

THE DEFINITE INTEGRAL

11.1 STATEMENT OF THE PROBLEM. LOWER AND UPPER SUMS

The **definite integral** is one of the basic concepts of mathematical analysis and is a powerful research tool in mathematics, physics, mechanics, and other disciplines. Calculation of areas bounded by curves, of arc lengths, volumes, work, velocity, path length, moments of inertia, and so forth reduce to the evaluation of a definite integral.

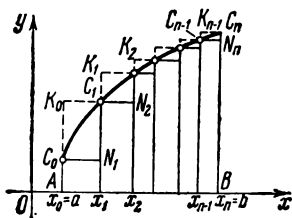


Fig. 210

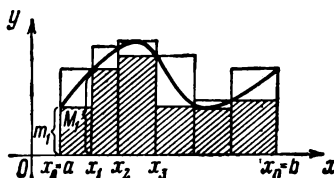


Fig. 211

Let a **continuous** function $y = f(x)$ be given on an interval $[a, b]$ (Figs. 210 and 211). Denote by m and M its smallest and largest values on this interval. Divide the interval $[a, b]$ into n subintervals:

$$a = x_0, x_1, x_2, \dots, x_{n-1}, x_n = b$$

so that

$$x_0 < x_1 < x_2 < \dots < x_n$$

and put

$$x_1 - x_0 = \Delta x_1, x_2 - x_1 = \Delta x_2, \dots, x_n - x_{n-1} = \Delta x_n$$

Then denote the smallest and greatest values of the function $f(x)$

on the subinterval $[x_0, x_1]$ by m_1 and M_1

on the subinterval $[x_1, x_2]$ by m_2 and M_2

.....

on the subinterval $[x_{n-1}, x_n]$ by m_n and M_n

Form the sums

$$\underline{s}_n = m_1 \Delta x_1 + m_2 \Delta x_2 + \dots + m_n \Delta x_n = \sum_{i=1}^n m_i \Delta x_i \quad (1)$$

$$\bar{s}_n = M_1 \Delta x_1 + M_2 \Delta x_2 + \dots + M_n \Delta x_n = \sum_{i=1}^n M_i \Delta x_i \quad (2)$$

The sum \underline{s}_n is called the *lower (integral) sum*, and the sum \bar{s}_n is called the *upper (integral) sum*.

If $f(x) \geq 0$, then the lower sum is numerically equal to the area of an "inscribed step-like figure" $AC_0N_1C_1N_2 \dots C_{n-1}N_nBA$ bounded by an "inscribed" broken line, the upper sum is equal numerically to the area of a "circumscribed step-like figure" $AK_0C_1K_1 \dots C_{n-1}K_{n-1}C_nBA$ bounded by a "circumscribed" broken line.

The following are some properties of upper and lower sums.

(a) Since $m_i \leq M_i$ for any i ($i = 1, 2, \dots, n$), by formulas (1) and (2) we have

$$\underline{s}_n \leq \bar{s}_n \quad (3)$$

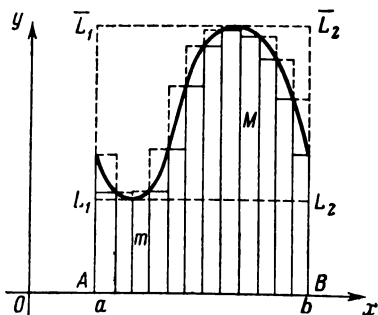


Fig. 212

(The equal sign occurs only when $f(x) = \text{const.}$)

(b) Since

$$m_1 \geq m, m_2 \geq m, \dots, m_n \geq m,$$

where m is the smallest value of $f(x)$ on $[a, b]$, we have

$$\begin{aligned} \underline{s}_n &= m_1 \Delta x_1 + m_2 \Delta x_2 + \dots + m_n \Delta x_n \geq m \Delta x_1 + m \Delta x_2 + \dots + m \Delta x_n \\ &= m(\Delta x_1 + \Delta x_2 + \dots + \Delta x_n) = m(b-a) \end{aligned}$$

Thus,

$$\underline{s}_n \geq m(b-a) \quad (4)$$

(c) Since

$$M_1 \leq M, M_2 \leq M, \dots, M_n \leq M$$

where M is the greatest value of $f(x)$ on $[a, b]$, we have

$$\begin{aligned} \bar{s}_n &= M_1 \Delta x_1 + M_2 \Delta x_2 + \dots + M_n \Delta x_n \leq M \Delta x_1 + M \Delta x_2 + \dots + M \Delta x_n \\ &= M(\Delta x_1 + \Delta x_2 + \dots + \Delta x_n) = M(b-a) \end{aligned}$$

Thus,

$$\bar{s}_n \leq M(b-a) \quad (5)$$

Combining the inequalities obtained, we get

$$m(b-a) \leq \underline{s}_n \leq \bar{s}_n \leq M(b-a)$$

If $f(x) \geq 0$, then the last inequality has a simple geometric meaning (Fig. 212), because the products $m(b-a)$ and $M(b-a)$ are, respectively, numerically equal to the areas of the "inscribed" rectangle AL_1L_2B and the "circumscribed" rectangle $\bar{A}\bar{L}_1\bar{L}_2B$.

