1
Functions of One Variable

1.1
Limits

It is often said that most mathematical errors, which get published, follow the word "clearly" and involve the improper interchange of two limits. In simple terms, a "limit" is the number that a function or sequence "approaches" as the input or index approaches some value. For example, we will say that the sequence $x_n = \frac{1}{n}$ approaches the limit 0 as n moves to infinity Or, in other words, we can make x_n arbitrarily small by choosing n big enough. We often write this as

$$\lim_{n\to\infty} x_n = 0$$

We can also take the limit of a function, for example, if $f(x) = x^2$ then

$$\lim_{x\to 2} f(x) = 4$$

A sequence of numbers x_n is said to converge to a limit x if we can make the difference $|x - x_n|$ arbitrarily small by making n big enough. If such a limit point does not exist, then we state that the sequence diverges. For example, the sequence of integers

$$x_n = 1, 2, 3, \ldots$$

is unbounded as $n \to \infty$, while the sequence

$$x_n = 1 + (-1)^n$$

oscillates and never settles down to a limit. More formally, we state

Definition 1.1. *Let f be a function defined on a real interval I then the limit as $x \to a$ exists if there exists a number l such that given a number $\epsilon > 0$ no matter how small, we can find a number $\delta > 0$, where for all $x \epsilon I$ satisfying*

$$|x - a| < \delta$$

we have

$$|f(x) - l| < \epsilon$$

A first Course in Mathematical Physics, First Edition. Colm T Whelan.

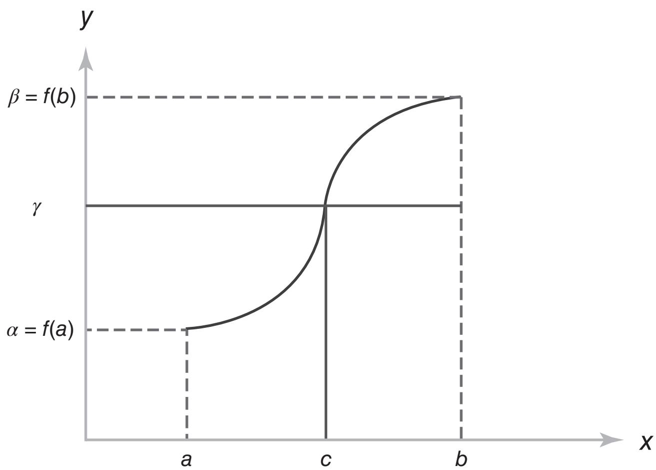

Figure 1.1 If $y = f(x)$ is a continuous function on $[a, b]$ if we pick any value, γ, that is between the value of $\alpha = f(a)$ and the value of $\beta = f(b)$ and draw a line straight out from this point, the line will hit the graph in at least one point with an x value between a and b.

Notice that we do not necessarily let x ever reach a but only get infinitesimally close to it. If in fact $f(a) = l$, then we state that the function is continuous. Intuitively, a function that is continuous on some interval $[a, b]$ will take on all values between $f(a)$ and $f(b)$ (Figure 1.1). For a more formal discussion see [1]. An intuitively obvious result is the intermediate value theorem.

Theorem 1.1. *Let f be a continuous function on a closed interval $[a, b]$ $\alpha = f(a), \beta = f(b)$. If γ is a number such that $\alpha < \gamma < \beta$, then there exists a number c such that*

$$f(c) = \gamma$$

Proof: For a formal proof see for example [1] ■

Consider the sequence of partial sums

$$S_n = \sum_{j=1}^{n} x_n \tag{1.1}$$

if the sequence of partial sums converges to some limit S as $n \to \infty$ then we say that the infinite series $\sum_{j=0}^{\infty} x_j$ is convergent.

Example 1.1. *The geometric series*

Let

$$s_n = \sum_{j=1}^{n} ax^j$$

then

$$\begin{aligned}(1-x)s_n &= \sum_{j=1}^{n} ax^j - ax^{j+1} \\ &= a - ax + ax - ax^2 + ax^2 + \cdots + ax^{n-1} - ax^n + ax^n - ax^{n+1} \\ &= a(1-x^{n+1})\end{aligned}$$

hence

$$s_n = \frac{a(1-x^{n+1})}{1-x}$$

Clearly, therefore, if $|x| < 1$ *the series converges. Its value being given by*

$$\sum_{j=1}^{\infty} ax^j = \frac{a}{1-x}$$

if $|x| \geq 1$, *the series diverges.*

1.2 Elementary Calculus

Assume that we are observing an object moving in one dimension. We measure its position to be x_0 at time $t = t_0$ and $x_0 + \Delta x$ at time $t = t_0 + \Delta t$, thus its average speed is

$$\bar{v} = \frac{(x_0 + \Delta x) - x_0}{(t_0 + \Delta t) - t_0} = \frac{\Delta x}{\Delta t} \tag{1.2}$$

Of course, this is only an average value the object could accelerate and decelerate during the time interval; if we need to know its speed at any given point, then we must shorten the time interval, and to know the "instantaneous" speed at a time $t = t_0$, we need to let Δt lead to zero, that is,

$$v(t) = \lim_{\Delta t \to 0} \frac{x_0 + \Delta x - x_0}{t_0 + \Delta t - t_0} \tag{1.3}$$

This motivates us to define the derivative of a function.

Definition 1.2. *If* f *is only a function of* x, *then the first derivative of* f *at* x *is defined to be*

$$\frac{df(x)}{dx} \equiv \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \tag{1.4}$$

If this limit exists, then the function is said to be differentiable. The function f *is said to be continuously differentiable if the derivative* $f'(x)$ *exists and is itself a continuous function.*

Frequently, we use the notation $f'(a)$ as a shorthand, that is,

$$f'(a) = \frac{df(x)}{dx}|_{x=a}$$

Example 1.2. *If $f(x) = x$, then*

$$f'(x) = \frac{df(x)}{dx} = \lim_{h\to 0} \frac{x+h-x}{h} = 1$$

If $f(x) = c$, where c is a constant, then $f(x+h) - f(x) = 0$ for all h, consequently $f'(x) = 0$. A partial converse to this result is as follows. If $f('x) = 0$ on some interval I, then $f(x) = c$ on I, where c is a constant. This is a consequence of the intermediate value theorem; see Problem 1.1. Clearly if

$$f'(x) = g'(x) \text{ on } I \tag{1.5}$$

then

$$f(x) = g(x) + c \text{ on } I \tag{1.6}$$

where c is a constant.

