ for any time $t \geq 0$. Namely, the sum $x^2(t) + y^2(t)$ is a constant function of the value 625 as a function of time t. As we know that $(f(x) + g(x))' = f'(x) + g'(x)$, differentiating both sides with respect to t, we get

$$2x\frac{dx(t)}{dt} + 2y\frac{dy(t)}{dt} = 0.$$

This equation yields the desired velocity.

Example 6.2.
Gas is escaping from a spherical balloon at the rate of 2 m^3/min. How fast is the surface area shrinking when the radius is 12 m?

The volume is given by $V(r) = \frac{4}{3}\pi r^3$ and the surface area is give by $S(r) = 4\pi r^2$. Note that the radius r is a function of time t. Differentiating these equations by t,

$$\begin{aligned} \frac{dV(t)}{dt} &= 4\pi r^2(t)\frac{dr(t)}{dt}, \quad \text{which is now equal to } -2 \\ \frac{dS(t)}{dt} &= 8\pi r(t)\frac{dr(t)}{dt}. \end{aligned}$$

Then $\frac{dr(t)}{dt} = -1/(2\pi r^2(t))$. Thus $\frac{dS(t)}{dt} = 8\pi r(t)(-1/(2\pi r^2(t))) = -4/r(t)$, which is $-1/3$ when $r(t) = 12$.

Example 6.3.
Water is running out of a conical funnel at the rate of 1 cm^3/sec. If the radius of the funnel is 4 cm and the height is 8 cm, find the rate at which the water level is dropping when it is 2 cm from the top.

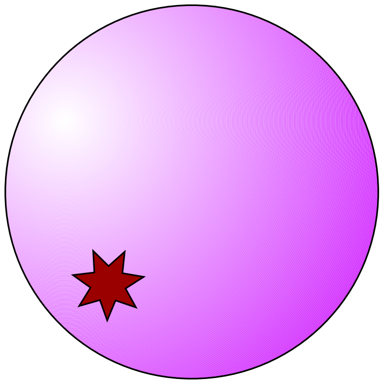

Figure 13: Balloon with a hole

Let $r(t)$ be the radius and $h(t)$ be the height of the surface of the water at time t, and let $V(t)$ be the volume of the water at time t. Then

$$V(t) = \frac{1}{3}\pi r^2(t)h(t) \quad \text{then using the relation } r(t) = \tfrac{1}{2}h(t) \text{ we have}$$
$$\frac{dV(t)}{dt} = \frac{1}{4}\pi h^2(t)\frac{dh(t)}{dt}, \quad \text{which is } -1.$$

Thus

$$\frac{dh(t)}{dt} = \frac{-4}{\pi h^2(t)}.$$

When the water level is 2 from the top, $h = 6$, then, at that time, $\frac{dh(t)}{dt} = \frac{-1}{9\pi}$.

Example 6.4.
A light hands H m above a street. An object h m tall directly under the light moves in a straight line along the street at v m/sec. Find a formula for the velocity $V(t)$ of the tip of the shadow cast by the object on the street at t seconds.

After t seconds, the object has moved a distance vt. Let $y(t)$ denote the position of the tip. Then

$$\frac{y(t) - vt}{y(t)} = \frac{h}{H}.$$

Hence

$$y(t) = \frac{Hvt}{H - h}, \quad \text{implying} \quad V(t) = \frac{dy(t)}{dt} = \frac{Hv}{H - h} = \frac{1}{1 - h/H}v.$$

7 Antiderivatives

Definition 7.1.
If $F'(x) = f(x)$, then F is called an *antiderivative* of f.

Remark 7.2.
Suppose $f'(x) = 0$ for any x. Then $f(x) = C$ for some constant C.

Proof. Suppose there are tow distinct points x_1 and x_2 such that $f(x_1) \neq f(x_2)$. Then by the law of mean we have $x \in (x_1, x_2)$ at which the derivative $f'(x)$ satisfies

$$f'(x) = \frac{f(x_1) - f(x_2)}{x_1 - x_2},$$

which is not zero. This is a contradiction. □

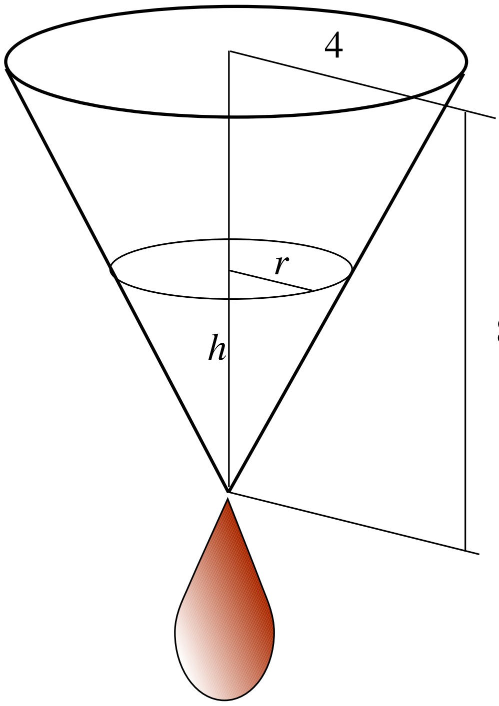

Figure 14: Conical funnel

Proposition 7.3.

1. *In general, if $F(x)$ is an antiderivative of $f(x)$, then $F(x) + C$ is also an antiderivative of $f(x)$, where C is any constant.*

2. *On the other hand, if $F(x)$ is an antiderivative of $f(x)$, and if $G(x)$ is any other antiderivative of $f(x)$, then $G(x) = F(x) + C$, for some constant C.*

Definition 7.4 (Notation and Terminology).
$\int f(x)dx$ will denote any antiderivative of $f(x)$. In this notation, $f(x)$ is called the *integrand*. An antiderivative $\int f(x)dx$ is also called an *indefinite integral*.

Proposition 7.5 (LAWS FOR ANTIDERIVATIVES). *The constant C on the right hand side means the functions on both sides differ within a constant number.*

Law 1. $\int 0dx = C.$

Law 2. $\int 1dx = x + C.$

Law 3. $\int adx = ax + C.$

Law 4. $\int x^r dx = \frac{x^{r+1}}{r+1} + C$ *for any rational number* $r \neq -1$.

Law 5. $\int af(x)dx = a\int f(x)dx\ +\ C.$

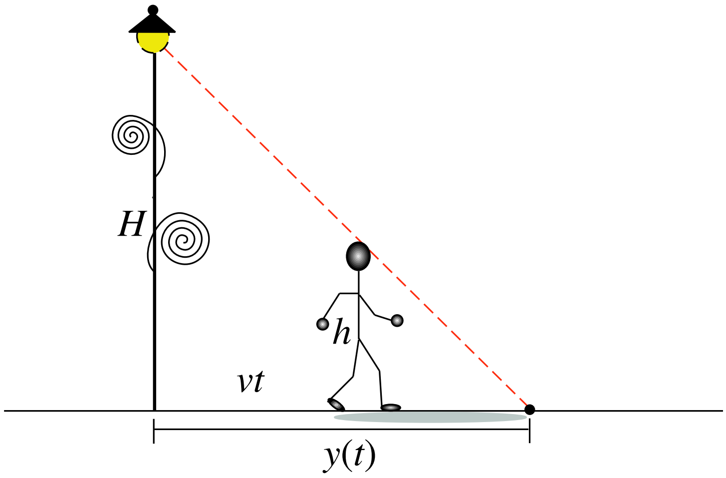

Figure 15: Moving shadow on the street

Law 6. $\int (f(x)+g(x))dx = \int f(x)dx + \int g(x)dx + C.$

Law 7. $\int (f(x)-g(x))dx = \int f(x)dx - \int g(x)dx + C.$

Law 8. $\int (g(x))^r g'(x)dx = \frac{1}{r+1}(g(x))^{r+1} + C$ *for any rational number* $r \neq -1$

Proof.

$$\begin{aligned} D_x\left(\frac{1}{r+1}(g(x))^{r+1}\right) &= \frac{1}{r+1}D_x((g(x))^{r+1}) \\ &= \frac{1}{r+1}(r+1)g(x)^r g'(x) \\ &= g(x)^r g'(x). \end{aligned}$$

□

Law 9. Substitution Method
Let $F(u) = \int f(u)du$. *Then*

$$F(g(x)) = \int f(g(x))g'(x)dx.$$

Proof. Let $u = g(x)$. Then

$$D_x(F(g(x))) = D_u(F(u))\frac{du}{dx} = f(u)g'(x) = f(g(x))g'(x),$$

meaning $F(g(x)) = \int f(g(x))g'(x)dx.$ □

Example 7.6. $\int x\sin(x^2)dx$?
$f(u) = \sin u, g(x) = x^2, g'(x) = 2x \quad \Rightarrow$

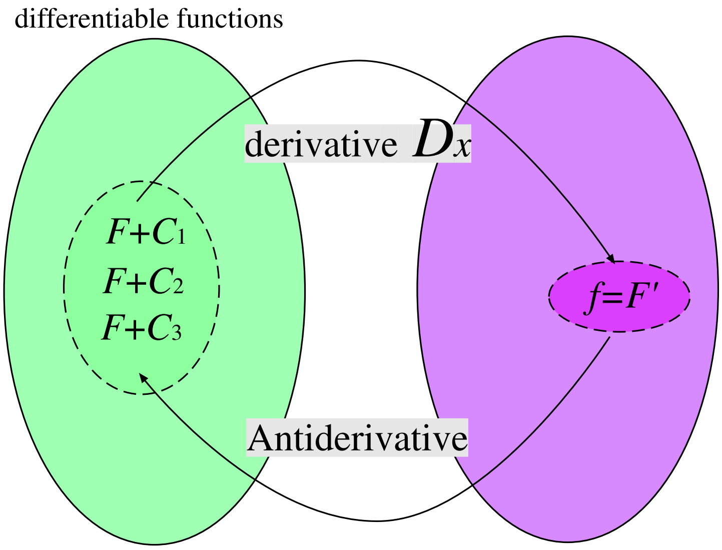

Figure 16: Derivative vs. Antiderivative

$$\begin{aligned}\int x\sin(x^2)dx &= \frac{1}{2}\int f(g(x))g'(x)dx \\ &= \frac{1}{2}(-\cos(x^2)) + C = -\frac{1}{2}\cos(x^2) + C.\end{aligned}$$

Law 10.

$$\int f(x)g'(x)dx = f(x)g(x) - \int f'(x)g(x)dx.$$

Proof. Since $D_x(f(x)g(x)) = f'(x)g(x) + f(x)g'(x)$, we have $f(x)g'(x) = D_x(f(x)g(x)) - f'(x)g(x)$ and hence

$$\begin{aligned}\int f(x)g'(x)dx &= \int D_x(f(x)g(x))dx - \int f'(x)g(x)dx \\ &= f(x)g(x) - \int f'(x)g(x)dx.\end{aligned}$$

□

Proposition 7.7.

$$\int \sin x dx = -\cos x + C$$

$$\int \cos x dx = \sin x + C$$

$$\int \sec^2 x dx = \tan x + C$$

$$\int \tan x \sec x dx = \sec x + C$$

$$\int \csc^2 x dx = -\cot x + C$$

$$\int \cot x \csc x dx = -\csc x + C$$

$$\int \frac{1}{\sqrt{1-x^2}} dx = \sin^{-1} x + C$$

$$\int \frac{1}{1+x^2} dx = \tan^{-1} x + C$$

$$\int \frac{1}{x\sqrt{x^2-1}} dx = \sec^{-1} x + C$$

$$\int \frac{1}{\sqrt{a^2-x^2}} dx = \sin^{-1}(\frac{x}{a}) + C \ \textit{for } a > 0$$

$$\int \frac{1}{a^2+x^2} dx = \frac{1}{a} \tan^{-1}(\frac{x}{a}) + C \ \textit{for } a > 0$$

$$\int \frac{1}{x\sqrt{x^2-a^2}} dx = \frac{1}{a} \sec^{-1}(\frac{x}{a}) + C \ \textit{for } a > 0$$

Example 7.8. The acceleration of gravity on the earth is 9.8 m/sec^2, meaning that the velocity v, a function of time t, of a body falling freely changes at the rate of

$$D_t(v(t)) = 9.8.$$

If the body is dropped from rest, what will its velocity be t seconds after it is released.

$$D_t(v(t)) = 9.8 \quad \Rightarrow \quad \int D_t(v(t))dt = \int 9.8dt$$
$$\Rightarrow \quad v(t) + C_1 = 9.8t + C_2 \text{ or } v(t) = 9.8t + C.$$

Since $v(0) = 0$, $C = 0$ and $v(t) = 9.8t$.

Example 7.9. Find the curve whose slope at the point (x, y) is $3x^2$ if the curve is required to pass through $(1, -1)$.

$$D_x(f(x)) = 3x^2 \quad \Rightarrow \quad \int D_x(f(x))dx = \int 3x^2 dx \quad \Rightarrow \quad f(x) = x^3 + C.$$

Since $f(1) = -1$, we have $C = -2$.