11.2 THE DEFINITE INTEGRAL.

PROOF OF THE EXISTENCE OF A DEFINITE INTEGRAL

We continue examining the question of the preceding section. In each of the subintervals $[x_0, x_1]$, $[x_1, x_2]$, ..., $[x_{n-1}, x_n]$ take

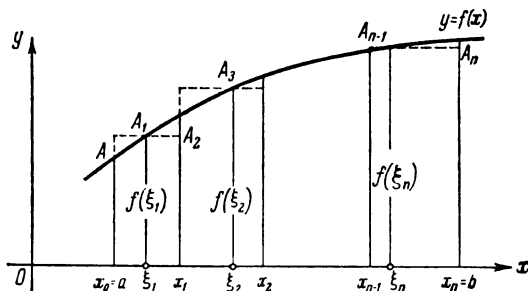


Fig. 213

a point and denote the points by $\xi_1, \xi_2, \dots, \xi_n$ (Fig. 213):

$$x_0 < \xi_1 < x_1, x_1 < \xi_2 < x_2, \dots, x_{n-1} < \xi_n < x_n$$

At each of these points find the value of the function $f(\xi_1), f(\xi_2), \dots, f(\xi_n)$. Form the sum

$$s_n = f(\xi_1) \Delta x_1 + f(\xi_2) \Delta x_2 + \dots + f(\xi_n) \Delta x_n = \sum_{i=1}^n f(\xi_i) \Delta x_i \quad (1)$$

This sum is called the *integral sum* of the function $f(x)$ on the interval $[a, b]$. Since for an arbitrary ξ_i belonging to the interval $[x_{i-1}, x_i]$ we will have

$$m_i \leq f(\xi_i) \leq M_i$$

and all $\Delta x_i > 0$, it follows that

$$m_i \Delta x_i \leq f(\xi_i) \Delta x_i \leq M_i \Delta x_i$$

and consequently

$$\sum_{i=1}^n m_i \Delta x_i \leq \sum_{i=1}^n f(\xi_i) \Delta x_i \leq \sum_{i=1}^n M_i \Delta x_i$$

or

$$\underline{s}_n \leq s_n \leq \bar{s}_n \quad (2)$$

The geometric meaning of the latter inequality for $f(x) \geq 0$ consists in the fact that the figure whose area is equal to s_n is bounded by a broken line lying between the "inscribed" broken line and the "circumscribed" broken line.

The sum s_n depends upon the way in which the interval $[a, b]$ is divided into the subintervals $[x_{i-1}, x_i]$ and also upon the choice of points ξ_i inside the resulting subintervals.

Let us now denote by $\max [x_{i-1}, x_i]$ the largest of the subintervals $[x_0, x_1], [x_1, x_2], \dots, [x_{n-1}, x_n]$. Let us consider different partitions of the interval $[a, b]$ into subintervals $[x_{i-1}, x_i]$ such that $\max [x_{i-1}, x_i] \rightarrow 0$. Obviously, the number of subintervals n then approaches infinity. Choosing the appropriate values of ξ_i , it is possible, for each partition, to form the integral sum

$$s_n = \sum_{i=1}^n f(\xi_i) \Delta x_i \quad (3)$$

We consider a certain sequence of partitions for which $\max \Delta x_i \rightarrow 0$ as $n \rightarrow \infty$. We choose the values ξ_i for each partition. Let us suppose that this sequence of integral sums s_n^* tends to a certain limit

$$\lim_{\max \Delta x_i \rightarrow 0} s_n^* = \lim_{\max \Delta x_i \rightarrow 0} \sum_{i=1}^n f(\xi_i) \Delta x_i = s \quad (4)$$

We are now able to state the following.

Definition 1. If for arbitrary partitions of the interval $[a, b]$ such that $\max \Delta x_i \rightarrow 0$ and for any choice of the points ξ_i on the subintervals $[x_{i-1}, x_i]$ the integral sum

$$s_n = \sum_{i=1}^n f(\xi_i) \Delta x_i \quad (5)$$

tends to one and the same limit s , then this limit is termed the *definite integral* of the function $f(x)$ on the interval $[a, b]$ and is denoted by

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

Thus, by definition,

$$\lim_{\max \Delta x_i \rightarrow 0} \sum_{i=1}^n f(\xi_i) \Delta x_i = \int_a^b f(x) dx \quad (6)$$

* In this case the sum is an ordered variable quantity.

The number a is termed the *lower limit* of the integral, b , the *upper limit* of the integral. The interval $[a, b]$ is called the *interval of integration* and x is the *variable of integration*.

Definition 2. If for a function $f(x)$ the limit (6) exists, then we say the function is *integrable on the interval* $[a, b]$.

Note that the lower sum s_n and the upper sum s_n are particular cases of the sum (5) and so if $f(x)$ is integrable, then the lower and upper sums tend to the same limit s and therefore, by (6), we can write

$$\lim_{\max \Delta x_i \rightarrow 0} \sum_{i=1}^n m_i \Delta x_i = \int_a^b f(x) dx \quad (7)$$

$$\lim_{\max \Delta x_i \rightarrow 0} \sum_{i=1}^n M_i \Delta x_i = \int_a^b f(x) dx \quad (7')$$

If we construct the graph of the integrand $y = f(x)$, then in the case of $f(x) \geq 0$ the integral

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

will be numerically equal to the *area* of a so-called *curvilinear trapezoid* bounded by the given curve, the straight lines $x=a$ and $x=b$, and the x -axis (Fig. 214).