1.2.1 Differentiation Products and Quotients

Assuming $f(x)$ can be written as the product of two functions, for example,

$$f(x) = u(x)v(x)$$

then

$$\frac{df(x)}{dx} = \lim_{h\to 0} \frac{u(x+h)v(x+h) - u(x)v(x)}{h} \tag{1.7}$$

We may rewrite the numerator in (1.7) as

$$u(x+h)[v(x+h) - v(x)] + v(x)[u(x+h) - u(x)]$$

in the limit as $h \to 0$ $u(x+h) \to u(x)$ and

$$\lim_{h\to 0} \frac{u(x+h) - u(x)}{h} = \frac{du(x)}{dx}$$

$$\lim_{h\to 0} \frac{v(x+h) - v(x)}{h} = \frac{dv(x)}{dx}$$

it immediately follows that

$$\frac{d(u(x)v(x))}{dx} = u(x)\frac{dv(x)}{dx} + v(x)\frac{du(x)}{dx} \tag{1.8}$$

It is also possible to show, Problem 1.10, that

$$\frac{d}{dx}\frac{f(x)}{g(x)} = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2} \tag{1.9}$$

Lemma 1.1.

$$\frac{dx^N}{dx} = Nx^{N-1}$$

Proof: If $N = 1$, clearly true since $\frac{dx}{dx} = 1$; assume to be true for all $N \leq N_0$, consider

$$f(x) = x^{N_0+1} = xx^{N_0}$$

then from (1.8)

$$\frac{df(x)}{dx} = x^{N_0}\frac{dx}{dx} + x\frac{dx^{N_0}}{dx} \tag{1.10}$$

Now, by assumption

$$\frac{dx^{N_0}}{dx} = N_0x^{N_0-1}$$

and as we have already seen

$$\frac{dx}{dx} = 1$$

hence

$$\frac{df(x)}{dx} = x^{N_0} + xN_0x^{N_0-1} = x^{N_0}[N_0 + 1] \tag{1.11}$$

Hence, by principle of induction, true for all integers. ■

1.2.2 Chain Rule

Assume that

$$f(x) = u(v(x))$$

For example,

$$\begin{aligned} v(x) &= \quad x^2 \\ u(y) &= \sqrt{1-y} \\ \Rightarrow f(x) &= \sqrt{1-x^2} \end{aligned}$$

then for such a function:

Lemma 1.2. *If v is differentiable at the point x and that u is differentiable at the point $y = v(x)$, then*

$$\frac{df(x)}{dx} = u'(v(x))v'(x) \tag{1.12}$$

or in other words

$$\frac{df}{dx} = \frac{df}{du}\frac{du}{dx}$$

Proof:

$$\frac{u(v(x+h)-u(v(x))}{h} = \frac{u(v(x+h)-u(v(x))}{v(x+h)-v(x)}\frac{v(x+h)-v(x)}{h} \tag{1.13}$$

We can now take the limit as $h \to 0$ and we have the result. In fact, to be more rigorous, we should worry about the possibility of $v(x+h)-v(x)$ passing through 0. For a treatment where this problem is explicitly dealt with see [2]. ■

Example 1.3. *Newton's second law can be written as*

$$\begin{aligned} F = ma &= m\frac{dv}{dt} \\ &= m\frac{dx}{dt}\frac{dv}{dx} = mv\frac{dv}{dx} \\ &= \frac{d[\frac{1}{2}mv^2]}{dx} \end{aligned} \tag{1.14}$$

Thus, force can be defined as the rate of mass times acceleration or as the rate of change of the kinetic energy with distance

1.2.3
Inverse Functions

Consider the functions shown in Figure 1.2. Both are continuous but for $f_2(x)$ the equation

$$y = f_2(x)$$

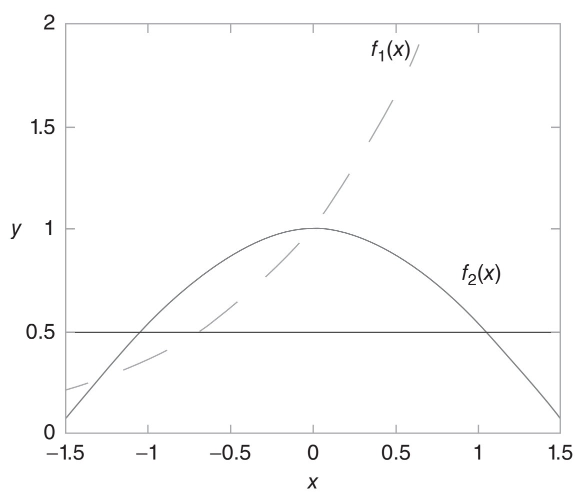

Figure 1.2 Over the range shown, the function $f_1(x)$ is invertible, $f_2(x)$ is not.

does not have a unique solution for $y = \frac{1}{2}$, while the equivalent equation for f_1 will have such a solution. The difference between the two functions is that f_1 is strictly increasing over the entire interval but f_2 is not.

Definition 1.3. *If f is a continuous strictly increasing function on $[a, b]$, with $\alpha = f(a), \beta = f(b)$, then from the intermediate value theorem, Theorem 1.1, we know that the set*

$$\{y = f(x) | a \leq x \leq b\}$$

forms the interval $[\alpha, \beta]$. We may define a function g

$$\begin{aligned} g\,:\,[\alpha, \beta] &\to [a, b] \\ g(f(x)) &= x \\ f(g(y)) &= y \end{aligned}$$

It is clear that we could just as well have constructed an inverse for a strictly decreasing function. The only time we will have a problem is when $x_1 \neq x_2$ but $f(x_1) = f(x_2)$. Usually, we write g as f^{-1}. Even for f_2, we can define an inverse if we agree to only look at the intervals $-1.5 \leq x < 0$ and $0 < x \leq 1.5$ separately. We can thus talk about a local inverse, that is, given a point x_0 if we can find an interval around it for which the function f is strictly increasing or decreasing, then we can find an inverse valid in this region. We know that (see Problem 1.3) a function is strictly increasing or decreasing on an interval once its derivative does not change sign; so if the derivative is continuous and we are not at a point x_0, where $f'(x_0) = 0$, then we can always find an interval, however, small so that the function is locally invertible. More formally, we may state the inverse function theorem.