8 The Definite Integral, Area under a Curve

Definition 8.1 (AREA UNDER A CURVE, DEFINITE INTEGRAL).
Choose points $x_1, x_2, \cdots, x_{n-1}$ between a and b. Let $x_0 = a$ and $x_n = b$. Thus

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$$

The interval $[a, b]$ is divided into n subintervals $[x_0, x_1], [x_1, x_2], \cdots, [x_{n-1}, x_n]$. Denote the lengths of these subintervals by $\Delta_1 x, \Delta_2 x, \cdots, \Delta_n x$. Hence, if $1 \leq k \leq n$,

$$\Delta_k x = x_k - x_{k-1}$$

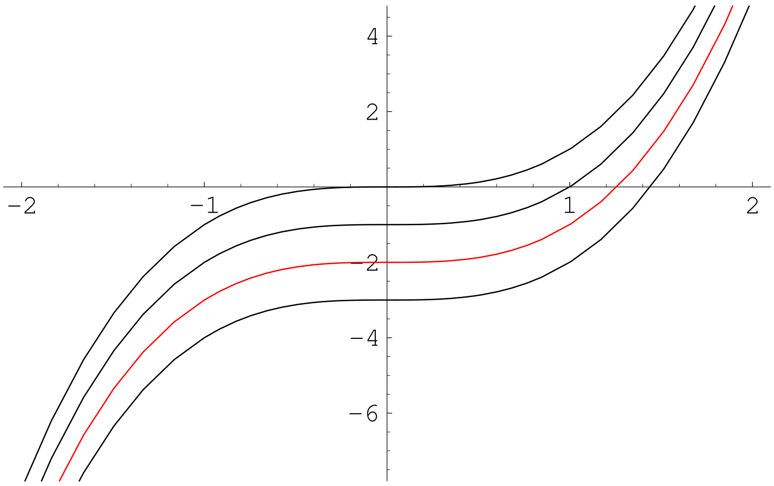

Figure 17: $f(x) = x^3 - 2$

Draw vertical line segments $x = x_k$ from the x axis up to the graph. This divides the region $\Re$ into n strips. Letting A be the area under the curve and $\Delta_k A$ denote the area of the kth strip, we obtain

$$A = \sum_{k=1}^{n} \Delta_k A$$

We can approximate the area $\Delta_k A$ in the following manner. Select any point x_k^* in the kth subinterval $[x_{k-1}, x_k]$. Draw the vertical line segment from the point x_k^* on the x axis up to the graph; the length of this segment is $f(x_k^*)$. The rectangle with base $\Delta_k x$ and height $f(x_k^*)$ has area $f(x_k^*)\Delta_k x$, which is approximately the area $\Delta_k A$ of the kth strip. Hence, the total area A under the curve is approximately the sum

$$\sum_{k=1}^{n} f(x_k^*)\Delta_k x = f(x_1^*)\Delta_1 x + f(x_2^*)\Delta_2 x + \cdots + f(x_n^*)\Delta_n x.$$

If successive approximations can be made as close as one wishes to a specific number, then that number will be denoted by

$$\int_a^b f(x)dx$$

and will be called the *definite integral* of f from a to b. Such a number does not exist in all cases, but it does exist, for example, when the function f is continuous on $[a, b]$. When $\int_a^b f(x)dx$ exists, its value is equal to the area A under the curve.

For any (not necessarily nonnegative) function f on $[a, b]$, sums of the form can be defined without using the notion of area. If there is a number to which these sums can be made as close as we wish, as n gets larger and larger and as the maximum of the lengths $\Delta_k x$ approaches 0, then that number is denoted $\int_a^b f(x)dx$ and is called the *definite integral* of f on $[a, b]$. When $\int_a^b f(x)dx$ exists, we say that f is *integrable* on $[a, b]$.

Proposition 8.2 (PROPERTIES OF THE DEFINITE INTEGRAL).

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

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

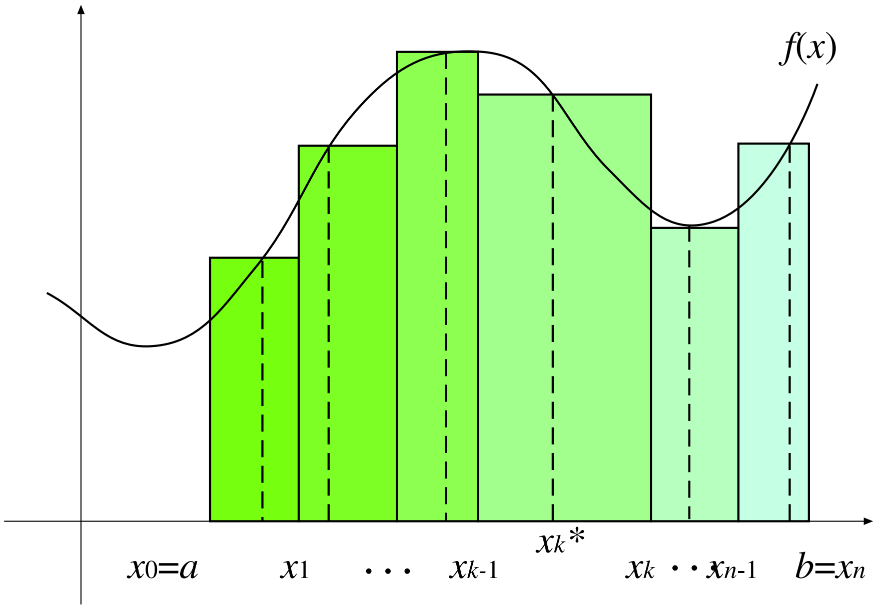

Figure 18: Area under a curve

(3) $\displaystyle\int_a^b (f(x)+g(x))dx = \int_a^b f(x)dx + \int_a^b g(x)dx$

(4) $\displaystyle\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx,$ *where* $a < c < b$

(5) *(i)* $\displaystyle\int_a^a f(x)dx = 0$

(ii) $\displaystyle\int_b^a f(x)dx = -\int_a^b f(x)dx$

Proof. Suppose that $a < b$ and you have intermediate points

$$b = x_0 > x_1 > \cdots > x_{n-1} > x_n = a.$$

Then

$$\begin{aligned}\int_b^a f(x)dx &= \lim \sum_{k=1}^{n} f(x_k^*)(x_k - x_{k-1}) \\ &= -\lim \sum_{k=1}^{n} f(x_k^*)(x_{k-1} - x_k) \\ &= -\int_a^b f(x)dx.\end{aligned}$$

□

Example 8.3.

$$\int_0^1 x^2 dx = \frac{1}{3}$$

Proof. Divide $[0,1]$ into n equal subintervals, i.e. $\Delta_k x = 1/n$. In the kth subinterval $[(k-1)/n, k/n]$, let x_k^* be the right endpoint k/n. Thus

$$\sum_{k=1}^{n} f(x_k^*)\Delta_k x = \sum_{k=1}^{n} \left(\frac{k}{n}\right)^2 \left(\frac{1}{n}\right) = \frac{1}{n^3}\sum_{k=1}^{n} k^2.$$

Since $\sum_{k=1}^{n} k^2 = n(n+1)(2n+1)/6$ (this can be seen by the induction), we have

$$\begin{aligned}\sum_{k=1}^{n} f(x_k^*)\Delta_k x &= \frac{1}{n^3}\frac{n(n+1)(2n+1)}{6} \\ &= \frac{1}{6}\left(\frac{n+1}{n}\right)\left(\frac{2n+1}{n}\right) \\ &= \frac{1}{6}\left(1+\frac{1}{n}\right)\left(2+\frac{1}{n}\right),\end{aligned}$$

which converges to $1/3$ as n goes to infinity.

□

What happens if we choose the left endpoint as x_k^*?

9 The Fundamental Theorem of Calculus

Theorem 9.1 (MEAN-VALUE THEOREM FOR INTEGRALS).
Let f be continuous on $[a,b]$. Then there exists c in $[a,b]$ such that

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

Proof. Since f is continuous, there are $m = \min\{\, f(x) \mid x \in [a,b]\,\}$ and $M = \max\{\, f(x) \mid x \in [a,b]\,\}$. Then

$$m(b-a) \le \int_a^b f(x)dx \le M(b-a) \quad \text{i.e.} \quad m \le \left(\int_a^b f(x)dx\right)/(b-a) \le M,$$

namely the value $(\int_a^b f(x)dx)/(b-a)$ lies between m and M. Note that there are $x_m \in [a,b]$ such that $m = f(x_m)$ and $x_M \in [a,b]$ such that $M = f(x_M)$. Apply the intermediate value theorem, we have c in the interval between x_m and x_M, that is contained in $[a,b]$ such that

$$\left(\int_a^b f(x)dx\right)/(b-a) = f(c).$$

□

Theorem 9.2. *Let f be continuous on $[a,b]$. Then for $x \in [a,b]$*

$$D_x\left(\int_a^x f(t)dt\right) = f(x).$$

Proof. Let $h(x) = \displaystyle\int_a^x f(t)dt$ and for δ consider the difference $h(x+\delta) - h(x)$. Then

$$\begin{aligned}h(x+\delta) - h(x) &= \int_a^{x+\delta} f(t)dt - \int_a^x f(t)dt \\ &= \int_x^{x+\delta} f(t)dt \\ &= \delta f(x_\delta^*)\end{aligned}$$

for some $x_\delta^* \in [x, x+\delta]$ by the mean value theorem for integrals. Dividing both sides by δ yields

$$\frac{h(x+\delta)-h(x)}{\delta} = f(x_\delta^*),$$

and taking the limit as $\delta \to 0$, we have

$$\lim_{\delta \to 0} \frac{h(x+\delta)-h(x)}{\delta} = \lim_{\delta \to 0} f(x_\delta^*) = f(x).$$

Here we used the fact that $x_\delta^* \to x$ as $\delta \to 0$ and f is continuous. Therefore we obtain the desired result

$$D_x\left(\int_a^x f(t)dt\right) = D_x(h(x)) = f(x).$$

□

Theorem 9.3 (FUNDAMENTAL THEOREM OF CALCULUS).
Let f be continuous on $[a,b]$, and let $F(x) = \int f(x)dx$, that is, F is an antiderivative of f. Then

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

Proof. By definition and Theorem 9.2 we have

$$D_x(F(x)) = f(x) = D_x\left(\int_a^x f(t)dt\right).$$

Then $F(x) = \int_a^x f(t)dt + K$ for some constant K. Since $F(a) = \int_a^a f(t)dt + K = 0 + K = K$, we see $K = F(a)$. □

The above result can be written as

$$\int_a^b f(x)dx = \int f(x)dx\Big]_a^b.$$

Example 9.4.

1.
$$\int_0^1 x^2dx = \int x^2dx\Big]_0^1 = \frac{1}{3}x^3\Big]_0^1 = \frac{1}{3}1^3 - \frac{1}{3}0^3 = \frac{1}{3}.$$

2.
$$\int_a^b x^r dx = \int x^r dx\Big]_a^b = \frac{1}{r+1}x^{r+1}\Big]_a^b = \frac{1}{r+1}(b^{r+1} - a^{r+1}).$$

10 The Natural Logarithm

We learned that

$$\int x^r dx = \frac{1}{r+1}x^{r+1} + C$$

when $r \neq -1$. What happens when $r = -1$?

Definition 10.1 (THE NATURAL LOGARITHM).
We define the function denoted by ln by

$$\ln x = \int_1^x \frac{1}{t}dt$$

for each $x > 0$.

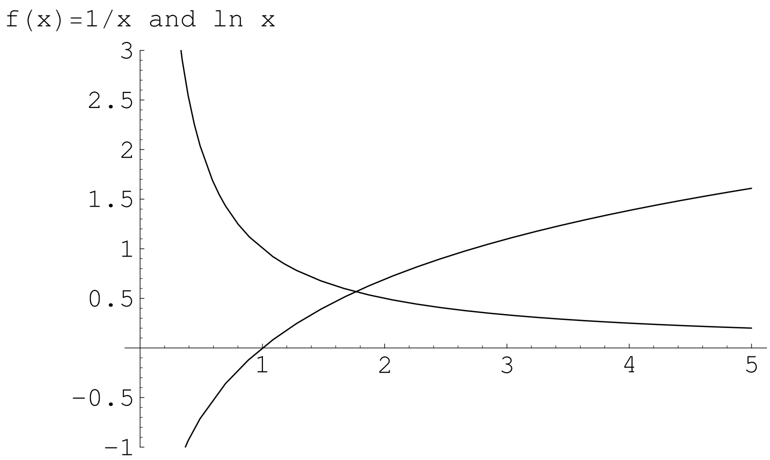

Figure 19: Hyperbolic function vs. Natural logarithm

Proposition 10.2.

$$D_x(\ln x) = \frac{1}{x} \quad \textit{for } x > 0.$$

Proof. By Theorem 9.2. □

Proposition 10.3 (PROPERTIES OF THE NATURAL LOGARITHM).