For this reason, if it is required to compute the area of a curvilinear trapezoid bounded by the curve $y = f(x)$, the straight lines $x=a$ and $x=b$, and the x -axis, this area Q is computed by means of the integral

$$Q = \int_a^b f(x) dx \quad (8)$$

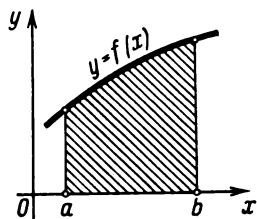


Fig. 214

We will prove the following theorem.

Theorem 1. If a function $f(x)$ is continuous on an interval $[a, b]$, then it is integrable on that interval.

Proof. Again partition the interval $[a, b]$ ($a < b$) into subintervals $[x_0, x_1], [x_1, x_2], \dots, [x_{i-1}, x_i], \dots, [x_{n-1}, x_n]$. Form the lower and upper sums:

$$\underline{s}_n = \sum_{i=1}^n m_i \Delta x_i \quad (9)$$

$$\bar{s}_n = \sum_{i=1}^n M_i \Delta x_i \quad (10)$$

For what follows we will need certain properties of upper and lower sums.

Property 1. *If the number of subintervals into which $[a, b]$ is partitioned by adding points of division is increased, the lower sum can only increase and the upper sum can only decrease.*

Proof. Let the interval $[a, b]$ be partitioned into n' subintervals by adding new points of division ($n' > n$). If some subinterval $[x_{k-1}, x_k]$ is split up into several parts, say p_k parts, then in the new lower sum $\underline{s}_{n'}$ the subinterval $[x_{k-1}, x_k]$ will be associated with p_k summands, which we denote by $\underline{s}_{p_k}^*$. In the sum \underline{s}_n , this subinterval will correspond to one term $m_k(x_k - x_{k-1})$. But then an inequality, similar to the inequality (4) of Sec. 11.1, holds true for the sum $\underline{s}_{p_k}^*$ and the quantity $m_k(x_k - x_{k-1})$. We can write

$$\underline{s}_{p_k}^* \geq m_k(x_k - x_{k-1})$$

Writing down the appropriate inequalities for each subinterval and summing the left and right members, we get

$$\underline{s}_{n'} \geq \underline{s}_n \quad (n' > n) \quad (11)$$

This completes the proof of Property 1.

Property 2. *In the case of an unlimited increase in the number of subintervals accomplished by adding new division points, the lower sum (9) and the upper sum (10) tend to certain limits \underline{s} and \bar{s} .*

Proof. By inequality (6) of Sec. 11.1 we can write

$$\underline{s}_n \leq M(b-a)$$

That is, \underline{s}_n is bounded for all n . By Property 1, \underline{s}_n increases monotonically with increasing n . Hence, by Theorem 7 on limits (see Sec. 2.5), this variable has a limit, which we denote by \underline{s} :

$$\lim_{n \rightarrow \infty} \underline{s}_n = \underline{s} \quad (12)$$

Similarly, we find that \bar{s}_n is bounded below and decreases monotonically. Consequently, \bar{s}_n has a limit, which we denote by \bar{s} :

$$\lim_{n \rightarrow \infty} \bar{s}_n = \bar{s}$$

Property 3. *If a function $f(x)$ is continuous on a closed interval $[a, b]$, then the limits \underline{s} and \bar{s} , defined by Property 2 on the condition that $\max \Delta x_i \rightarrow 0$, are equal.*

We denote this common limit by s :

$$\underline{s} = \bar{s} = s \quad (13)$$

Proof. Let us consider the difference between the upper and lower sums:

$$\begin{aligned}\bar{s}_n - \underline{s}_n &= (M_1 - m_1) \Delta x_1 + (M_2 - m_2) \Delta x_2 + \dots \\ &\quad + (M_i - m_i) \Delta x_i + \dots + (M_n - m_n) \Delta x_n = \sum_{i=1}^n (M_i - m_i) \Delta x_i \quad (14)\end{aligned}$$

Denote by ϵ_n the maximum difference $(M_i - m_i)$ for a given partition:

$$\epsilon_n = \max (M_i - m_i)$$

It may be proved (though we will not do so) that if a function $f(x)$ is continuous on a closed interval, then for any mode of partition of the interval $[a, b]$, $\epsilon_n \rightarrow 0$ provided $\max \Delta x_i \rightarrow 0$:

$$\lim_{\max \Delta x_i \rightarrow 0} \epsilon_n = 0 \quad (15)$$

The property of a continuous function on a closed interval as expressed by equation (15) is called *uniform continuity* of the function.

We will thus make use of the theorem that *a function continuous on a closed interval is uniformly continuous on that interval*.