Theorem 1.2. *For functions of a single real variable, if f is a continuously differentiable function with nonzero derivative at the point x_0, then f is invertible in a neighborhood of x_0, the inverse is continuously differentiable and*

$$\frac{df^{-1}(y)}{dy} = \frac{1}{f'(x)}$$

where $x = f^{-1}(y)$.

Proof: If f has a nonzero derivative at x_0, then it follows that there is a interval around x_0 where it is either increasing or decreasing then (Problem 1.4), f^{-1} is continuous. Let $\alpha < y_0 < \beta; y_0 = f(x_0), y = f(x)$, then

$$\frac{g(y) - g(y_0)}{y - y_0} = \frac{x - x_0}{f(x) - f(x_0)} = \frac{1}{\frac{f(x) - f(x_0)}{x - x_0}}$$

since f^{-1} is continuous the result follows. ■

1.3 Integration

There are number of equivalent ways of looking at integrals. Perhaps the most intuitive is to consider

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

as the area in the plane bounded by the curves $y = f(x), y = 0, y = f(a), y = f(b)$. Conventionally, we often describe this quantity as the area under the curve $y = f(x)$; see Figure 1.3. As a first approximation, we could simply assume that the function $y = f(x)$ could be approximated by its initial value $y = f(a)$ over the entire range, and we would have

$$I \approx f(a)(b - a)$$

See Figure 1.3(a). Now, we clearly lose some area by this approximation. We can improve it by taking a point c, with $a \le c \le b$ and approximating the integral by two rectangles of area $f(a)(c - a)$ and $f(b)(b - c)$.

We can continue this process by adding more and more subintervals. If

$$a = \chi_0 < \chi_1 < \chi_2 < \cdots < \chi_n = b \tag{1.15}$$

then we can approximate

$$I \approx \sum_{j=1}^{n} f(x_i)(\chi_i - \chi_{i-1}) \tag{1.16}$$

where

$$\chi_{i-1} \le x_i \le \chi_i$$

If we make these intervals arbitrarily small, that is, let $n \to \infty$, then we should get an accurate measure of the area under the curve. This prompts the following definition.

Definition 1.4. *The integral from a to b of the function f is given by*

$$I = \int_a^b f(x) = \lim_{n\to\infty} \sum_{j=1}^{n} f(x_i)(\chi_i - \chi_{i-1}) \tag{1.17}$$

where

$$a = \chi_0 < \chi_1 < \chi_2 < \cdots < \chi_{i-1} \le x_i \le \chi_i < \cdots < \chi_n = b$$

We can use the intermediate value theorem to establish the following theorem.

Theorem 1.3. *Mean value theorem for integrals. Let f be continuous on* $[a, b]$, *then there exists a c* ϵ (a, b) *s.t.*

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

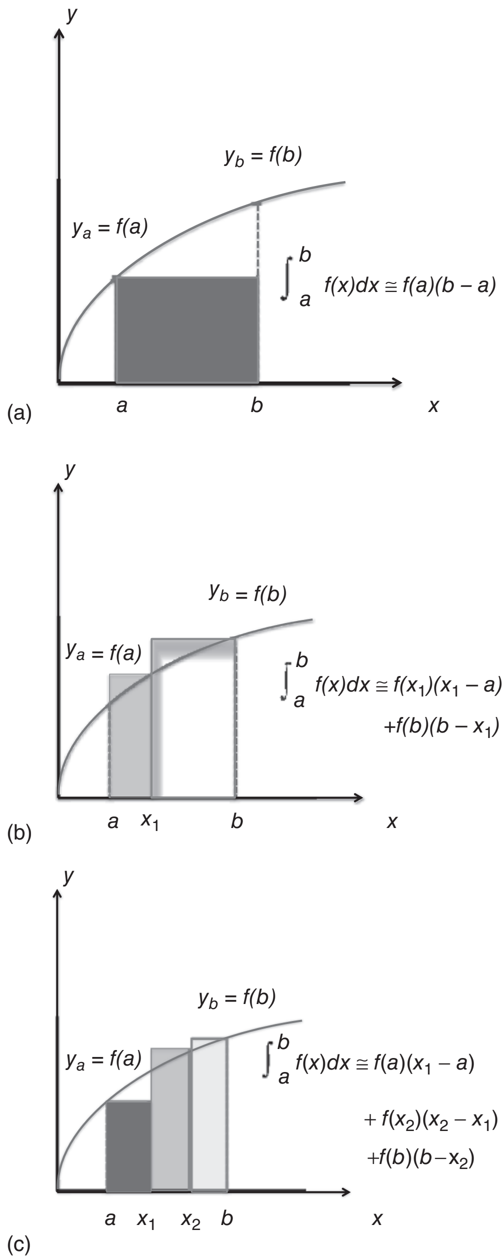

Figure 1.3 (a) Approximating the value of the integral as $f(a)(b-a)$, (b) picking a point x_1 where $a \le x_1 \le b$ and approximating integral as $\int_a^b f(x)dx \approx f(x_1)(x_1 - a) + f(x_1)(b - x_1)$, and (c) picking another point x_2 where $x_1 \le x_1 \le b$ and approximating integral as $\int_a^b f(x)dx \approx f(x_1)(x_1 - a) + f(x_2)(x_2 - x_1) + f(b)(b - x_2)$.

Proof: From Definition 1.2, if m is the minimum value of f on $[a, b]$ and M its maximum, then

$$m(b-a) \le \int_a^b f(x)dx \le M(b-a)$$

Hence, by the intermediate value theorem, there exists $c \in [a, b]$ s.t.