(1) $\ln 1 = 0$

(2) If $x > 1$, *then* $\ln x > 0$

(3) If $0 < x < 1$, *then* $\ln x < 0$

(4) (a) $D_x(\ln |x|) = \frac{1}{x}$

Proof. If $x > 0$, trivial. Suppose $x < 0$, then $|x| = -x$ and hence

$$D_x(\ln |x|) = D_x(\ln(-x)) = \frac{1}{-x} \cdot (-1) = \frac{1}{x}.$$

□

(b) $\displaystyle\int \frac{1}{x} dx = \ln |x| + C$ *for* $x \neq 0$

(5) $\ln uv = \ln u + \ln v$ *for any* $u, v > 0$.

Proof. Let a be a positive constant and consider the function $\ln ax$ and its derivative. By the chain rule we see

$$D_x(\ln ax) = \frac{1}{ax} D_x(ax) = \frac{1}{ax} a = \frac{1}{x} = D_x(\ln x),$$

holds for any $x > 0$, implying that $\ln ax = \ln x + K$ for some constant K. Take $x = 1$ then $\ln a = \ln 1 + K = K$. Thus we have

$$\ln ax = \ln a + \ln x.$$

Since this holds for any positive constant a, we have the result. □

(6) $\ln \frac{1}{v} = -\ln v$.

Proof.
$$0 = \ln 1 = \ln \frac{v}{v} = \ln v \left(\frac{1}{v}\right) = \ln v + \ln \frac{1}{v}.$$
□

(7) $\ln\left(\frac{u}{v}\right) = \ln u - \ln v$ *for any* $u, v > 0$.

Proof.
$$\ln\left(\frac{u}{v}\right) = \ln\left(u\frac{1}{v}\right) = \ln u + \ln\left(\frac{1}{v}\right) = \ln u - \ln v.$$
□

Proposition 10.4.

(1) $\ln x$ *is an increasing function.*

(2) $\ln u = \ln v$ *implies* $u = v$.

Proof. $\ln u = \ln v$ implies $\ln\left(\frac{u}{v}\right) = 0$. Then $\frac{u}{v} = 1$. □

(3) $\lim_{x\to+\infty} \ln x = +\infty$

(4) $\lim_{x\to 0^+} \ln x = -\infty$

(5) $\int \frac{g'(x)}{g(x)} dx = \ln|g(x)| + C$

Proof. By the chain rule
$$D_x(\ln|g(x)|) = \frac{1}{g(x)}g'(x) = \frac{g'(x)}{g(x)}.$$
□

11 Exponential and Logarithmic Functions

Definition 11.1 (EXPONENTIAL FUNCTION).
$\exp x$ is the inverse of $\ln x$.

Proposition 11.2 (PROPERTIES OF EXPONENTIAL FUNCTION).

(1) $\exp x > 0$ *for all* x

(2) $\ln(\exp x) = x$

(3) $\exp(\ln x) = x$

(4) $\exp x$ *is an increasing function.*

(5) $D_x(\exp x) = \exp x$

Proof. Let $f(x) = \ln x$ and $f^{-1}(y) = \exp y$. Then
$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))} \quad \textit{or} \quad D_y(\exp y) = \frac{1}{1/\exp y} = \exp y.$$
□

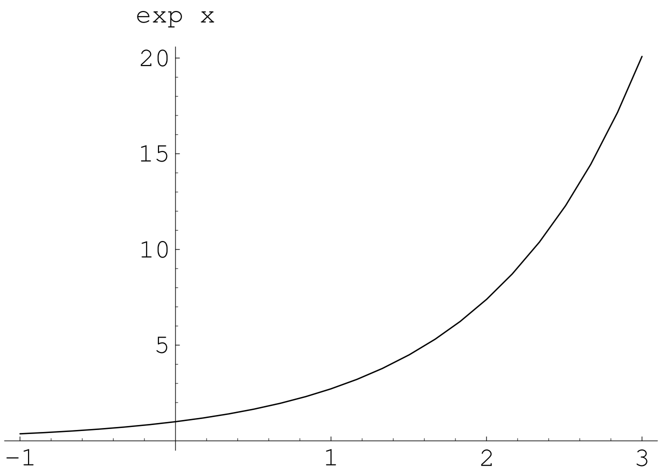

Figure 20: Exponential function

Proposition 11.3.

(1) $\int \exp x dx = \exp x + C$

(2) $\int \exp(-x)dx = -\exp(-x) + C$

(3) $\exp 0 = 1$

(4) $\exp(u+v) = (\exp u)(\exp v)$

Proof. Let $x = \exp u, y = \exp v$, i.e., $u = \ln x, v = \ln y$. Then $\exp(u+v) = \exp(\ln x + \ln y) = \exp(\ln(xy)) = xy = (\exp u)(\exp v)$. □

(5) $\exp(u-v) = \dfrac{\exp u}{\exp v}$

(6) $x < \exp x$ *for all* x

(7) $\lim_{x\to+\infty} \exp x = +\infty$

(8) $\lim_{x\to-\infty} \exp x = 0$

Definition 11.4.
Let e be the number such that $\ln e = 1$. This number e is called *Napier's constant*[1] or the *base of natural logarithm.*

Proposition 11.5.

1. $e \sim 2.71828182845904523536028747135266249775724709369995957496 6967\cdots\cdots\cdots$

[1] John Napier (1550 - 1617) was a Scottish mathematician. He is most remembered as the inventor of logarithms and of the decimal point. He was born in Merchiston Tower, Edinburgh. Napier is relatively little known outside mathematical circles where he made what is undoubtedly an extremely important advance in the history of mathematics. Logarithms made calculations by hand much easier and thereby opened the way to many later scientific advances. His work, Mirifici Logarithmorum Canonis Descriptio, contained thirty-seven pages of explanatory matter and ninety pages of tables, which facilitated the furtherment of astronomy, dynamics and physics. Napier's powers of invention were not confined to logarithms. He published a small treatise on a simple way to perform multiplication, the Rabdologiae, introducing a calculating device which became known as Napier's 'Rods' or 'Bones'. In an appendix he explained another method of multiplication and division using metal plates, which is one of the earliest known attempts at a mechanical means of calculation.

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

3. $\exp x = e^x$

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

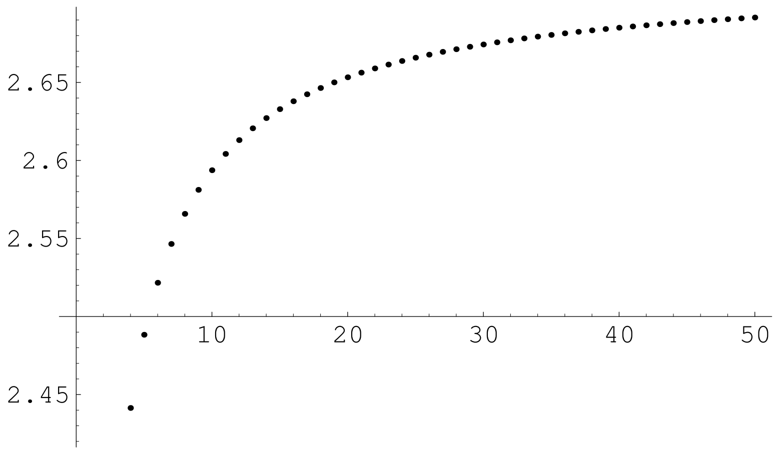

Figure 21: Convergence of $(1 + 1/n)^n$ to e

Definition 11.6.
For a real number a we define

$$a^x = \exp(x \ln a).$$

Proposition 11.7.

$$D_x(a^x) = (\ln a)a^x$$

Proof. $D_x(a^x) = D_x(\exp(x \ln a)) = \exp(x \ln a) \cdot \ln a = (\ln a)a^x$. □

12 L'Hôpital's Rule

Lemma 12.1 (CAUCHY'S MEAN VALUE THEOREM).
Let f and g be continuous on $[a, b]$ and differentiable on (a, b). Suppose $g'(x) \neq 0$ for any $x \in (a, b)$. Then there is $c \in (a, b)$ such that

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

Proof. Note first that $g(b) \neq g(a)$, since otherwise there would be a point, say $d \in (a, b)$ with $g'(d) = 0$ by Rolle's Theorem, which contradicts the assumption. Let

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

and

$$F(x) = f(x) - f(a) - k(g(x) - g(a)).$$

Then F is continuous on $[a, b]$, differentiable on (a, b), and $F(a) = F(b) = 0$. By applying Rolle's Theorem there is $c \in (a, b)$ such that

$$F'(c) = f'(c) - kg'(c) = 0 \quad \text{or} \quad \frac{f'(c)}{g'(c)} = k.$$

□

Theorem 12.2 (L'HÔPITAL'S RULE).
Suppose that f and g are continuous on an interval $[a-\epsilon, a+\epsilon]$, differentiable on $(a-\epsilon, a+\epsilon)\setminus\{a\}$, and $g'(x)\neq 0$ for all $x\in(a-\epsilon, a+\epsilon)\setminus\{a\}$. If $f(a)=g(a)=0$ and there exists

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

then

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

Proof. Let x be a point in $(a, a+\epsilon)$. By Lemma 12.1 there is a point $c\in(a,x)$ such that

$$\frac{f(x)-f(a)}{g(x)-g(a)}=\frac{f'(c)}{g'(c)},$$

which together with $f(a)=g(a)=0$ yields

$$\frac{f(x)}{g(x)}=\frac{f'(c)}{g'(c)}.$$

The same equality will be obtained for $x\in(a-\epsilon,a)$. Since $x\to a$ implies $c\to a$, where the right hand side is assumed to have a limit. □

Example 12.3. Prove that

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

Proof. Let $f(x)=x-\ln(1+x)$ and $g(x)=x^2$. Clearly $f(x)\to 0$ and $g(x)\to 0$ as $x\to 0$, and also $f'(x)\to 0$ and $g'(x)\to 0$ as $x\to 0$.

$$f''(x)=\frac{1}{(1+x)^2}\qquad g''(x)=2.$$

Then

$$\lim_{x\to 0}\frac{f(x)}{g(x)}=\lim_{x\to 0}\frac{f'(x)}{g'(x)}=\lim_{x\to 0}\frac{f''(x)}{g''(x)}$$
$$=\lim_{x\to 0}\frac{1/(1+x)^2}{2}=\frac{1}{2}.$$

□

Corollary 12.4.
Suppose that $\lim_{x\to+\infty}f(x)=\lim_{x\to\infty}g(x)=0$ and there exists

$$\lim_{x\to+\infty}\frac{f'(x)}{g'(x)}$$

[1]Guillaume De l'Hôpital served as a cavalry officer but resigned because of nearsightedness. From that time on he directed his attention to mathematics. L'Hôpital was taught calculus by Johann Bernoulli from the end of 1691 to July 1692. L'Hôpital was a very competent mathematician and solved the brachystochrone problem. The fact that this problem was solved independently by Newton, Leibniz and Jacob Bernoulli puts l'Hôpital in very good company. L'Hôpital's fame is based on his book Analyse des infiniment petits pour l'intelligence des lignes courbes (1696) which was the first text-book to be written on the differential calculus. In the introduction l'Hôpital acknowledges his indebtedness to Leibniz, Jacob Bernoulli and Johann Bernoulli but l'Hôpital regarded the foundations provided by him as his own ideas. In this book is found the rule, now known as L'Hôpital's rule, for finding the limit of a rational function whose numerator and denominator tend to zero at a point.

Note that l'Hôpital's name is commonly seen spelled both "l'Hospital" and "l'Hôpital" , the two being equivalent in French spelling.

then

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

The same result holds when $x \to -\infty$.

Corollary 12.5.
Suppose that $\lim_{x\to a-0} f(x) = \lim_{x\to a-0} g(x) = 0$ *and there exists*

$$\lim_{x\to a-0} \frac{f'(x)}{g'(x)}$$

then

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

The same result holds when $x \to a + 0$.