Reverting to (14), we replace each difference $(M_i - m_i)$ on the right by ϵ_n , which is at least equal to the difference. This yields the inequality

$$\begin{aligned}\bar{s}_n - \underline{s}_n &\leq \epsilon_n \Delta x_1 + \epsilon_n \Delta x_2 + \dots + \epsilon_n \Delta x_n = \\ &= \epsilon_n (\Delta x_1 + \Delta x_2 + \dots + \Delta x_n) = \epsilon_n (b - a)\end{aligned}$$

Passing to the limit as $\max \Delta x_i \rightarrow 0$ ($n \rightarrow \infty$), we get

$$\lim_{\max \Delta x_i \rightarrow 0} (\bar{s}_n - \underline{s}_n) \leq \lim_{\max \Delta x_i \rightarrow 0} \epsilon_n (b - a) = (b - a) \lim_{\max \Delta x_i \rightarrow 0} \epsilon_n = 0 \quad (16)$$

That is,

$$\lim \bar{s}_n = \lim \underline{s}_n = s \quad (17)$$

or $\bar{s} = \underline{s} = s$, which completes the proof.

Property 4. Let s_{n_1} and \bar{s}_{n_1} be the lower and upper sums corresponding to partitions of the interval $[a, b]$ into n_1 and n_2 subintervals, respectively. We then have the inequality

$$\underline{s}_{n_1} \leq \bar{s}_{n_2} \quad (18)$$

for arbitrary n_1 and n_2 .

Proof. Consider a partition of $[a, b]$ into $n_3 = n_1 + n_2$ subintervals, where the division points are those of the first and second partitions.

By inequality (3) of Sec. 11.1, we have

$$\underline{s}_{n_1} \leq \bar{s}_{n_1} \quad (19)$$

On the basis of Property 1, we have

$$\underline{s}_{n_1} \leq \underline{s}_{n_2} \quad (20)$$

$$\bar{s}_{n_1} \leq \bar{s}_{n_2} \quad (21)$$

Using relations (20) and (21), we can extend inequality (19)

$$\underline{s}_{n_1} \leq \underline{s}_{n_2} \leq \bar{s}_{n_2} \leq \bar{s}_{n_1}$$

or

$$\underline{s}_{n_1} \leq s_{n_2}$$

which completes the proof.

Property 5. If a function $f(x)$ is continuous on an interval $[a, b]$, then for any sequence of partitions of $[a, b]$ into subintervals $[x_{i-1}, x_i]$, not necessarily by means of adjoining new points of division, provided $\max \Delta x_i \rightarrow 0$, the lower (integral) sum \underline{s}_m^* and the upper (integral) sum \bar{s}_m^* tend to the limit s defined by Property 3.

Proof. Let us consider a sequence of partitions of the sequence of upper sums \bar{s}_n defined by Property 2. For arbitrary values of n and m [by inequality (18)], we can write

$$\underline{s}_m^* \leq \bar{s}_n$$

Passing to the limit as $n \rightarrow \infty$ we can write, by (15),

$$\underline{s}_m^* \leq s$$

Similarly we can prove that $s \leq \bar{s}_m^*$.

Thus,

$$\underline{s}_m^* \leq s \leq \bar{s}_m^*$$

or

$$s - \underline{s}_m^* \geq 0, \quad \bar{s}_m^* - s \geq 0 \quad (22)$$

Consider the limit of the difference:

$$\lim_{\max \Delta x_i \rightarrow 0} (\bar{s}_m^* - \underline{s}_m^*)$$

Since the function $f(x)$ is continuous on the closed interval $[a, b]$ we will prove [in the same way as for Property 3, see equation (16)] that

$$\lim_{\max \Delta x_i \rightarrow 0} (\bar{s}_m^* - \underline{s}_m^*) = 0$$

We rewrite this relation as

$$\lim_{\max \Delta x_i \rightarrow 0} [(\bar{s}_m^* - s) + (s - \underline{s}_m^*)] = 0$$

By (22), each of the differences in the square brackets is nonnegative. Hence,

$$\lim_{\max \Delta x_i \rightarrow 0} (\bar{s}_m^* - s) = 0, \quad \lim_{\max \Delta x_i \rightarrow 0} (s - \underline{s}_m^*) = 0$$

and we finally get

$$\lim_{\max \Delta x_i \rightarrow 0} \underline{s}_m^* = s, \quad \lim_{\max \Delta x_i \rightarrow 0} \bar{s}_m^* = s \quad (23)$$

This completes the proof.

Now we can prove the foregoing theorem. Let a function $f(x)$ be continuous on an interval $[a, b]$. We consider an arbitrary sequence of (integral) sums

$$s_n = \sum_{i=1}^n f(\xi_i) \Delta x_i$$

such that $\max \Delta x_i \rightarrow 0$ and ξ_i is an arbitrary point in the subinterval $[x_{i-1}, x_i]$.

We consider the appropriate sequences of upper and lower sums s_n and \bar{s}_n for the given sequence of partitions. The relations (2) hold true for each partition:

$$\underline{s}_n < s_n < \bar{s}_n$$

Passing to the limit as $\max \Delta x_i \rightarrow 0$ and using equations (23) and Theorem 4, Sec. 2.5, we get

$$\lim_{\max \Delta x_i \rightarrow 0} s_n = s$$

where s is the limit defined by Property 3.

As already stated, this limit is termed the definite integral $\int_a^b f(x) dx$. Thus, if $f(x)$ is continuous on the interval $[a, b]$, then

$$\lim_{\max \Delta x_i \rightarrow 0} \sum f(\xi_i) \Delta x_i = \int_a^b f(x) dx \quad (24)$$

It may be noted that there are both integrable and nonintegrable functions in the class of discontinuous functions.