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

■

We can define a function

$$F(x) = \int_a^x f(y)dy \tag{1.18}$$

Theorem 1.4.

$$\frac{dF(x)}{dx} = f(x)$$

Proof:

$$F(x+h) = \int_a^{x+h} f(y)dy$$

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

$$F(x+h) - F(x) = \int_x^{x+h} f(y)dy \tag{1.19}$$

Now, if h is sufficiently small, we can take $f(y) = f(x)$ over the entire interval and

$$\frac{F(x+h) - F(x)}{h} = \frac{x+h-x}{h} f(x) = f(x) \tag{1.20}$$

take the limit $h \to 0$ and the result follows immediately. ■

We note that the constant a is entirely arbitrary. Theorem 1.4 is rather grandly known as the fundamental theorem of calculus, and it essentially states that integration is the inverse process to differentiation. It has an important corollary.

Corollary 1.1. *Assume that g is continuously differentiable function that maps the real interval $[a, b]$ onto the real interval I and that f is a continuous function that maps I into $\mathbb{R}$. Then,*

$$\int_{g(a)}^{g(b)} f(x) = \int_a^b f(g(t))g'(t)dt$$

Proof: The function $f(g(t))g'(t)$ is continuous on $[a, b]$ just as f is; therefore, both $\int_{g(a)}^{g(b)} f(x)dx$ and $\int_a^b f(g(t))g'(t)dt$ exist. All we have to show is that they are, in fact, equal. Since f is continuous, the fundamental theorem tells us that it has a differentiable “anti-derivative” F. If we apply the chain rule to $h(t) = F(g(t))$, then

$$\begin{aligned}
h'(t) &= F'(g(t))g'(t) \\
&= f(g(t)g'(t) \\
\Rightarrow \int_a^b f(g(t))g'(t)dt &= \int_a^b h'(t)dt \\
&= h(b) - h(a) \\
&= F(g(b)) - F(g(a)) \\
&= \int_{g(b)}^{g(a)} f(x)dx
\end{aligned}$$

Note that we have used the fundamental theorem twice more. ■

We remark that the result looks very much neater if we use the $\frac{dy}{dx}$ notation, that is, if we write

$$\begin{aligned}
\frac{dx}{dt} &= g'(t) \\
dx &= g'(t)dt
\end{aligned}$$

then our result becomes

$$\int_a^b f(x)dx = \int_{g(a)}^{g(b)} f(g(t))\frac{dg(t)}{dt}dt \tag{1.21}$$

If we rewrite (1.8) in the form

$$\frac{d(u(x)v(x))}{dx} = u(x)v'(x) + v(x)u'(x) \tag{1.22}$$

and integrate, we have

$$u(x)v(x) = \int u(x)v'(x)dx + \int v(x)u'(x)dx$$

$$\int u(x)v'(x)dx = u(x)v(x) - \int u'(x)v(x)dx \tag{1.23}$$

The result (1.23) gives us the useful method of evaluating integrals, known as integration by parts.We have seen here that the derivative of a constant, any constant, is zero. Thus, if we know that

$$\frac{dF(x)}{dx} = f(x)$$

then all we can state is that

$$F(x) = \int_a^x f(y)dy + c \tag{1.24}$$

where c is some constant. From our definition of the integral (Definition 1.2), we see that

$$\int_a^a f(x)dx = 0 \tag{1.25}$$

hence

$$F(a) = c$$

We may write (1.24) as

$$F(x) - F(a) = \int_a^x f(y)dy \tag{1.26}$$

1.4
The Binomial Expansion

Definition 1.5. *If n is a nonzero positive integer, then we define*

$$n! = n \cdot (n-1) \cdot (n-2) \cdots 3 \cdot 2 \cdot 1$$

we take $0! = 1$.

Definition 1.6. *If n, m are integers, $n \geq m$, we define the binomial coefficients*

$$\binom{n}{m} = \frac{n!}{m!(n-m)!}$$

Lemma 1.3.

$$\binom{N}{m} + \binom{N}{m-1} = \binom{N+1}{m}$$

Proof:

$$\begin{aligned}\binom{N}{m} + \binom{N}{m-1} &= \frac{N!}{(N-m)!m!} + \frac{N!}{(N-(m-1))!(m-1)!}\\ &= \frac{N!}{(m-1)!(N-m)!}\left[\frac{1}{m} + \frac{1}{N-m+1}\right]\\ &= \frac{N!}{(m-1)!(N-m)!}\left[\frac{N+1}{m(N-m+1)}\right] = \frac{(N+1)!}{m!(N-m+1)!}\end{aligned}$$

■

Theorem 1.5. *If x is a real number and n is an integer, then*

$$(1+x)^n = \sum_{m=0}^{n} \binom{n}{m} x^m \tag{1.27}$$

Proof: We will proceed by induction. We first note that if $n = 1$ then the right-hand side of (1.27) reduces to

$$\binom{1}{0} x^0 + \binom{1}{1} x = 1 + x$$

Assume that (1.27) is true for $n = N$, then

$$\begin{aligned}
(1+x)^{N+1} &= (1+x)\sum_{m=0}^{N}\binom{N}{m}x^m \\
&= \sum_{m=0}^{N}\binom{N}{m}x^m + \binom{N}{m}x^{m+1} \\
&= 1 + \sum_{m=1}^{N}\left[\binom{N}{m} + \binom{N}{m-1}\right]x^m + x^{N+1} \\
&= \sum_{m=0}^{N+1}\binom{N+1}{m}x^m + x^{N+1}
\end{aligned}$$

where we have made use of the result of Lemma 1.3; thus, by principle of induction true for all n. ■

Corollary 1.2. *If x, y are real numbers and n is an integer, then*

$$(y+x)^n = \sum_{m=0}^{n}\binom{n}{m}x^m y^{n-m} \tag{1.28}$$

Proof:

$$(1+z)^n = \sum_{m=0}^{n}\binom{n}{m}z^m$$

Let $z = \frac{x}{y}$ and result follows. ■

1.5 Taylor's Series

Very often in physical problems you need to find a relatively simple approximation to a complex function or you need to estimate the size of a function. One of the most commonly used techniques is to approximate a function by a polynomial.