L'Hôpital's rule occasionally fails to yield useful results, as in the case of the function $\lim_{x\to+\infty} x/(x^2+1)^{1/2}$. Repeatedly applying the rule gives expressions which oscillate and never converge:

$$\begin{aligned}\lim_{x\to\infty} \frac{x}{(x^2+1)^{1/2}} &= \lim_{x\to\infty} \frac{1}{x(x^2+1)^{-1/2}} \\ &= \lim_{x\to\infty} \frac{(x^2+1)^{1/2}}{x} \\ &= \lim_{x\to\infty} \frac{x(x^2+1)^{-1/2}}{1} \\ &= \lim_{x\to\infty} \frac{x}{(x^2+1)^{1/2}}. \qquad ???\end{aligned}$$

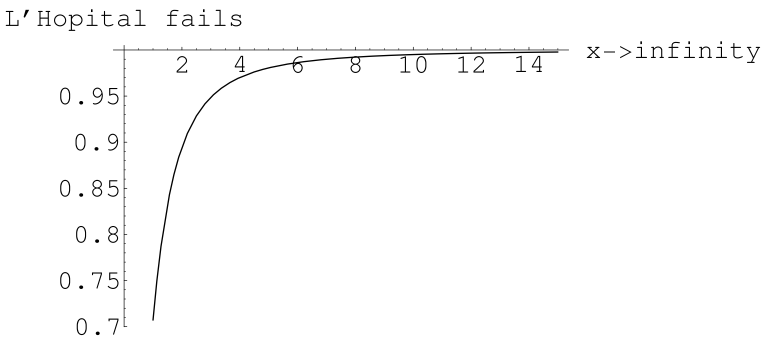

Figure 22: $\lim_{x\to+\infty} x/(x^2+1)^{1/2} = 1$

13 Arc Length

Proposition 13.1 (ARC LENGTH).
Let f *be the differentiable on* $[a, b]$. *Consider the part of the graph of* f *from* $(a, f(a))$ *to* $(b, f(b))$. *Let us find a formula for the length* L *of this curve. Divide* $[a, b]$ *into* n *equal subintervals, each of length* Δx. *To each point* x_k *in this subdivision there corresponds a point* $P_k(x_k, f(x_k))$ *on the curve. For large*

n,the sum $\overline{P_0P_1} + \overline{P_1P_2} + \cdots + \overline{P_{n-1}P_n} = \sum_{k=1}^{n} \overline{P_{k-1}P_k}$ *of the lengths of the line segments* $P_{k-1}P_k$ *is an approximation to the length of the curve.*

By the distance formula,

$$\overline{P_{k-1}P_k} = \sqrt{(x_k - x_{k-1})^2 + (f(x_k) - f(x_{k-1}))^2}$$

Now, $x_k - x_{k-1} = \Delta x$ *and, by the law of the mean,*

$$f(x_k) - f(x_{k-1}) = (x_k - x_{k-1})f'(x_k^*) = (\Delta x)f'(x_k^*)$$

for some x_k^* *in* (x_{k-1}, x_k). *Thus,*

$$\begin{aligned}
\overline{P_{k-1}P_k} &= \sqrt{(\Delta x)^2 + (\Delta x)^2(f'(x_k^*))^2} \\
&= \sqrt{(1 + (f'(x_k^*))^2)(\Delta x)^2} \\
&= \sqrt{1 + (f'(x_k^*))^2}\sqrt{(\Delta x)^2} = \sqrt{1 + (f'(x_k^*))^2}\Delta x \\
\sum_{k=1}^{n} \overline{P_{k-1}P_k} &= \sum_{k=1}^{n} \sqrt{1 + (f'(x_k^*))^2}\Delta x
\end{aligned}$$

The right-hand sum is an approximating sum for the definite integral

$$\int_a^b \sqrt{1 + (f'(x))^2} dx.$$

Therefore, letting $n \to +\infty$, *we get the arc length formula:*

$$L = \int_a^b \sqrt{1 + (f'(x))^2} dx.$$

14 Parametric Representation of Curves

Proposition 14.1 (PARAMETRIC EQUATIONS).
If the coordinates (x, y) *of a point* P *on a curve are given as functions* $x = f(u), y = g(u)$ *of a third variable or parameter,* u*, the equations* $x = f(u)$ *and* $y = g(u)$ *are called parametric equations of the curve.*

Proposition 14.2 (First Derivative).

$$\frac{dy}{dx} = \left(\frac{dy}{du}\right) \bigg/ \left(\frac{dx}{du}\right)$$

Proposition 14.3 (Second Derivative).

$$\frac{d^2y}{dx^2} = \left(\frac{d}{du}\left(\frac{dy}{dx}\right)\right) \bigg/ \frac{dx}{du}$$

Proposition 14.4 (ARC LENGTH FOR A PARAMETRIC CURVE).
If a curve is given by parametric equations $x = f(t)$*,* $y = g(t)$*, then the length of the arc of the curve between the points corresponding to parameter values* t_1 *and* t_2 *is*

$$L = \int_{t_1}^{t_2} \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} dt$$

15 Taylor and Maclaurin Series, Taylor's Formula with Remainder

Definition 15.1 (TAYLOR AND MACLAURIN SERIES).
Let f be a function that is infinitely differentiable at $x = c$, that is, the derivatives $f^{(n)}(c)$ exist for all positive integers n. The *Taylor series* for f about c is the power series

$$\sum_{n=0}^{+\infty} a_n(x-c)^n = a_0 + a_1(x-c) + a_2(x-c)^2 + \cdots$$

where

$$a_n = \frac{f^{(n)}(c)}{n!} \quad \text{for all } n \quad \text{and} \quad n! = n \times (n-1) \times (n-2) \times \cdots \times \times 1.$$

Note that $f^{(0)}$ is taken to mean the function f itself, so that $a_0 = f(c)$. The *Maclaurin*[2] *series for f* is the Taylor series for f about 0, that is, the power series

$$\sum_{n=0}^{+\infty} b_n x^n = b_0 + b_1 x + b_2 x^2 + \cdots$$

where

$$b_n = \frac{f^{(n)}(0)}{n!} \quad \text{for all } n.$$

Example 15.2.

1. $\sin x \quad : \quad x - \dfrac{x^3}{3!} + \dfrac{x^5}{5!} - \dfrac{x^7}{7!} + \cdots$

2. $\exp x \quad : \quad 1 + \dfrac{x}{1!} + \dfrac{x^2}{2!} + \dfrac{x^3}{3!} + \cdots$

3. $\ln(1+x) \quad : \quad x - \dfrac{x^2}{2} + \dfrac{x^3}{3} - \dfrac{x^4}{4} + \cdots$

Theorem 15.3 (TAYLOR'S FORMULA WITH REMAINDER).
Let f be a function such that its $(n+1)$st derivative $f^{(n+1)}$ exists in an open interval containing a and b. Then there is some c between a and b such that

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

where,

$$R_n(c) = \frac{f^{(n+1)}(c)}{(n+1)!}(b-a)^{n+1},$$

which is called the remainder term.

[2] Colin Maclaurin was born in February of 1698 in Kilmodan, Argyllshire, Scotland. He attended the University of Glasgow at age eleven (not unusual back then) and graduated at age fourteen. After graduation he remained at Glasgow to study divinity for a period and in 1717, at age nineteen, he became professor of mathematics at Marischal College in the University of Aberdeen. In 1725 he was appointed deputy of the mathematical professor at Edinburgh, James Gregory (nephew of the famous James Gregory), upon the recommendation of Isaac Newton, who actually offered to pay Maclaurin's salary, being so impressed with his work. Eventually, Maclaurin went on to succeed Gregory. (An interesting aside, the Maclaurin series for many trigonometric functions was developed and published by James Gregory before Maclaurin was even born, but Maclaurin wasn't aware of this and published them in Methodus incrementorum directa et inversa.) In 1733 he married Married Anne Stewart, the daughter of the Solicitor General of Scotland. He actively opposed the Young Pretender of the Jacobite Rebellion in 1745 and assisted in the defence of Edinburgh but had to flee to York upon the approach of the Highlanders. He then returned after the Jacobite army marched south, but all this running around exhausted him and he got sick and died on June 14, 1746.

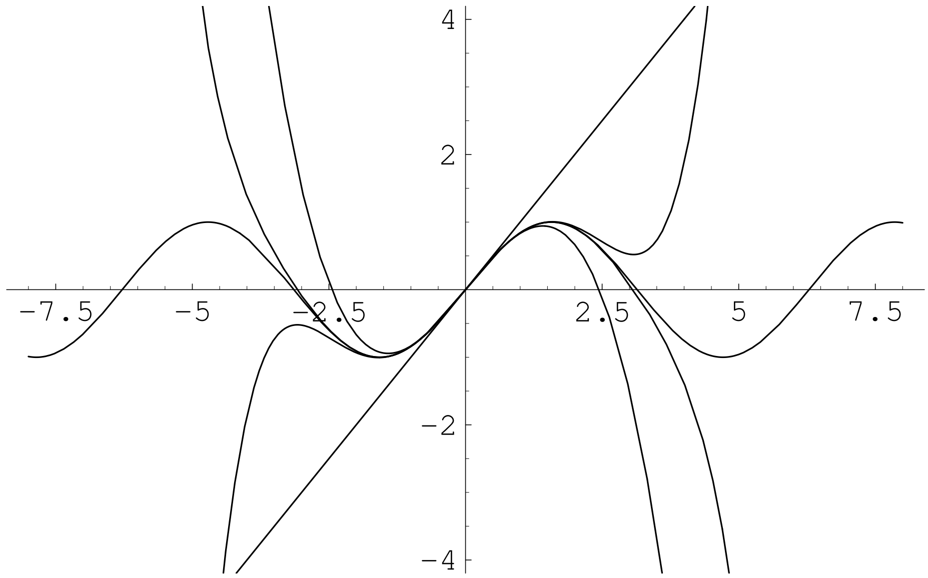

Figure 23: Taylor series of $\sin x$

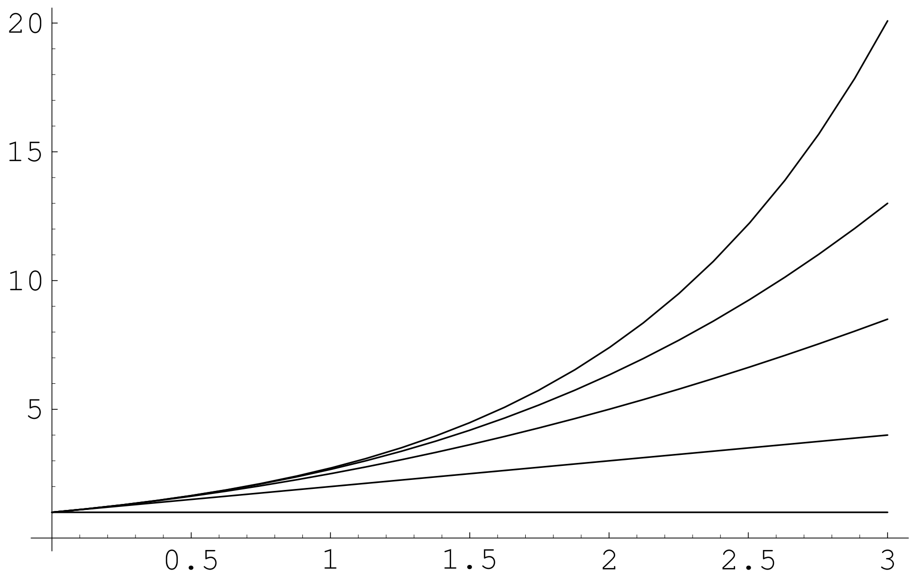

Figure 24: Taylor series of $\exp x$

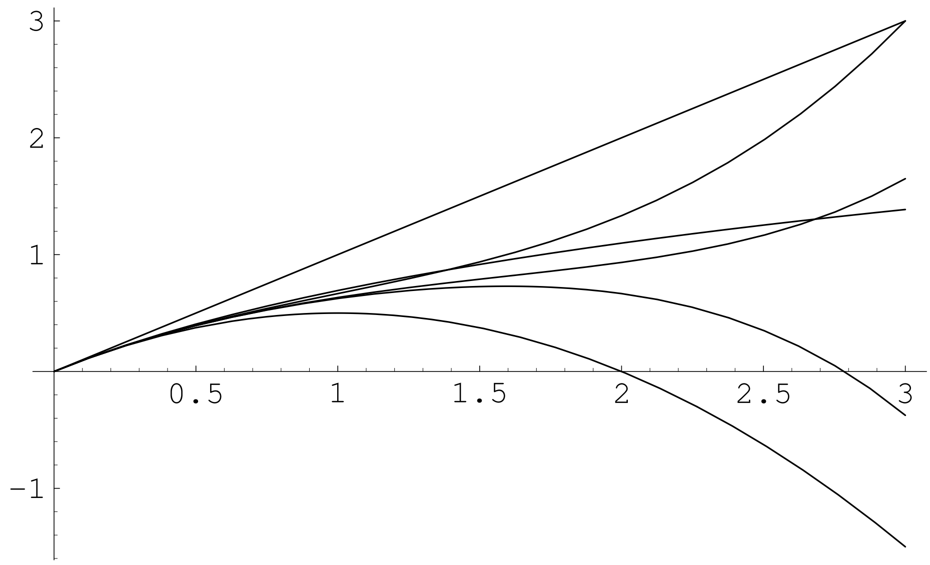

Figure 25: Spot the function $\ln(1+x)$! Taylor series of $\ln(1+x)$

Proof. Let

$$F(x) = f(b) - \left\{ f(x) + f'(x)(b-x) + \frac{f''(x)}{2!}(b-x)^2 + \cdots \right.$$
$$\left. + \frac{f^{(n)}(x)}{n!}(b-x)^n \right\} - K(b-x)^{n+1}.$$

Then $F(b) = 0$. We can choose K such that $F(a) = 0$. Note that $F(a) = 0$ is equivalent to $f(b) = f(a) + f'(a)(b-a) + \cdots + \frac{f^{(n)}(a)}{n!}(b-a)^n + K(b-a)^{n+1}$. Since $F(a) = F(b) = 0$ for this K, applying Rolle's theorem we obtain c between a and b such that $F'(c) = 0$, i.e.

$$0 = F'(c) = -f'(c) + \left\{ f'(c) - \frac{b-c}{1!} f''(c) \right\} + \left\{ \frac{b-c}{1!} f''(c) - \frac{(b-c)^2}{2!} f'''(c) \right\} + \cdots$$
$$+ \left\{ \frac{(b-c)^{n-1}}{(n-1)!} f^{(n)}(c) - \frac{(b-c)^n}{n!} f^{(n+1)}(c) \right\} + K(n+1)(b-c)^n$$
$$= -\frac{(b-c)^n}{n!} f^{(n+1)}(c) + K(n+1)(b-c)^n.$$

Then the constant K chosen satisfies

$$K = \frac{f^{(n+1)}(c)}{(n+1)!}.$$

□

Example 15.4.