Note 1. It will be noted that the definite integral depends only on the form of the function $f(x)$ and the limits of integration, and not on the variable of integration, which may be denoted by

any letter. Thus, without changing the magnitude of a definite integral it is possible to replace the letter x by any other letter:

$$\int_a^b f(x) dx = \int_a^b f(t) dt = \dots = \int_a^b f(z) dz$$

Note 2. When introducing the concept of the definite integral $\int_a^b f(x) dx$ we assumed that $a < b$. In the case where $b < a$ we will, by definition, have

$$\int_a^b f(x) dx = - \int_b^a f(x) dx \quad (25)$$

Thus, for instance,

$$\int_5^0 x^2 dx = - \int_0^5 x^2 dx$$

Note 3. In the case of $a = b$ we assume, by definition, that for any function $f(x)$ we have

$$\int_a^a f(x) dx = 0 \quad (26)$$

This is natural also from the geometric standpoint. Indeed, the base of a curvilinear trapezoid has length equal to zero; consequently, its area is zero too.

Example 1. Compute the integral $\int_a^b kx dx$ ($b > a$).

Solution. Geometrically, the problem is equivalent to computing the area Q of a trapezoid bounded by the lines $y = kx$, $x = a$, $x = b$, $y = 0$ (Fig. 215).

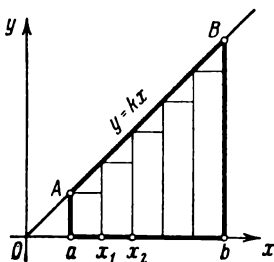


Fig. 215

The function $y = kx$ under the integral sign is continuous. Therefore, in order to compute the definite integral we have the right, as was stated above, to divide the interval $[a, b]$ in any way and choose arbitrary intermediate points ξ_k . The result of computing a definite integral is independent of the way in which the integral sum is formed, provided that the subinterval approaches zero.

Divide the interval $[a, b]$ into n equal subintervals.

The length Δx of each subinterval is $\Delta x = \frac{b-a}{n}$; this number is the partition unit.

The division points have coordinates:

$$x_0 = a, \quad x_1 = a + \Delta x, \\ x_2 = a + 2\Delta x, \quad \dots, \quad x_n = a + n\Delta x$$

For the points ξ_k take the left end points of each subinterval:

$$\xi_1 = a, \quad \xi_2 = a + \Delta x, \quad \xi_3 = a + 2\Delta x, \quad \dots, \quad \xi_n = a + (n-1)\Delta x$$

Form the integral sum (1). Since $f(\xi_i) = k\xi_i$, we have

$$\begin{aligned} s_n &= k\xi_1 \Delta x + k\xi_2 \Delta x + \dots + k\xi_n \Delta x \\ &= ka \Delta x + [k(a + \Delta x)] \Delta x + \dots + \{k[a + (n-1)\Delta x]\} \Delta x \\ &= k\{a + (a + \Delta x) + (a + 2\Delta x) + \dots + [a + (n-1)\Delta x]\} \Delta x \\ &= k\{na + [\Delta x + 2\Delta x + \dots + (n-1)\Delta x]\} \Delta x \\ &= k\{na + [1 + 2 + \dots + (n-1)] \Delta x\} \Delta x \end{aligned}$$

where $\Delta x = \frac{b-a}{n}$. Taking into account that

$$1 + 2 + \dots + (n-1) = \frac{n(n-1)}{2}$$

(as the sum of an arithmetic progression), we get

$$s_n = k \left[na + \frac{n(n-1)}{2} \frac{b-a}{n} \right] \frac{b-a}{n} = k \left[a + \frac{n-1}{n} \frac{b-a}{2} \right] (b-a)$$

Since $\lim_{n \rightarrow \infty} \frac{n-1}{n} = 1$, it follows that

$$\lim_{n \rightarrow \infty} s_n = Q = k \left[a + \frac{b-a}{2} \right] (b-a) = k \frac{b^2 - a^2}{2}$$

Thus,

$$\int_a^b kx \, dx = k \frac{b^2 - a^2}{2}$$

The area of $ABba$ (Fig. 215) is readily computed by the methods of elementary geometry. The result will be the same.

Example 2. Evaluate $\int_0^b x^2 \, dx$.

Solution. The given integral is equal to the area Q of a curvilinear trapezoid bounded by a parabola $y = x^2$, the ordinate $x = b$, and the straight line $y = 0$ (Fig. 216).