Theorem 1.6. *Let f be a real function, which is continuous and has continuous derivatives up to the $n+1$ order, then*

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f^{(2)}(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_n(x) \tag{1.29}$$

where $n! = n\cdot(n-1)\cdot(n-2)\cdots 3\cdot 2\cdot 1$ and

$$R_n(x) = \int_a^x \frac{f^{(n+1)}(t)}{n!}(x-t)^n\, dt \tag{1.30}$$

Proof: We proceed by induction. For $n = 0$, (1.29) reduces to

$$f(x) - f(a) = \int_a^x f'(t)dt \tag{1.31}$$

which is just a statement of the fundamental theorem of calculus.Now assuming that (1.29) is true for $n = N$,

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f^{(2)}(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(N)}(a)}{N!}(x-a)^N + \int_a^x \frac{f^{(N+1)}(t)}{N!}(x-t)^N \, dt \tag{1.32}$$

Now, use integration by parts to evaluate the integral on the left-hand side of (1.32) Let

$$u(t) = \frac{f^{N+1}(t)}{N!}, \quad dv = (x-t)^N$$

then

$$\int_a^x \frac{f^{(N+1)}(t)}{N!}(x-t)^N$$

$$dt = -\frac{f^{N+1}(t)}{N+1.N!}(x-t)^{N+1}|^x{}_a + \int_a^x \frac{f^{N+2}(t)}{(N+1)!}(x-t)^{N+1} \, dt \tag{1.33}$$

Hence, by principle of induction, results true for all n. ■

Clearly, if R_n goes to zero uniformly as $n \to \infty$, then we can find an infinite series. Examples:

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots$$
$$\sin x = 1 - x + \frac{x^3}{3!} + \cdots$$

An alternative form for the remainder term can be derived by making use of the mean value theorem for integrals, that is,

$$R_{n+1} = \int_a^x \frac{f^{(n+1)}(t)}{n!}(x-t)^n \, dt = f^{(n+1)}(\alpha)\frac{(x-a)(x-\alpha)^n}{n!} \tag{1.34}$$

where α is some number, $a \le \alpha \le x$. The form (1.34) is the Cauchy form of the remainder term.

An alternative form was derived by Lagrange

$$R_{n+1}(x) = \frac{f^{(n+1)}(\beta)}{(n+1)!}(x-a)^{n+1} \tag{1.35}$$

with $a \le \beta \le x$.

Corollary 1.3. *If $f(x)$ is a differentiable function defined on some interval, I, and*

$$f'(x) = 0 \text{ for all } x \in I$$

then$f(x)$ is constant on I.

Proof: f is differentiable; so applying (1.29) and(1.34), we have

$$f(x) = f(0) + R_1 = f(0)$$

for all $x \epsilon I$. ■

1.6
Extrema

Let us assume that $F(x)$ is a continuous function with a continuous first derivate. Assume further that F has a local maximum at some point x_0; hence, for some infinitesimal increment $|h|$,

$$F(x_0 \pm |h|) < F(x_0) \tag{1.36}$$

hence

$$\frac{F(x_0 + |h|) - F(x_0)}{|h|} < 0$$
$$\frac{F(x_0 - |h|) - F(x_0)}{-|h|} > 0 \tag{1.37}$$

We can make $|h|$ arbitrarily small. Hence, when we take limit from the left and right and since we have assumed the derivative is continuous, we must have

$$\frac{dF(x_0)}{dx} \equiv \frac{dF(x)}{dx}|_{x_0} = 0 \tag{1.38}$$

Following a similar argument, it is immediately obvious that if x_0 corresponds to a minimum (1.38) also holds. Now, assume that F is a continuous function with continuous first and second derivatives and that there is a point x_0 in its domain where (1.38) holds. Then, using the Taylor's expansion (1.29), we have

$$\begin{aligned} F(x) &= F(x_0) + (x - x_0)\frac{dF(x_0)}{dx} + \frac{1}{2}\frac{d^2F(x_0)}{dx^2}(x - x_0)^2 + O(|x - x_0|^3) \\ &= F(x_0) + \frac{1}{2}\frac{d^2F(x_0)}{dx^2}(x - x_0)^2 + O(|x - x_0|^3) \end{aligned} \tag{1.39}$$

Now, since $(x - x_0)^2 > 0$ and since we can choose x arbitrarily close to x_0, we see at once that

$$F \text{ has a maximum at } x_0 \text{ if } \frac{d^2F(x_0)}{dx^2} < 0$$
$$F \text{ has a minimum at } x_0 \text{ if } \frac{d^2F(x_0)}{dx^2} > 0$$

1.7
Power Series

The geometric series we considered earlier is an example of a power series.

Definition 1.7. *A power series is a function of a variable x defined as an infinite sum*

$$\sum_{n=0}^{\infty} a_n x^n \tag{1.40}$$

where a_n are numbers.

In the case of the geometric series, all the a_ns are equal. Just because we write down a series of the form (1.40) it does not mean that such a thing is well defined. It is, in essence, a limit of the sequence of partial sums and this limit may or may not exist. We have already seen that the geometric series converges if and only if $|x| < 1$. The interval of convergence is the range of values $a < x < b$ for which (1.40) converges. Note this is an open interval, that is, we need to consider the end points separately. In our example the geometric series diverges at both end points! We have seen in the previous section that if a function f has a Taylor expansion with remainder term R_n which uniformly goes to zero on some interval $I = \{x|a < x < b\}$, then f can be represented by a power series on this interval. Power series are extremely useful. We will state some results and refer the reader to [1] for proof, see also [3].

- A power series may be differentiated or integrated term by term: the resulting series converges to the derivative or the integral of the function represented by the original series within the same interval of convergence.
- Two power series may be added, subtracted, or multiplied by a constant, and this will converge at least within the same interval of convergence, i.e. Suppose that

  $$s_1(X) = \sum_{n=0}^{\infty} a_n X^n$$
  $$s_2(X) = \sum_{m=0}^{\infty} b_m X^m$$

  are both convergent within the interval I and α, β are numbers. Then,

  $$s_3 = \sum_{n=0}^{\infty} (\alpha a_n + \beta b_n) X^n$$

  is convergent within the interval I and

  $$s_3(X) = \alpha s_1(X) + \beta s_2(X)$$

- The power series of a function is unique, that is, if

  $$f(X) = \sum_{n=0}^{\infty} a_n X^n$$
  $$f(X) = \sum_{m=0}^{\infty} b_m X^m$$

  then $a_n = b_n$ for all n.