1. $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots + (-1)^{n-1} \frac{x^{2n-1}}{(2n-1)!} + R_{2n+1}$,

   where $R_{2n+1} = (-1)^n \frac{x^{2n+1}}{(2n+1)!} \cos \theta x, \; 0 < \theta < 1$

2. $\exp x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^{n-1}}{(n-1)!} + R_n$,

   where $R_n = \frac{x^n}{n!} \exp \theta x, \; 0 < \theta < 1$

3. $\ln(1+x) = x - \dfrac{x^2}{2} + \dfrac{x^3}{3} - \dfrac{x^4}{4} + \cdots + (-1)^{n-2}\dfrac{x^{n-1}}{n-1} + R_n$,
where $R_n = (-1)^{n-1}\dfrac{x^n}{n}\left(\dfrac{1}{1+\theta x}\right)^n$, $0 < \theta < 1$

Example 15.5. The function

$$f(x) = \begin{cases} 0 & \text{if } x = 0 \\ \exp(-\frac{1}{x^2}) & \text{if } x \neq 0 \end{cases}$$

is so flat at the origin that all its derivatives there are zero (Try to prove it! It's not so easy as you guess.), implying its Maclaurin series is zero everywhere.

$$f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n + \cdots \equiv 0 \neq f(x) \text{ if } x \neq 0.$$

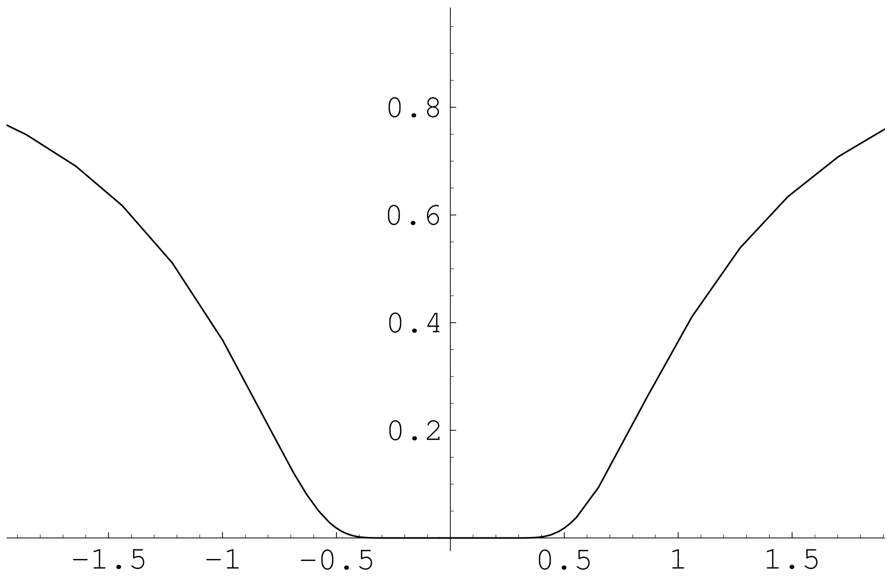

Figure 26: $\exp(-1/x^2)$ is so flat at the origin

16 Multivariable Functions, Limit and Continuity

Definition 16.1 (MULTIVARIABLE FUNCTIONS).
Suppose D is the set of n-tuples or n-dimensional vectors of real numbers $(x_1, x_2, \ldots, x_n)$. A *real valued function* f on D is a rule that assigns a real number

$$w = f(x_1, x_2, \ldots, x_n)$$

to each element in D. The set D is called the *domain* of the function.

We focus on functions with *two* variables, which will be denoted by x and y.

Definition 16.2 (EUCLIDEAN NORM).
For a two dimensional vector (x, y) the nonnegative real number $\sqrt{x^2 + y^2}$ is the *Euclidean norm* of (x, y), which is denoted by $\|(x, y)\|$ and defined in general as $\sqrt{\sum_{i=1}^{n} x_i^2}$ for n-dimensional vector $x = (x_1, x_2, \ldots, x_n)$.

Definition 16.3 (NEIGHBORHOOD).
The δ-*neighborhood* of (x_0, y_0) is the set

$$B((x_0, y_0), \delta) = \{ (x, y) \mid \|(x, y) - (x_0, y_0)\| < \delta \}.$$

Remark 16.4. Since

$$\|(x, y) - (x_0, y_0)\| = \|(x - x_0, y - y_0)\| = \sqrt{(x - x_0)^2 + (y - y_0)^2},$$

we see that

$$B((x_0, y_0), \delta) = \{ (x, y) \mid (x - x_0)^2 + (y - y_0)^2 < \delta^2 \}$$

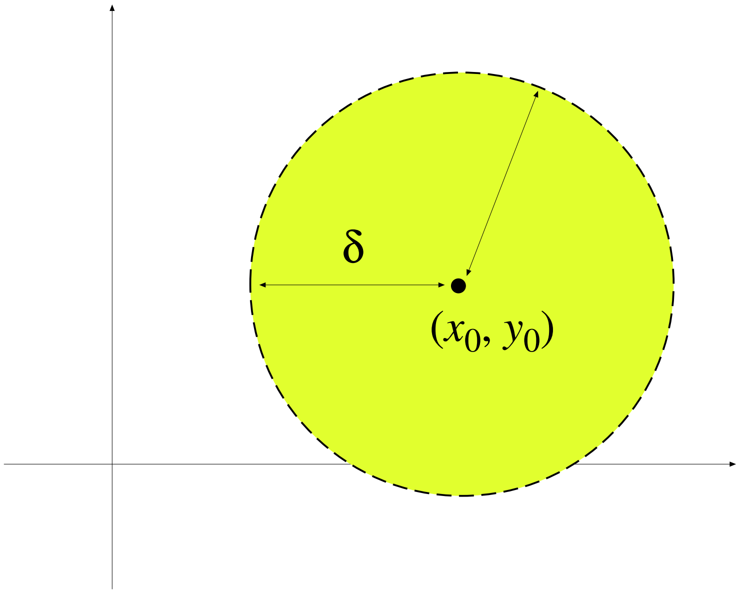

Figure 27: δ-neighborhood $B((x_0, y_0), \delta)$ of (x_0, y_0)

Definition 16.5 (LIMIT).
We say that a function $f(x, y)$ approaches the *limit* L as (x, y) approaches (x_0, y_0) and write

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = L$$

if, for every $\epsilon > 0$, there is a corresponding number $\delta > 0$ such that for all (x, y) in the domain of f,

$$0 < \|(x, y) - (x_0, y_0)\| < \delta \quad \Rightarrow \quad |f(x, y) - L| < \epsilon.$$

It should be noted that the function value at (x_0, y_0) does not affect the limit. It can be undefined there.

Proposition 16.6.
Suppose $\lim_{(x,y)\to(x_0,y_0)} f(x, y) = L$ *and* $\lim_{(x,y)\to(x_0,y_0)} g(x, y) = M$. *Then*

(1) $\displaystyle\lim_{(x,y)\to(x_0,y_0)} \{f(x, y) \pm g(x, y)\} = L \pm M$

(2) $\displaystyle\lim_{(x,y)\to(x_0,y_0)} f(x, y) \cdot g(x, y) = L \cdot M$

(3) $\displaystyle\lim_{(x,y)\to(x_0,y_0)} \frac{f(x, y)}{g(x, y)} = \frac{L}{M}$ *if* $M \neq 0$

Definition 16.7 (CONTINUITY).
A function f is *continuous at the point* (x_0, y_0) if

1. f is defined at (x_0, y_0)
2. $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y)$ exists
3. $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y) = f(x_0, y_0)$.

A function f is *continuous* if it is continuous at every point of its domain.

Example 16.8. The function

$$f(x,y) = \begin{cases} \dfrac{2xy}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

is continuous at every point except the origin $(0,0)$, at which it is not continuous.

Proof. Let m be an arbitrary value and consider the line $y = mx$. Then

$$f(x, mx) = \frac{2x(mx)}{x^2 + (mx)^2} = \frac{2m}{1+m^2} \quad \text{whenever } x \neq 0.$$

Therefore f has this number as its limit as (x, y) approaches $(0, 0)$ along the line. This limit changes with m, and so $\lim_{(x,y)\to(0,0)} f(x, y)$ fails to exist. □

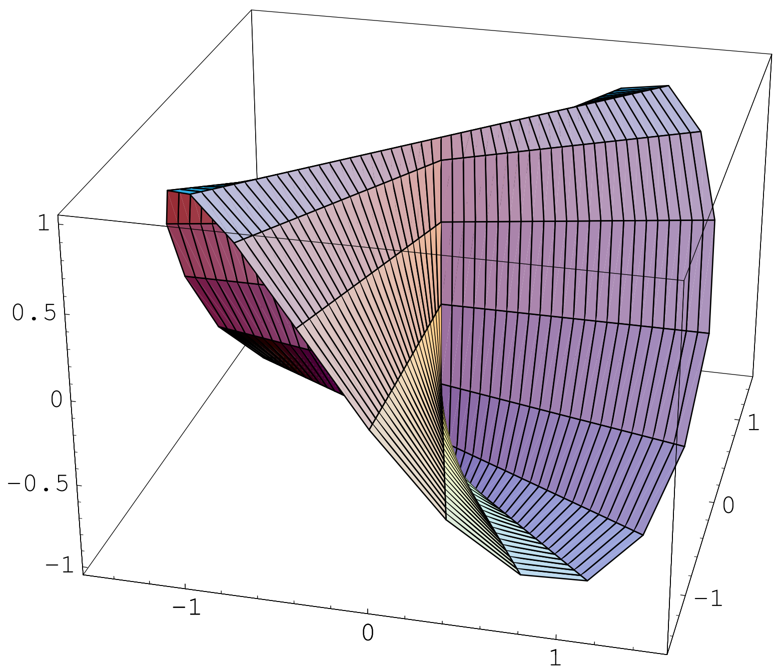

Figure 28: $\lim_{(x,y)\to(0,0)} f(x, y)$ does not exist

Proposition 16.9 (TWO-PATH TEST).
If a function f has different limits along two different paths as (x, y) approaches (x_0, y_0), then $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y)$ does not exist.

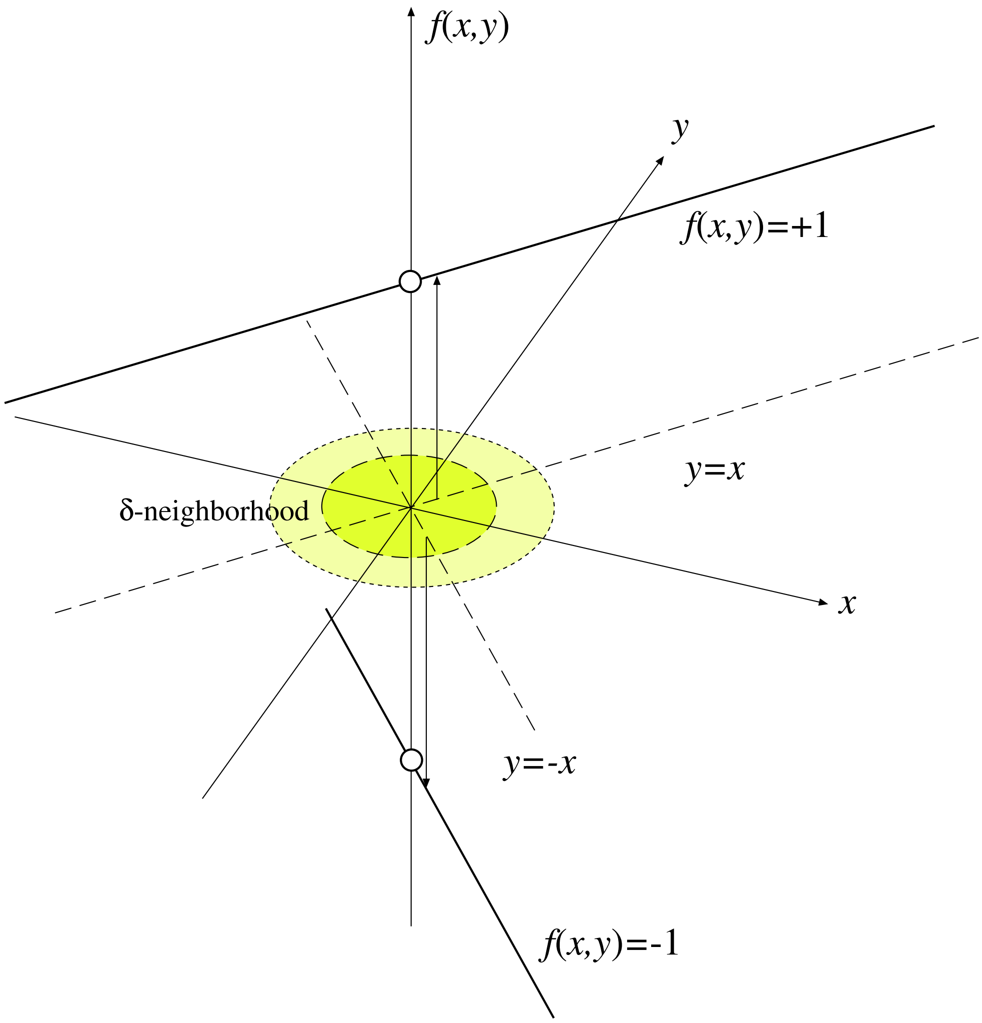

Figure 29: Why $\lim_{(x,y)\to(0,0)} f(x,y)$ does not exist

Example 16.10. The function

$$f(x,y) = \frac{2x^2y}{x^4+y^2}$$

has no limit as (x,y) approaches $(0,0)$.