Divide the interval $[a, b]$ into n equal parts by the points

$$x_0 = 0, \quad x_1 = \Delta x, \quad x_2 = 2\Delta x, \quad \dots, \quad x_n = b = n\Delta x, \quad \Delta x = \frac{b}{n}$$

For the ξ_i points take the right extremities of each subinterval. Form the integral sum

$$\begin{aligned} s_n &= x_1^2 \Delta x + x_2^2 \Delta x + \dots + x_n^2 \Delta x \\ &= [(\Delta x)^2 \Delta x + (2\Delta x)^2 \Delta x + \dots + (n\Delta x)^2 \Delta x] = (\Delta x)^3 [1^2 + 2^2 + \dots + n^2] \end{aligned}$$

As we know,

$$1^2 + 2^2 + 3^2 + \dots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

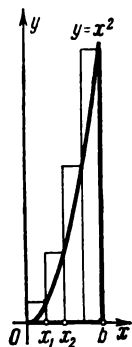


Fig. 216

therefore

$$s_n = \frac{b^3}{n^3} \frac{n(n+1)(2n+1)}{6} = \frac{b^3}{6} \left(1 + \frac{1}{n}\right) \left(2 + \frac{1}{n}\right)$$

$$\lim_{n \rightarrow \infty} s_n = Q = \int_0^b x^2 dx = \frac{b^3}{3}$$

Example 3. Evaluate $\int_a^b m dx$ ($m = \text{const.}$).

Solution.

$$\begin{aligned} \int_a^b m dx &= \lim_{\max \Delta x_i \rightarrow 0} \sum_{i=1}^n m \Delta x_i = \lim_{\max \Delta x_i \rightarrow 0} m \sum_{i=1}^n \Delta x_i \\ &= m \lim_{\max \Delta x_i \rightarrow 0} \sum_{i=1}^n \Delta x_i = m(b-a) \end{aligned}$$

Here, $\sum_{i=1}^n \Delta x_i$ is the sum of the lengths of the subintervals into which the interval $[a, b]$ was divided. No matter what the method of partition, the sum is equal to the length of the segment $b-a$.

Example 4. Evaluate $\int_a^b e^x dx$.

Solution. Again divide the interval $[a, b]$ into n equal parts:

$$x_0 = a, x_1 = a + \Delta x, \dots, x_n = a + n\Delta x, \Delta x = \frac{b-a}{n}$$

Take the left extremities as the points ξ_i . Then form the sum

$$\begin{aligned} s_n &= e^a \Delta x + e^{a+\Delta x} \Delta x + \dots + e^{a+(n-1)\Delta x} \Delta x \\ &= e^a (1 + e^{\Delta x} + e^{2\Delta x} + \dots + e^{(n-1)\Delta x}) \Delta x \end{aligned}$$

The expression in the brackets is a geometric progression with common ratio $e^{\Delta x}$ and first term 1; therefore

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

Then we have

$$n\Delta x = b-a, \quad \lim_{\Delta x \rightarrow 0} \frac{\Delta x}{e^{\Delta x} - 1} = 1$$

(By l'Hospital's rule $\lim_{z \rightarrow 0} \frac{z}{e^z - 1} = \lim_{z \rightarrow 0} \frac{1}{e^z} = 1$.) Thus,

$$\lim_{n \rightarrow \infty} s_n = Q = e^a (e^{b-a} - 1) \cdot 1 = e^b - e^a$$

that is,

$$\int_a^b e^x dx = e^b - e^a$$

Note 4. The foregoing examples show that the direct evaluation of definite integrals as the limits of integral sums involves great difficulties. Even when the integrands are very simple (kx , x^2 , e^x), this method involves cumbersome computations. The finding of definite integrals of more complicated functions leads to still greater difficulties. The natural problem that arises is to find some practically convenient way of evaluating definite integrals. This method, which was discovered by Newton and Leibniz, utilizes the profound relationship that exists between integration and differentiation. The following sections of this chapter are devoted to the exposition and substantiation of this method.

11.3 BASIC PROPERTIES OF THE DEFINITE INTEGRAL

Property 1. *A constant factor may be taken outside the sign of the definite integral: if $A = \text{const}$, then*

$$\int_a^b Af(x) dx = A \int_a^b f(x) dx \quad (1)$$

Proof.

$$\begin{aligned} \int_a^b Af(x) dx &= \lim_{\max \Delta x_i \rightarrow 0} \sum_{i=1}^n Af(\xi_i) \Delta x_i \\ &= A \lim_{\max \Delta x_i \rightarrow 0} \sum_{i=1}^n f(\xi_i) \Delta x_i = A \int_a^b f(x) dx \end{aligned}$$

Property 2. *The definite integral of an algebraic sum of several functions is equal to the algebraic sum of the integrals of the summands. Thus, in the case of two terms*

$$\int_a^b [f_1(x) + f_2(x)] dx = \int_a^b f_1(x) dx + \int_a^b f_2(x) dx \quad (2)$$

Proof.

$$\begin{aligned} \int_a^b [f_1(x) + f_2(x)] dx &= \lim_{\max \Delta x_i \rightarrow 0} \sum_{i=1}^n [f_1(\xi_i) + f_2(\xi_i)] \Delta x_i \\ &= \lim_{\max \Delta x_i \rightarrow 0} \left[\sum_{i=1}^n f_1(\xi_i) \Delta x_i + \sum_{i=1}^n f_2(\xi_i) \Delta x_i \right] \\ &= \lim_{\max \Delta x_i \rightarrow 0} \sum_{i=1}^n f_1(\xi_i) \Delta x_i + \lim_{\max \Delta x_i \rightarrow 0} \sum_{i=1}^n f_2(\xi_i) \Delta x_i \\ &= \int_a^b f_1(x) dx + \int_a^b f_2(x) dx \end{aligned}$$

The proof is similar for any number of terms.