1.8 Basic Functions

1.8.1 Exponential

Assume that there exists a function exp(x), which is its own derivative, and whose value at $x = 0$ is 1, that is,

$$\frac{d\exp(x)}{dx} = \exp(x); \quad \exp(0) = 1 \tag{1.41}$$

We can then construct a Taylor's series

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots R_n. \tag{1.42}$$

For $x > 0$, each term is greater than zero, so the exponential function is positive. Since it is its own derivative, its derivative is positive and, therefore, is strictly increasing. We need to consider the remainder term

$$R_n = \frac{x^n \exp(\xi)}{n!}, \quad 0 \leq \xi \leq x$$

It can be shown that the remainder term can be made arbitrarily small by choosing n large enough, that is,

$$\lim_{n \to o} R_n = 0$$

hence

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \tag{1.43}$$

We take (1.43) to be the defining equation for the exponential function.

Definition 1.8. *The exponential function is defined for all real x by the power series*

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \tag{1.44}$$

We assume that (1.44) is uniformly convergent and hence differentiable term by term. Thus,

$$\begin{aligned}\frac{d\exp(x)}{dx} &= \sum_{n=0}^{\infty} n\frac{x^{(n-1)}}{n!} \\ &= \sum_{n=0}^{\infty} \frac{x^{(n-1)}}{(n-1)!} \\ &= \exp(x)\end{aligned}$$

Thus, the exponential function is its own derivative, as expected; furthermore

$$\exp(0) = 1 \tag{1.45}$$

Using the chain rule, we see at once that

$$\frac{d\exp(-x)}{dx} = -\exp(-x) \tag{1.46}$$

Consider

$$f(x) = \exp(x)\exp(-x) \Rightarrow \frac{df(x)}{dx} = 0$$

where we have made use of (1.46) and the product rule.Therefore,

$$\exp(x)\exp(-x) = c$$

where c is a constant, now considering $x = 0$ and using (1.45) we have

$$\exp(x)\exp(-x) = 1 \tag{1.47}$$

From this, it follows that

$$\exp(x) \neq 0 \tag{1.48}$$

for all x and

$$\exp{(x)}^{-1} = \exp(-x) \tag{1.49}$$

so $\exp(-x)$ is a strictly decreasing, positive function of x for $x > 0$. Clearly

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

Lemma 1.4. *If $g(x)$ is another function s.t.*

$$g'(x) = g(x)$$

then $g(x) = c\exp(x)$ where c is a constant.

Proof: Let

$$f(x) = \frac{g(x)}{\exp(x)}$$

then $f'(x) = 0$; hence $f(x) = c$ where c is a constant; hence, $c\exp(x) = g(x)$. ■

The immediate consequence of Lemma 1.4 is that $\exp(x)$ is uniquely defined by the requirement that it be equal to its derivative and the initial condition $\exp(0) = 1$, that is, if $g(x)$ is s.t. $g'(x) = g(x)$ and $g(0) = 1$, we know that $g(x) = c\exp(x)$. Substituting the values at 0 shows that $c = 1$; hence, $g(x) = \exp(x)$.

Corollary 1.4.

$$\exp(x + y) = \exp(x)\exp(y) \tag{1.50}$$

Proof: For any number y, define

$$g(x) = \exp(x + y)$$

Clearly, $g'(x) = g(x)$. Hence, from Lemma 1.4

$$\exp(x + y) = c\exp(x)$$

Replacing $x = 0$ shows that $c = \exp(y)$. ■

Hence, for every positive integer n $\exp(nx) = \exp{(x)}^n$, we further note

$$\exp(-x) = \exp{(x)}^{-1}$$

All of these prompt us to write

$$\exp(x) = e^x \tag{1.51}$$

where for consistency we define the irrational number "e" to be

$$e = \sum_{n=0}^{\infty} \frac{1}{n!} \cong 2.718282$$

The function $\exp(x)$ is a rapidly increasing function of x. If you have a situation where the rate at which a population grows or decreases is proportional to the population at a given time, then we have

$$\frac{dN(t)}{dt} = \mu N(t) \tag{1.52}$$

If the constant, μ, is positive, then we have a growing population, which would be characteristic of an animal population with plenty of available food and no predators. If μ is negative, then we have a decreasing population, for example, radioactive nuclei that decay probabilistically.

Example 1.4. *Radioactive Decay. Radioactive nuclei decay according to the law*

$$\frac{dN(t)}{dt} = -\lambda N(t) \tag{1.53}$$

$N(t)$ being the number of atoms at time t, λ is known as the decay constant and is characteristic of a given species. Rewriting (1.53) using our chain rule, we have

$$\frac{d(-\lambda N(t))}{d(-\lambda t)} = -\lambda N(t) \tag{1.54}$$

Hence

$$Ce^{-\lambda t}$$

is the solution, where C is a constant Equation (1.53) states that the rate of decay is proportional to the number of radioactive nuclei present. Let N_0 be the number present at time $t = 0$. Hence

$$N(t) = N_0 e^{-\lambda t}$$

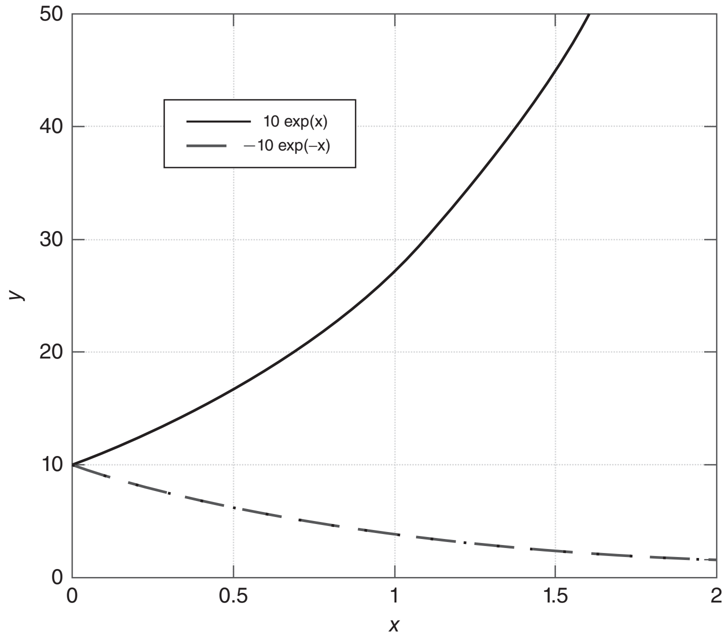

Figure 1.4 Plot of $N(t) = 10e^{\mu t}$. $\mu = +1$, solid line; $\mu = -1$, dashed line.