Proof. Let (x,y) approach to $(0,0)$ along the curve $y = kx^2$. □

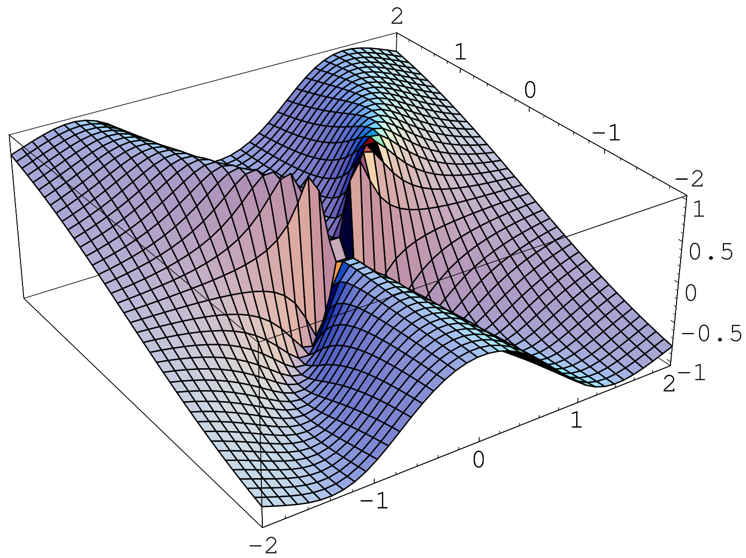

Figure 30: $f(x,y) = \frac{2x^2y}{x^4+y^2}$

Example 16.11. Another example:

$$f(x,y) = \frac{x}{\sqrt{x^2+y^2}}$$

17 Partial Derivatives

Let (x_0, y_0) be a point in the domain of f. The vertical plane defined by $y = y_0$ will cut the surface $z = f(x,y)$ in the curve $z = f(x, y_0)$. We define the *partial derivative of* f as the ordinary derivative of $f(x, y_0)$ with respect to x at the point $x = x_0$.

Definition 17.1 (PARTIAL DERIVATIVE).
The *partial derivative of* f *with respect to* x *at the point* (x_0, y_0) is

$$\frac{\partial f}{\partial x}(x_0,y_0) = \frac{d}{dx} f(x,y_0)\Big|_{x=x_0} = \lim_{h\to 0} \frac{f(x_0+h,y_0) - f(x_0,y_0)}{h}$$

provided the limit exists.

There are several notations for a partial derivative:

$$\frac{\partial f}{\partial x}(x_0,y_0), \quad \frac{\partial f}{\partial x}\Big|_{(x_0,y_0)}, \quad f_x(x_0,y_0)$$

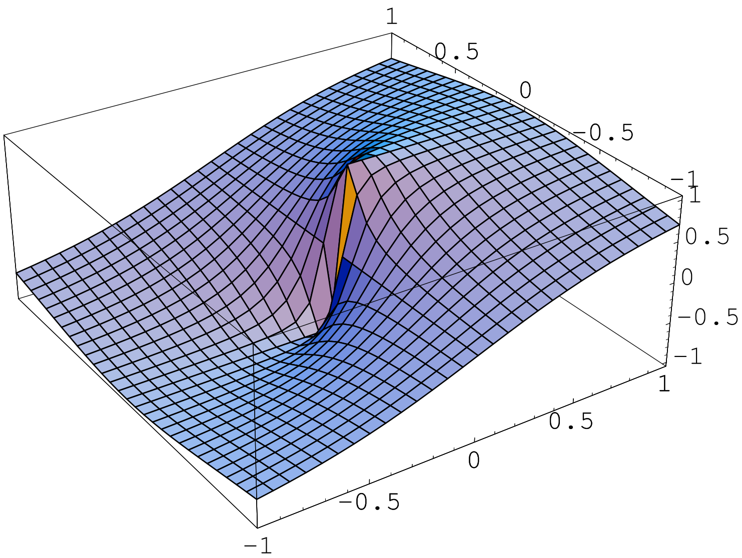

Figure 31: $f(x, y) = \frac{x}{\sqrt{x^2+y^2}}$

Remark 17.2. A function f can have partial derivatives with respect to both x and y at a point without being continuous there. This is different from functions of a single variable, where the existence of derivative implies continuity. The function

$$f(x, y) = \begin{cases} 0 & \text{if } xy \neq 0 \\ 1 & \text{if } xy = 0 \end{cases}$$

is not continuous at $(0, 0)$ while $f_x(0, 0) = f_y(0, 0) = 0$.

Example 17.3.
The partial derivatives of $f(x, y) = y \sin xy$ are

$$\frac{\partial f}{\partial x} = y \frac{\partial}{\partial x} \sin xy = y \cos xy \frac{\partial}{\partial x} xy = y^2 \cos xy$$
$$\frac{\partial f}{\partial y} = y \frac{\partial}{\partial y} \sin xy + \sin xy \frac{\partial}{\partial y} y = y \cos xy \frac{\partial}{\partial y} xy + \sin xy = xy \cos xy + \sin xy$$

Theorem 17.4 (EULER'S THEOREM).
If f and its partial derivatives f_x, f_y, f_{xy} and f_{yx} are defined on an open region containing a point (a, b) and are all continuous at (a, b), then

$$f_{xy}(a, b) = f_{yx}(a, b).$$

Proof. Omitted. □

Example 17.5.
If the continuity requirement is dropped, it is possible to construct functions for which mixed partials $f_{xy}(a, b)$ and $f_{yx}(a, b)$ are not equal:

$$f(x, y) = \begin{cases} \dfrac{xy(x^2 - y^2)}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0) \end{cases}$$

for which $f_{xy}(0, 0) = -1$ and $f_{yx}(0, 0) = +1$.

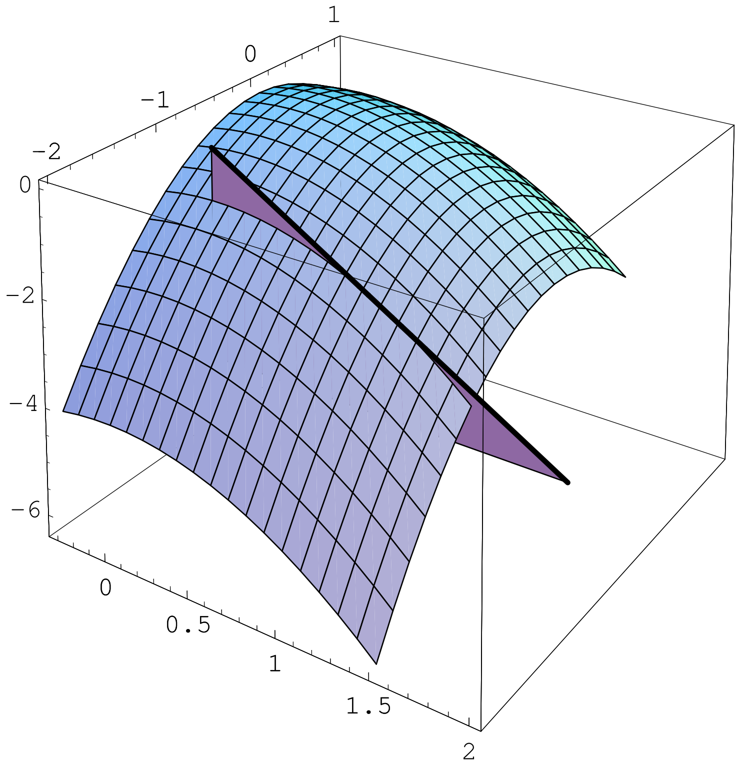

Figure 32: Partial derivative

18 Review of Derivatives of One Variable

Definition 18.1.
The limit of the *average rate of change*

$$\lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

is called the *derivative of f at x_0* provided it exits. The limit is denoted by $f'(x_0)$, $\frac{d}{dx} f(x_0)$, $\frac{df}{dx}(x_0)$ or $\left.\frac{df}{dx}\right|_{x=x_0}$. The function is said to be *differentiable at x_0* if the derivative exists.

Proposition 18.2.
A real number α is the derivative of f at $x = x_0$, that is $\alpha = f'(x_0)$ if and only if there is a real-valued function $\epsilon = \epsilon(x_0, \Delta x)$ of x_0 and Δx such that

$$f(x_0 + \Delta x) = f(x_0) + \alpha \Delta x + \epsilon \cdot \Delta x \quad \textit{and} \quad \epsilon \to 0 \textit{ as } \Delta x \to 0.$$

Proof. Suppose the existence of ϵ. Then

$$f(x_0 + \Delta x) - f(x_0) = \alpha \Delta x + \epsilon \cdot \Delta x \quad \text{hence} \quad \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = \alpha + \epsilon.$$

By $\lim_{\Delta x \to 0} \epsilon = 0$ we have

$$\lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = \alpha.$$

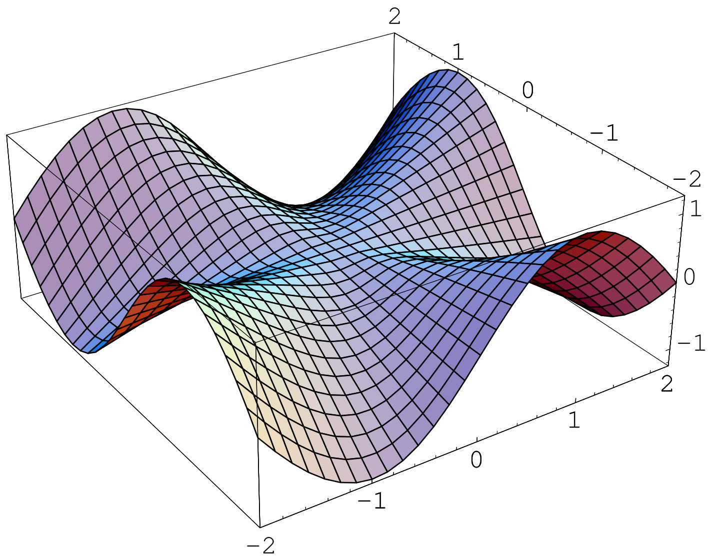

Figure 33: A function that does not meet the conditions of Euler's theorem

On the other hand, let ϵ be defined by

$$\begin{aligned}\epsilon &= \frac{f(x_0+\Delta x) - f(x_0) - \alpha\Delta x}{\Delta x} \\ &= \frac{f(x_0+\Delta x) - f(x_0)}{\Delta x} - \alpha \\ &= \frac{f(x_0+\Delta x) - f(x_0)}{\Delta x} - f'(x_0).\end{aligned}$$

Clearly this $\epsilon \to 0$ as $\Delta x \to 0$. □

Definition 18.3.
The *derivative* $f'(x_0)$ *of* f *at* x_0 is the real number such that

$$f(x_0+\Delta x) = f(x_0) + f'(x_0)\Delta x + \epsilon \cdot \Delta x$$

holds for some function ϵ satisfying $\epsilon \to 0$ as $\Delta x \to 0$.