Properties 1 and 2, though proved only for the case $a < b$, hold also for $a \geq b$.

However, the following property holds only for $a < b$:

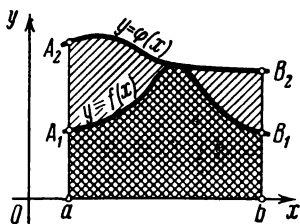
Property 3. If on an interval $[a, b]$ ($a < b$), the functions $f(x)$ and $\varphi(x)$ satisfy the condition $f(x) \leq \varphi(x)$, then

$$\int_a^b f(x) dx \leq \int_a^b \varphi(x) dx \quad (3)$$

Proof. Let us consider the difference

$$\begin{aligned} \int_a^b \varphi(x) dx - \int_a^b f(x) dx &= \int_a^b [\varphi(x) - f(x)] dx \\ &= \lim_{\max \Delta x_i \rightarrow 0} \sum_{i=1}^n [\varphi(\xi_i) - f(\xi_i)] \Delta x_i. \end{aligned}$$

Here, each difference $\varphi(\xi_i) - f(\xi_i) \geq 0$, $\Delta x_i \geq 0$. Thus, each term of the sum is nonnegative, the entire sum is nonnegative, and its limit is nonnegative; that is,



$$\int_a^b [\varphi(x) - f(x)] dx \geq 0$$

or

$$\int_a^b \varphi(x) dx - \int_a^b f(x) dx \geq 0$$

Fig. 217

whence follows inequality (3).

If $f(x) > 0$ and $\varphi(x) > 0$, then this property is nicely illustrated geometrically (Fig. 217). Since $\varphi(x) \geq f(x)$, the area of the curvilinear trapezoid aA_1B_1b does not exceed the area of the curvilinear trapezoid aA_2B_2b .

Property 4. If m and M are the smallest and greatest values of a function $f(x)$ on an interval $[a, b]$ and $a \leq b$, then

$$m(b-a) \leq \int_a^b f(x) dx \leq M(b-a) \quad (4)$$

Proof. It is given that

$$m \leq f(x) \leq M$$

On the basis of Property (3) we have

$$\int_a^b m dx \leq \int_a^b f(x) dx \leq \int_a^b M dx \quad (4')$$

But

$$\int_a^b m dx = m(b-a), \quad \int_a^b M dx = M(b-a)$$

(see Example 3, Sec. 11.2). Putting these expressions into inequality (4'), we get inequality (4).

If $f(x) \geq 0$, this property can easily be illustrated geometrically (Fig 218). The area of the curvilinear trapezoid $aABb$ lies between the areas of the rectangles aA_1B_1b and aA_2B_2b .

Property 5 (Mean-value theorem). If a function $f(x)$ is continuous on an interval $[a, b]$, then there is a point ξ on this interval such that the following equation holds:

$$\int_a^b f(x) dx = (b-a) f(\xi) \quad (5)$$

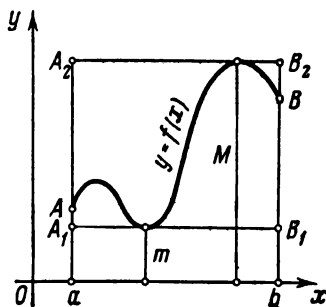


Fig. 218

Proof. For definiteness let $a < b$. If m and M are, respectively, the smallest and greatest values of $f(x)$ on $[a, b]$, then by virtue of (4)

$$m \leq \frac{1}{b-a} \int_a^b f(x) dx \leq M$$

whence

$$\frac{1}{b-a} \int_a^b f(x) dx = \mu, \quad \text{where } m \leq \mu \leq M$$

Since $f(x)$ is continuous on $[a, b]$, it takes on all intermediate values between m and M . Therefore, for some value ξ ($a \leq \xi \leq b$) we will have $\mu = f(\xi)$, or

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

Property 6. For any three numbers a, b, c the equation

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx \quad (6)$$

is true, provided all these three integrals exist.

Proof. First suppose that $a < c < b$, and form the integral sum of the function $f(x)$ on the interval $[a, b]$.

Since the limit of the integral sum is independent of the way in which the interval $[a, b]$ is divided into subintervals, we divide $[a, b]$ into subintervals such that the point c is the division point. Then we partition the sum \sum_a^b , which corresponds to the in-

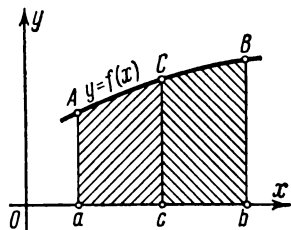


Fig. 219

interval $[a, b]$, into two sums: \sum_a^c , which corresponds to $[a, c]$, and \sum_c^b , which corresponds to $[c, b]$. Then

$$\sum_a^b f(\xi_i) \Delta x_i = \sum_a^c f(\xi_i) \Delta x_i + \sum_c^b f(\xi_i) \Delta x_i$$

Now, passing to the limit as $\max \Delta x_i \rightarrow 0$, we get relation (6).

If $a < b < c$, then on the basis of what has been proved we can write

$$\int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx \quad \text{or} \quad \int_a^b f(x) dx = \int_a^c f(x) dx - \int_b^c f(x) dx$$

but by formula (4), Sec. 11.2, we have

$$\int_b^c f(x) dx = - \int_c^b f(x) dx$$

Therefore,

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$$

This property is similarly proved for any other arrangement of points a , b , and c .