is the solution. A frequently asked question is how long will it take for half of the atoms to decay, which is the "half life" of the nuclear species, t_h (Figure 1.4). It is the solution of

$$\begin{aligned}\frac{N_0}{2} &= N_0 e^{-\lambda t_h}\\ \Rightarrow e^{-\lambda t_h} &= \frac{1}{2}\end{aligned}$$

1.8.2 Logarithm

The function $f(x) = \exp(x)$ is strictly increasing for all x and thus by Theorem 1.2 an inverse function g exists and

$$\frac{d\exp^{-1}(y)}{dy} = \frac{1}{f'(g(y))} = \frac{1}{f(g(y))} = \frac{1}{y}$$

Now, since $f(0) = 1$, we must have $g(1) = 0$ (Figure 1.5). The function g will be denoted by ln or $\log_e$.

Lemma 1.5.

$$\ln\ (xa) = \ln\ x + \ln\ a. \tag{1.55}$$

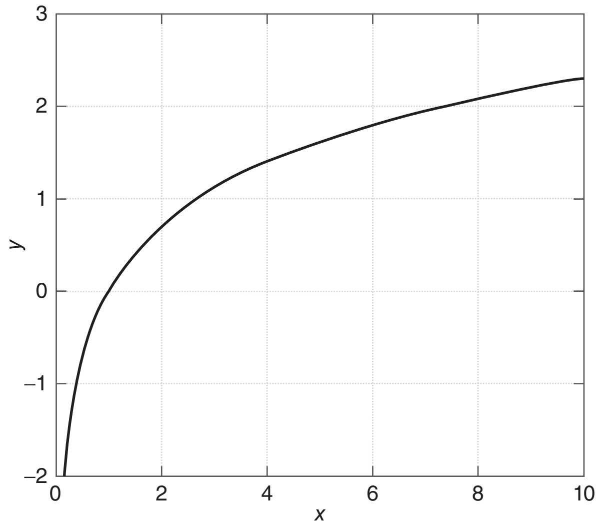

Figure 1.5 Plot of $y = \ln x$.

Proof: Consider

$$F(x) = \ln(ax) - \ln(x)$$

$$\begin{aligned} F'(x) &= \frac{d\ln(ax)}{d(ax)}\frac{dax}{dx} - \frac{d\ln(x)}{dx} \\ &= \frac{1}{ax}a - \frac{1}{x} \\ &= 0 \end{aligned}$$

Thus, $F(x) = c$, where c is constant. Now take $x = 1$ and we have $F(1) = \ln(a) = c$. ∎

Corollary 1.5.

$$\ln(x^n) = n\ln\ x$$

Proof: Clearly true for $n = 1$, assume to be true for $n = N$, that is, assume $\ln\ x^N = N\ln\ x$.

$$\ln(x^{N+1}) = \ln(x(x^N)) = \ln\ x + \ln\ x^N = \ln\ x + N\ln\ x = (N+1)\ln\ x$$

Thus, by principle of induction, true for all N. ∎

Assume that $n \neq 1$, then from Corollary 1.2

$$\ln(1^n) = n\ln(1)$$

which yields a contradiction unless

$$\begin{aligned} 0 &= \ln(1) \\ &= \ln(xx^{-1}) \\ &= \ln x + \ln x^{-1} \\ &\Rightarrow \\ \ln\ x^{-1} &= -\ln\ x \end{aligned} \tag{1.56}$$

It follows from (1.56) that as $x \to \infty$ $\ln(\frac{1}{x}) \to -\infty$.

1.9 First-Order Ordinary Differential Equations

Assume that we are presented with the ordinary differential equation

$$\frac{dy}{dt} + p(t)y = g(t) \tag{1.57}$$

then we can solve it by the following method: define

$$r(t) = \exp\left(\int_a^t p(x)\ dx\right) \tag{1.58}$$

Notice that x here is a dummy variable and a is an arbitrary constant, which we can choose later. Employing the chain rule and the fundamental theorem, we have

$$\frac{dr}{dt} = \exp\left(\int_a^t p(x)\ dx\right)\frac{d\int_a^t p(x)}{dt} = p(t)\exp\left(\int_a^t p(x)\ dx\right) \tag{1.59}$$

now multiplying (1.57) by $r(t)$, we immediately arrive at

$$\frac{dr(t)y(t)}{dt} = g(t)r(t) \tag{1.60}$$

which we can now integrate directly. The term $r(t)$ is known as an integrating factor.

Example 1.5. *Solve*

$$t\frac{dy(t)}{dt} + 2y = 4t^2 \tag{1.61}$$

subject to

$$y(1) = 2$$

Divide (1.61) by t to put it in the form (1.57), then

$$\frac{dy(t)}{dt} + 2\frac{y}{t} = 4t$$

$$r(t) = \exp\left(\int_a^t \frac{2}{x}dx\right)$$

$$\Rightarrow r(t) = \exp(2\ln t - 2\ln a) \tag{1.62}$$

With loss of generality, we may take the arbitrary constant a to be unity and then

$$r(t) = \exp(2\ln t) = \exp(\ln t^2) = t^2 \tag{1.63}$$

Multiplying (1.61) by t^2, we have

$$\begin{aligned} t^2\frac{dy}{dt} + 2ty &= 4t^3 \\ \Rightarrow \frac{dt^2 y}{dt} &= 4t^3 \\ \Rightarrow t^2 y(t) &= t^4 + c \end{aligned} \tag{1.64}$$

where c is a constant. Now substitute the initial condition $y(1) = 2$ and we have $c = 1$.