Example 18.4.
$(f \cdot g)'(x_0) = f(x_0)g'(x_0) + f'(x_0)g(x_0)$

Proof.

$$\begin{aligned}(f\cdot g)(x_0+\Delta x) &= f(x_0+\Delta x)\cdot g(x_0+\Delta x) \\ &= (f(x_0) + f'(x_0)\Delta x + \epsilon_f\cdot\Delta x)\cdot(g(x_0) + g'(x_0)\Delta x + \epsilon_g\cdot\Delta x) \\ &= f(x_0)\cdot g(x_0) + f(x_0)\cdot g'(x_0)\Delta x + f'(x_0)\Delta x\cdot g(x_0) \\ &\quad + f'(x_0)\Delta x\cdot g'(x_0)\Delta x \\ &\quad + \epsilon_f\cdot\Delta x\cdot(g(x_0) + g'(x_0)\Delta x + \epsilon_g\cdot\Delta x) \\ &\quad + \epsilon_g\cdot\Delta x\cdot(f(x_0) + f'(x_0)\Delta x + \epsilon_f\cdot\Delta x) \\ &= f(x_0)\cdot g(x_0) + (f(x_0)\cdot g'(x_0) + f'(x_0)\cdot g(x_0))\,\Delta x \\ &\quad + (f'(x_0)\Delta x\cdot g'(x_0))\Delta x \\ &\quad + (\epsilon_f\cdot(g(x_0) + g'(x_0)\Delta x + \epsilon_g\cdot\Delta x))\Delta x \\ &\quad + (\epsilon_g\cdot(f(x_0) + f'(x_0)\Delta x + \epsilon_f\cdot\Delta x))\Delta x.\end{aligned}$$

We see that $(\epsilon_f \cdot (g(x_0) + g'(x_0)\Delta x + \epsilon_g \cdot \Delta x)) \to 0$ as $\Delta x \to 0$. In fact

$$\begin{aligned}
&\lim_{\Delta x\to 0} \epsilon_f \cdot (g(x_0) + g'(x_0)\Delta x + \epsilon_g \cdot \Delta x) \\
&= \lim_{\Delta x\to 0} \epsilon_f \cdot \lim_{\Delta x\to 0} (g(x_0) + g'(x_0)\Delta x + \epsilon_g \cdot \Delta x) \\
&= 0 \cdot (g(x_0) + 0 + 0) = 0.
\end{aligned}$$

□

19 Total Differential, Gradient

Definition 19.1.
The two-dimensional vector (α, β) is the *gradient of f at* (x_0, y_0) if there are functions ϵ_x and ϵ_y such that

$$\begin{aligned}
f(x_0 + \Delta x, y_0 + \Delta y) &= f(x_0, y_0) + \alpha\Delta x + \beta\Delta y + \epsilon_x \cdot \Delta x + \epsilon_y \cdot \Delta y \\
&= f(x_0, y_0) + \begin{pmatrix}\alpha, & \beta\end{pmatrix}\begin{pmatrix}\Delta x \\ \Delta y\end{pmatrix} + \begin{pmatrix}\epsilon_x, & \epsilon_y\end{pmatrix}\begin{pmatrix}\Delta x \\ \Delta y\end{pmatrix}.
\end{aligned}$$

and

$$\epsilon_x, \epsilon_y \to 0 \text{ as } \Delta x \to 0 \text{ and } \Delta y \to 0.$$

We say f *is differentiable at* $(x, y) = (x_0, y_0)$ if there is a gradient there. Since the gradient vector is unique, we denote it by $\nabla f(x_0, y_0)$.

Proposition 19.2.
If the partial derivatives f_x and f_y of f are defined in a neighborhood of (x_0, y_0) and continuous at (x_0, y_0), then f is differentiable at (x_0, y_0).

Proof. By definition

$$\begin{aligned}
f(x_0 + \Delta x, y_0 + \Delta y) &= f(x_0, y_0 + \Delta y) + f_x(x_0, y_0 + \Delta y)\Delta x + \epsilon_x\Delta x \\
&= (f(x_0, y_0) + f_y(x_0, y_0)\Delta y + \epsilon_y\Delta y) + f_x(x_0, y_0 + \Delta y)\Delta x + \epsilon_x\Delta x \\
&= f(x_0, y_0) + f_x(x_0, y_0 + \Delta y)\Delta x + f_y(x_0, y_0)\Delta y + \epsilon_x\Delta x + \epsilon_y\Delta y \\
&= f(x_0, y_0) + f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y \\
&\quad + \{f_x(x_0, y_0 + \Delta y) - f_x(x_0, y_0)\}\Delta x + \epsilon_x\Delta x + \epsilon_y\Delta y.
\end{aligned}$$

Since $f_x(x_0, y_0 + \Delta y)$ is continuous at (x_0, y_0) the difference $f_x(x_0, y_0 + \Delta y) - f_x(x_0, y_0)$ converges to zero as $\Delta y \to 0$. □

Proposition 19.3.
If the function is differentiable at (x_0, y_0), then

(1) it is continuous there

(2) $\nabla f(x_0, y_0) = (f_x(x_0, y_0), f_y(x_0, y_0))$

Definition 19.4.
The *linearization* of f at (x_0, y_0) where f is differentiable is the affine function

$$\begin{aligned}
L(x, y) &= f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) \\
&= f(x_0, y_0) + \nabla f(x_0, y_0)\begin{pmatrix}x - x_0 \\ y - y_0\end{pmatrix}
\end{aligned}$$

The change $f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$ in the linearization is called the *total differential* of f and written as

$$df = f_x(x_0, y_0)dx + f_y(x_0, y_0)dy.$$

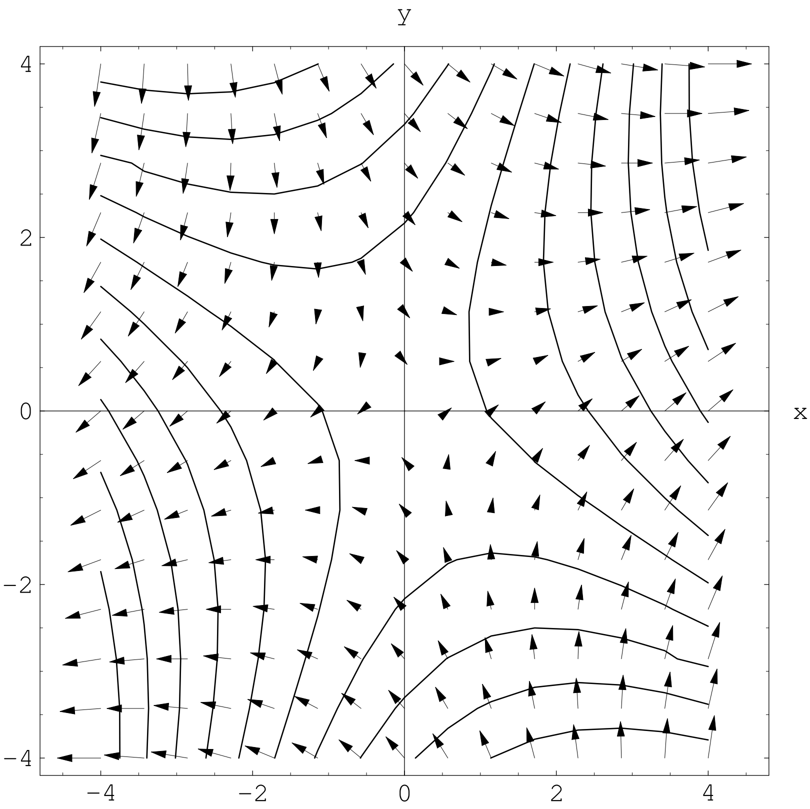

Figure 34: Gradient Field of $f(x,y) = 4x^2 + 6xy - 3y^2$

Example 19.5.
Your company is making right circular cylindrical storage tanks that are 25 m high with a radius of 5 m. How sensitive are the tank's volumes to small variation in height and radius?

The volume is a function of height h and radius r, i.e.

$$V = \pi r^2 h.$$

Then

$$dV = V_h(h_0, r_0)dh + V_r(h_0, r_0)dr = \pi r_0^2 dh + 2\pi r_0 h_0 dr = \pi(5^2 dh + 2 \times 5 \times 25 dr)$$

Definition 19.6 (MORE THAN TWO VARIABLES).
Let f be a real-valued function of n variables $(x_1, x_2, \ldots, x_n)$. The n-dimensional vector $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is the *gradient of f at* $(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$ if there are functions $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ such that

$$\begin{aligned} f(\bar{x}_1 + \Delta x_1, \ldots, \bar{x}_n + \Delta x_n) &= f(\bar{x}_1, \ldots, \bar{x}_n) + \alpha_1 \Delta x_1 + \cdots + \alpha_n \Delta x_n \\ &\quad + \epsilon_1 \cdot \Delta x_1 + \cdots + \epsilon_n \cdot \Delta x_n \\ &= f(\bar{x}_1, \ldots, \bar{x}_n) + \begin{pmatrix} \alpha_1 & \ldots & \alpha_n \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \vdots \\ \Delta x_n \end{pmatrix} \\ &\quad + \begin{pmatrix} \epsilon_1 & \ldots & \epsilon_n \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \vdots \\ \Delta x_n \end{pmatrix}. \end{aligned}$$

and

$$\epsilon_1, \ldots, \epsilon_n \to 0 \text{ as } \Delta x_1, \ldots, \Delta x_n \to 0.$$

We say f *is differentiable at* $(\bar{x}_1, \ldots, \bar{x}_n)$ if there is a gradient there. Since the gradient vector is unique, we denote it by $\nabla f(\bar{x}_1, \ldots, \bar{x}_n)$.

The *linearization* of f at $(\bar{x}_1, \ldots, \bar{x}_n)$ where f is differentiable is the affine function

$$L(x_1, \ldots, x_n) = f(\bar{x}_1, \ldots, \bar{x}_n) + \nabla f(\bar{x}_1, \ldots, \bar{x}_n) \begin{pmatrix} x_1 - \bar{x}_1 \\ \vdots \\ x_n - \bar{x}_n \end{pmatrix}.$$

The *total differential* is

$$\begin{aligned} df &= f_{x_1}(\bar{x}_1, \ldots, \bar{x}_n)dx_1 + \cdots + f_{x_n}(\bar{x}_1, \ldots, \bar{x}_n)dx_n \\ &= \nabla f(\bar{x}_1, \ldots, \bar{x}_n) \begin{pmatrix} dx_1 \\ \vdots \\ dx_n \end{pmatrix}. \end{aligned}$$

20 Rules

Proposition 20.1.
Suppose f and g are functions of two variables and differentiable at (x_0, y_0) and h be a real valued function of one variable which is differentiable at $f(x_0, y_0)$. Then

(1) $f \pm g$ is differentiable there and $\nabla(f \pm g)(x_0, y_0) = \nabla f(x_0, y_0) \pm \nabla g(x_0, y_0)$

(2) $f \cdot g$ is differentiable there and $\nabla(f \cdot g)(x_0, y_0) = f(x_0, y_0)\nabla g(x_0, y_0) + g(x_0, y_0)\nabla f(x_0, y_0)$

(3) $\frac{1}{f}$ is differentiable there and $\nabla\left(\frac{1}{f}\right)(x_0, y_0) = \frac{-1}{f(x_0, y_0)^2}\nabla f(x_0, y_0)$ provided $f(x_0, y_0) \neq 0$

(4) The composition hf is differentiable there and

$$\nabla(hf)(x_0, y_0) = h'(f(x_0, y_0))\nabla f(x_0, y_0).$$

Proof.

$$\begin{aligned}(hf)(x_0 + \Delta x, y_0 + \Delta y) &= h(f(x_0 + \Delta x, y_0 + \Delta y)) \\ &= h\left(f(x_0, y_0) + \nabla f(x_0, y_0)\begin{pmatrix}\Delta x \\ \Delta y\end{pmatrix} + \begin{pmatrix}\epsilon_x & \epsilon_y\end{pmatrix}\begin{pmatrix}\Delta x \\ \Delta y\end{pmatrix}\right) \\ &= h\left(f(x_0, y_0)\right) + h'(f(x_0, y_0))\left(\nabla f(x_0, y_0)\begin{pmatrix}\Delta x \\ \Delta y\end{pmatrix} + \begin{pmatrix}\epsilon_x & \epsilon_y\end{pmatrix}\begin{pmatrix}\Delta x \\ \Delta y\end{pmatrix}\right) \\ &\quad + \epsilon \cdot \left(\nabla f(x_0, y_0)\begin{pmatrix}\Delta x \\ \Delta y\end{pmatrix} + \begin{pmatrix}\epsilon_x & \epsilon_y\end{pmatrix}\begin{pmatrix}\Delta x \\ \Delta y\end{pmatrix}\right) \\ &= (hf)(x_0, y_0) + h'(f(x_0, y_0))\nabla f(x_0, y_0)\begin{pmatrix}\Delta x \\ \Delta y\end{pmatrix} \\ &\quad + \begin{pmatrix}\bar{\epsilon}_x & \bar{\epsilon}_y\end{pmatrix}\begin{pmatrix}\Delta x \\ \Delta y\end{pmatrix}\end{aligned}$$

□

21 Directional Derivative

Definition 21.1 (DIRECTIONAL DERIVATIVE).
For the direction vector (u, v) with $\|(u, v)\| = 1$, the *derivative of f at (x_0, y_0) in the direction (u, v)* is

$$\nabla_{(u,v)} f(x_0, y_0) = \lim_{s \to 0} \frac{f(x_0 + su, y_0 + sv) - f(x_0, y_0)}{s}$$

provided the limit exists.

Proposition 21.2.

$$\nabla_{(u,v)} f(x_0, y_0) = \nabla f(x_0, y_0)\begin{pmatrix}u \\ v\end{pmatrix}$$

Proof. Let $h(s) = f(x_0 + su, y_0 + sv)$, then $\nabla_{(u,v)} f(x_0, y_0) = h'(0)$ by definition, which is given by the chain rule

$$h'(0) = f_x(x_0, y_0)\frac{d(su)}{ds} + f_y(x_0, y_0)\frac{d(sv)}{ds} = \nabla f(x_0, y_0)\begin{pmatrix}u \\ v\end{pmatrix}.$$

□

22 Tangent to Level Curve

Suppose a differentiable function $f(x, y)$ has a constant value, say c, along a smooth curve $(g(t), h(t))$ parameterized by t, i.e. $f(g(t), h(t)) = c$. Differentiating both sides with respect to t leads to

$$\frac{d}{dt} f(g(t), h(t)) = \frac{d}{dt} c = 0,$$

which is by the chain rule

$$f_x(g(t), h(t))g'(t) + f_y(g(t), h(t))h'(t) = 0 \quad \text{or} \quad \nabla f(g(t), h(t))\begin{pmatrix}g'(t) \\ h'(t)\end{pmatrix} = 0.$$

Since $(g'(t), h'(t))$ is the tangent vector of the smooth curve, this equation says ∇f is normal to the tangent vector. Therefore the tangent line to the level curve at (x_0, y_0) is given by

$$\nabla f(x_0, y_0)\begin{pmatrix}x - x_0 \\ y - y_0\end{pmatrix} = 0.$$

Example 22.1.
Let $f(x,y) = \frac{x^2}{4} + y^2$ and consider its level curve of level 2. Find the tangent line of this level curve at $(x,y) = (-2,1)$.

$$\nabla f(-2,1) = (\frac{x}{2}, 2y)|_{(x,y)=(-2,1)} = (-1,2).$$

Then the tangent line is

$$(-1,2)\begin{pmatrix} x-(-2) \\ y-(1) \end{pmatrix} = (-1)(x+2) + 2(y-1) = 0.$$

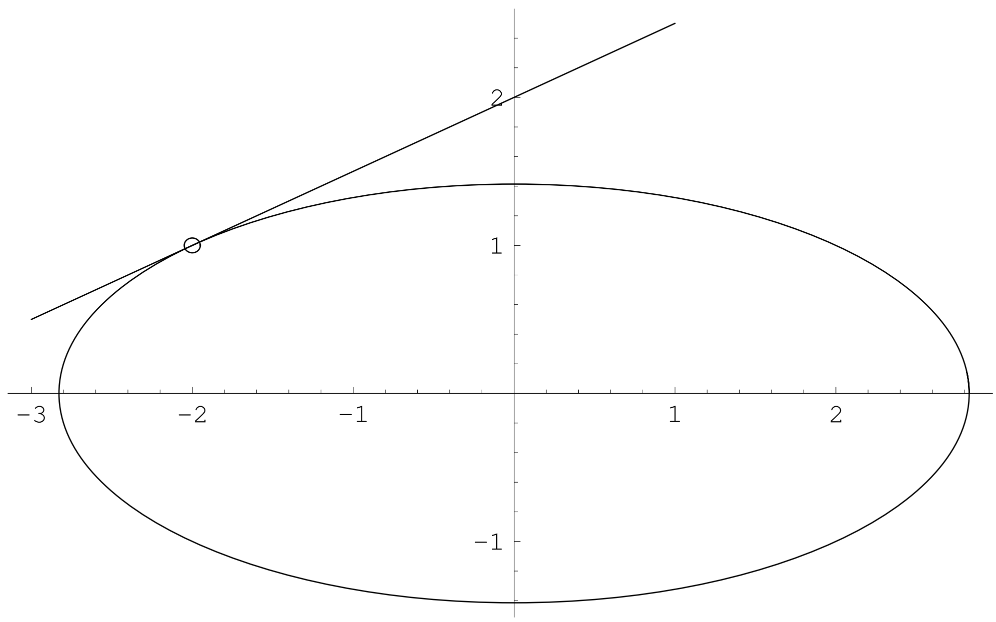

Figure 35: Eclipse $\frac{x^2}{4} + y^2 = 2$ and the tangent line at $(-2,1)$

23 Extreme Values and Saddle Points

Definition 23.1.
Let f be defined on a region R containing (a,b). Then

1. f takes a *local maximum at* (a,b) if there is a $\delta > 0$ such that $f(a,b) \geq f(x,y)$ for all points in $R \cap B((a,b),\delta)$,

2. f takes a *local minimum at* (a,b) if there is a $\delta > 0$ such that $f(a,b) \leq f(x,y)$ for all points in $R \cap B((a,b),\delta)$.

Proposition 23.2.
Suppose f has partial derivatives at an interior point (a,b) of its domain. If f has a local maximum or minimum at (a,b), $f_x(a,b) = f_y(a,b) = 0$. Therefore if in addition f is differentiable, $\nabla f(a,b) = (0,0)$.

Proof. We show $f_x(a,b) = 0$.

1. $x = a$ is an interior point of the domain of $f(x,b)$,

2. $f(x,b)$ is differentiable at $x = a$ and its derivative is $f_x(a,b)$,

3. $f(x,b)$ has a local maximum or minimum at $x = a$,

4. then $f_x(a,b) = 0$.

□

Definition 23.3.
An interior point of the domain of f

1. where both f_x and f_y are zero or
2. both of f_x and f_y do not exist

is a *critical point* of f.

Definition 23.4.
A differentiable function f defined on R has a *saddle point* at (a,b) if

1. (a,b) is a critical point of f,
2. for any $\delta > 0$
   (i) there is a point $(x,y) \in R \cap B((a,b),\delta)$ such that $f(x,y) > f(a,b)$ and
   (ii) there is a point $(x',y') \in R \cap B((a,b),\delta)$ such that $f(x',y') < f(a,b)$.

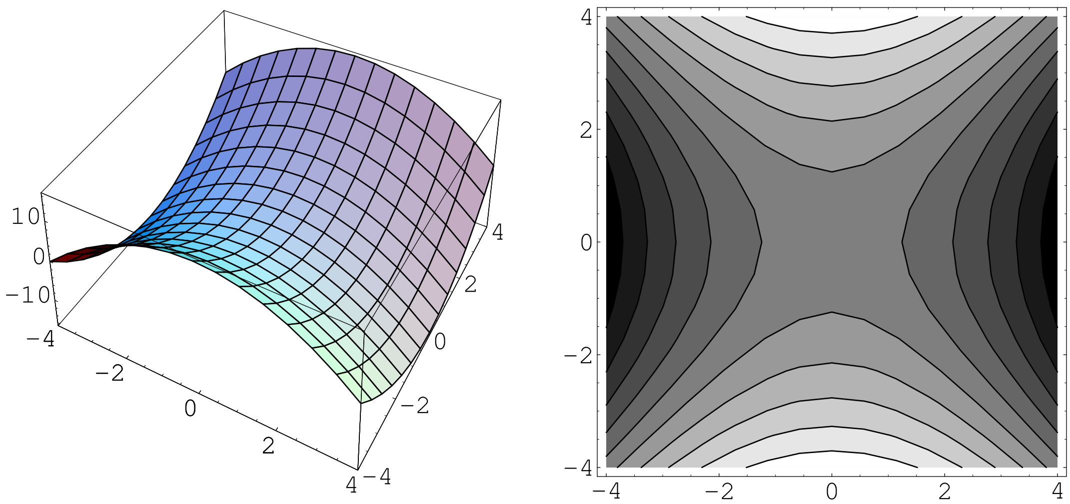

Figure 36: Saddle point

24 Lagrange Multipliers

How can we find the extreme values of a function whose domain is constrained to lie within some particular subset of the plane? Lagrange developed a method in 1755.

Example 24.1.
Find the point closest to the origin on the plane $2x + y - z - 5 = 0$.

$$\left|\begin{array}{lll} \text{minimize} & \sqrt{x^2+y^2+z^2} \quad \text{or} \quad x^2+y^2+z^2 \\ \text{subject to} & 2x+y-z-5=0 \end{array}\right.$$

If we regard x and y as the independent variables and write $z = 2x + y - 5$, the problem reduces to finding a point where

$$h(x, y) = x^2 + y^2 + (2x + y - 5)^2$$

has its minimum value. Any minimum must occur at points where

$$h_x(x, y) = 0 \quad \text{and} \quad h_y(x, y) = 0,$$

i.e.

$$10x + 4y = 20, \quad 4x + 4y = 10.$$

Solving this equations we have $x = 5/3, y = 5/6$.

Example 24.2.
Find the point closest to the origin on the cylinder $x^2 - z^2 - 1 = 0$.

$$\left| \begin{array}{ll} \text{minimize} & x^2 + y^2 + z^2 \\ \text{subject to} & x^2 - z^2 - 1 = 0 \end{array} \right.$$

Method #1
If we regard x and y as the independent variables and write $z^2 = x^2 - 1$, the problem reduces to finding a point where

$$h(x, y) = x^2 + y^2 + (x^2 - 1) = 2x^2 + y^2 - 1$$

has its minimum value. Any minimum must occur at points where

$$h_x(x, y) = 4x = 0 \quad \text{and} \quad h_y(x, y) = 2y = 0,$$

that is $x = y = 0$. But this point is not on the cylinder. What went wrong? Look at the constraint carefully, we see that $x^2 - 1 = z^2 \geq 0$. Then the variable x cannot move freely but is restricted to the "shadow" of the cylinder, i.e. $x \leq -1$ and $x \geq 1$.

We can avoid this trouble if we treat y and z as independent variables and minimize

$$k(y, z) = (z^2 + 1) + y^2 + z^2 = 1 + y^2 + 2z^2$$

to obtain

$$k_y(y, z) = 2y = 0, \quad k_z(y, z) = 4z = 0. \quad \text{i.e.} \quad y = z = 0,$$

which leads to $x^2 = z^2 + 1 = 1$ or $x = \pm 1$.

Method #2
Picture a small sphere centered at the origin expanding till it touches the cylinder. At each point of contact, the cylinder and sphere have the same tangent plane and also normal vector. Therefore if we set

$$f(x, y, z) = x^2 + y^2 + z^2 - a^2, \quad g(x, y, z) = x^2 - z^2 - 1$$

the gradients ∇f and ∇g are parallel at points of contact. That is

$$\nabla f(x, y, z) = \lambda \nabla g(x, y, z)$$

holds for some scalar λ. Then

$$(2x, 2y, 2z) = \lambda(2x, 0, -2z).$$

For what value of λ will a point (x, y, z) satisfy this equation as well as $z^2 = x^2 - 1$? Using the fact that no point on the cylinder has zero x-coordinate, we obtain $\lambda = 1$. For this value of λ $2z = \lambda(-2z)$ reduces to $z = 0$ and $2y = \lambda 0$ reduces to $y = 0$. Thus we conclude the points we seek have coordinates $(x, 0, 0)$. For this point to sit on the cylinder, $x = \pm 1$.

Proposition 24.3.
Suppose $f(x,y,z)$ and $g(x,y,z)$ are differentiable. To find the local maximum and minimum values of f subject to the constraint $g(x,y,z)=0$, i.e.

$$\left|\begin{array}{ll} \textit{maximize/minimize} & f(x,y,z) \\ \textit{subject to} & g(x,y,z)=0, \end{array}\right.$$

find the values of x,y,z and λ satisfying

$$\nabla f(x,y,z)=\lambda\nabla g(x,y,z) \quad \textit{and} \quad g(x,y,z)=0.$$

When the constraints are

$$g_1(x,y,z)=0, \quad g_2(x,y,z)=0,$$

find x,y,z,λ_1 and λ_2 satisfying

$$\nabla f(x,y,z)=\lambda_1\nabla g_1(x,y,z)+\lambda_2\nabla g_2(x,y,z), \quad g_1(x,y,z)=0$$

and

$$g_2(x,y,z)=0.$$

Remark 24.4. This proposition is guaranteed by the *implicit function theorem*, which we skip in this note.

Example 24.5.

$$\left|\begin{array}{ll} \text{minimize/maximize} & x^2+y^2+z^2 \\ \text{subject to} & x^2+y^2-1=0 \\ & x+y+z-1=0 \end{array}\right.$$

$$f(x,y,z)=x^2+y^2+z^2, \quad g_1(x,y,z)=x^2+y^2-1, \quad g_2(x,y,z)=x+y+z-1$$

Then the system we have to solve is

$$\begin{aligned} &(2x,2y,2z)=\lambda_1(2x,2y,0)+\lambda_2(1,1,1) \\ &x^2+y^2-1=0 \\ &x+y+z-1=0. \end{aligned}$$

From these equations we have

$$(1-\lambda_1)x=z, \quad (1-\lambda_1)y=z.$$

These are satisfied if either $\lambda_1=1$ and $z=0$ or $\lambda_1\neq 1$ and $x=y=z/(1-\lambda_1)$.
When $z=0$ we find two points $(1,0,0)$ and $(0,1,0)$.
When $x=y$ we obtain

$$x^2+x^2-1=0, \quad x+x+z-1=0,$$

which give

$$x=\pm\frac{\sqrt{2}}{2}, \quad z=1\mp\sqrt{2}.$$

The corresponding points on the ellipse are

$$\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2},1-\sqrt{2}\right) \quad \text{and} \quad \left(-\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2},1+\sqrt{2}\right).$$

You need to be careful to decide which is farthest from the origin.