1.10
Trigonometric Functions

Lemma 1.6. *Let $c(x), s(x)$ be continuous differentiable functions such that*

$$\begin{aligned} s'(x) &= c(x) \\ c'(x), &= -s(x) \\ s(0) &= 0 \\ c(0) &= 1 \end{aligned} \tag{1.65}$$

then

$$c^2(x) + s^2(x) = 1 \tag{1.66}$$

Proof: Let

$$c^2(x) + s^2(x) = F(x)$$

$$F'(x) = 2c(x)c'(x) + 2s(x)s'(x) = 0$$

thus $F(x)$ must be a constant. Substituting the values at $x = 0$, we have the result. ■

Lemma 1.7. *If we have two sets of functions $c(x), s(x)$ and $f(x)$, $g(x)$ s.t.*

$$\begin{aligned} c'(x) &= -s(x) & g'(x) &= -f(x) \\ s'(x) &= c(x) & f'(x) &= g(x) \\ c(0) &= 1 & g(0) &= 1 \\ s(0) &= 0 & f(0) &= 0 \end{aligned}$$

then $f(x) = s(x)$; $c(x) = g(x)$ for all x.

Proof: We know that both the pairs $(c, s), (f, g)$ must satisfy the relation (1.66)

$$c^2(x) + s^2(x) = 1$$
$$f^2(x) + g^2(x) = 1$$

The functions

$$F_1(x) = f(x)c(x) - s(x)g(x)$$
$$F_2(x) = f(x)s(x) + c(x)g(x)$$

are such that

$$\frac{dF_1(x)}{dx} = \frac{dF_2(x)}{dx} = 0$$

Hence

$$a = f(x)c(x) - s(x)g(x)$$
$$b = f(x)s(x) + c(x)g(x)$$

where a and b are constants. Substituting the values at $x = 0$ yields

$$0 = f(x)c(x) - s(x)g(x)$$
$$1 = f(x)s(x) + c(x)g(x)$$

Hence

$$0 = f(x)c^2(x) - c(x)s(x)g(x)$$
$$s(x) = f(x)s^2(x) + s(x)c(x)g(x)$$

Adding the last two lines yields

$$s(x) = f(x)$$

Hence

$$s'(x) = f'(x)$$

Hence

$$c(x) = g(x)$$

■

Clearly, the functions $c(x), s(x)$ have all the properties of the $\sin(x)$ and $\cos(x)$ of trigonometry. The rest of the properties that we know and love can be derived from the above-mentioned results. We can write down a Taylor series for both using (1.65), which leads us to the following definition.

Definition 1.9. *We define the* sin *and* cos *functions by the uniformly convergent power series*

$$\sin(X) = \sum_{n\geq 0} \frac{(-1)^n}{(2n+1)!} X^{2n+1}$$
$$\cos(X) = \sum_{n\geq 0} \frac{(-1)^n}{(2n)!} X^{2n} \tag{1.67}$$

Differentiating term by term, we see that these functions satisfy the conditions in Lemma 1.7 and are consequently unique. This may appear to be an odd way to discuss the sin and cos functions but there is an important lesson here in that perfectly good functions can be defined simply as the solution of differential equations.

Definition 1.10. *We define the* tan, sec, *and* csc *functions*

$$\tan(X) = \frac{\sin X}{\cos X}$$
$$\sec(X) = \frac{1}{\cos X}$$
$$\csc(X) = \frac{1}{\sin X}$$

1.10.1 L'Hôpital's Rule

We often times have to deal with the limit of a quotient:

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

where

$$g(a) = f(a) = 0$$

Perhaps surprisingly a finite limit may exist, since

$$\frac{f(a+h)}{g(a+h)} = \frac{f(a+h) - f(a)}{g(a+h) - g(a)} \tag{1.68}$$

where we have used the fact that, $g(a) = f(a) = 0$. Now dividing the numerator and denominator on the right-hand side of (1.68) by h and taking the limit as $h \to 0$ yields

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{\lim_{h\to 0}(f(a+h) - f(a))/h}{\lim_{h\to 0}(g(a+h) - g(a))/h} = \frac{f'(a)}{g'(a)} \tag{1.69}$$

The relation (1.69) is known as "l'Hôpital's rule."

Example 1.6.

$$\lim_{x \to o} \frac{\sin x}{x} = \frac{\cos 0}{1} = 1$$

Problems

1.1 *Prove Rolle's theorem*: Let g be continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b). Assume that $g(a) = g(b)$, then show that there exists at least one point x_0 in (a, b) *s.t.* $g'(x_0) = 0$.

1.2 *Prove the mean value theorem*: Let f be continuous on the the closed interval $[a, b]$ and differentiable on the open interval (a, b) then there exist at least one point x_0 in (a, b) $s.t.$

$$\frac{f(b) - f(a)}{b - a} = f'(x_0)$$

1.3 Show that f is strictly increasing (decreasing) on an interval (a, b) if and only if $f'(x) > 0 (f'(x) < 0)$ for all $x \epsilon (a, b)$.

1.4 Let f be continuous and strictly increasing on $[a, b]$. Show that the inverse function of f is continuous and strictly increasing.

1.5 Show that

$$f(x) = 4x^2 - 4x + 1 = 0$$

has exactly one solution.

1.6 Solve

$$\frac{dy}{dx} + xy = x$$

for y.

1.7 Consider

$$f(x) = x^4 + 2x^3 - 3x^2 - 4x + 4$$

Find:

- the extrema of f
- the zeros of f
- the intervals on which f is increasing, decreasing

Plot the function.

1.8 Assume that $y = x^x$. Find $\dfrac{dy}{dx}$.

1.9 a) Evaluate

$$\int \frac{dx}{\sqrt{24 - 16x^2}}$$

b) Evaluate

$$\int \frac{dx}{24 + 16x^2}$$

1.10 From first principles, prove that

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

1.11 Evaluate

$$\int e^x \cos(x) dx$$