Fig. 219 illustrates Property 6, geometrically, or the case where $f(x) > 0$ and $a < c < b$: the area of the trapezoid $aABb$ is equal to the sum of the areas of the trapezoids $aACc$ and $cCBb$.

11.4 EVALUATING A DEFINITE INTEGRAL. THE NEWTON-LEIBNIZ FORMULA

In a definite integral

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

let the lower limit a be fixed and let the upper limit b vary. Then the value of the integral will vary as well: that is, the integral is a function of the upper limit.

So as to retain customary notations, we shall denote the upper limit by x , and to avoid confusion we shall denote the variable of integration by t . (This change in notation does not change the value of the integral.) We get the integral $\int_a^x f(t) dt$. For constant a , this integral will be a function of the upper limit x . We denote this function by $\Phi(x)$:

$$\Phi(x) = \int_a^x f(t) dt \quad (1)$$

If $f(t)$ is a nonnegative function, the quantity $\Phi(x)$ is numerically equal to the area of the curvilinear trapezoid $aAXx$ (Fig. 220). It is obvious that this area varies with x .

Let us find the derivative of $\Phi(x)$ with respect to x , i. e., the derivative of the definite integral (1) with respect to the upper limit.

Theorem 1. *If $f(x)$ is a continuous function and $\Phi(x) = \int_a^x f(t) dt$, then we have the equation*

$$\Phi'(x) = f(x)$$

In other words, the derivative of a definite integral with respect to the upper limit is equal to the integrand in which the value of the upper limit replaces the variable of integration (provided that the integrand is continuous).

Proof. Let us give the argument x a positive or negative increment Δx ; then (taking into account Property 6 of a definite integral) we get

$$\Phi(x + \Delta x) = \int_a^{x+\Delta x} f(t) dt = \int_a^x f(t) dt + \int_x^{x+\Delta x} f(t) dt$$

The increment of the function $\Phi(x)$ is equal to

$$\Delta\Phi = \Phi(x + \Delta x) - \Phi(x) = \int_a^x f(t) dt + \int_x^{x+\Delta x} f(t) dt - \int_a^x f(t) dt$$

that is,

$$\Delta\Phi = \int_x^{x+\Delta x} f(t) dt$$

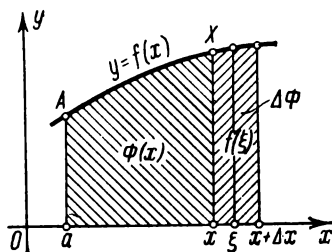


Fig. 220

Apply to the latter integral the mean-value theorem (Property 5 of a definite integral):

$$\Delta\Phi = f(\xi)(x + \Delta x - x) = f(\xi)\Delta x$$

where ξ lies between x and $x + \Delta x$.

Find the ratio of the increment of the function to the increment of the argument:

$$\frac{\Delta\Phi}{\Delta x} = \frac{f(\xi)\Delta x}{\Delta x} = f(\xi)$$

Hence,

$$\Phi'(x) = \lim_{\Delta x \rightarrow 0} \frac{\Delta\Phi}{\Delta x} = \lim_{\Delta x \rightarrow 0} f(\xi)$$

But since $\xi \rightarrow x$ as $\Delta x \rightarrow 0$, we have

$$\lim_{\Delta x \rightarrow 0} f(\xi) = \lim_{\xi \rightarrow x} f(\xi)$$

and due to the continuity of the function $f(x)$,

$$\lim_{\xi \rightarrow x} f(\xi) = f(x)$$

Thus, $\Phi'(x) = f(x)$, and the theorem is proved.

The geometric illustration of this theorem (Fig. 220) is simple; the increment $\Delta\Phi = f(\xi)\Delta x$ is equal to the area of a curvilinear trapezoid with base Δx , and the derivative $\Phi'(x) = f(x)$ is equal to the length of the segment xX .

Note. One consequence of the theorem that has been proved is that *every continuous function has an antiderivative*. Indeed, if the function $f(t)$ is continuous on the interval $[a, x]$, then, as was pointed out in Sec. 11.2, in this case the definite integral $\int_a^x f(t) dt$ exists, which is to say that the following function exists:

$$\Phi(x) = \int_a^x f(t) dt$$

But from what has already been proved, it is the antiderivative of $f(x)$.

Theorem 2. If $F(x)$ is some antiderivative of a continuous function $f(x)$, then the formula

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

holds.

This formula is known as the *Newton-Leibniz formula*.*

Proof. Let $F(x)$ be some antiderivative of the function $f(x)$. By Theorem 1, the function $\int_a^x f(t) dt$ is also an antiderivative of $f(x)$. But any two antiderivatives of a given function differ by a constant C^* . And so we can write

$$\int_a^x f(t) dt = F(x) + C^* \quad (3)$$

For an appropriate choice of C^* , this equation holds for all values of x , that is, it is an identity. To determine the constant C^* put $x=a$ in the identity; then

$$\int_a^a f(t) dt = F(a) + C^*$$

or

$$0 = F(a) + C^*$$

whence

$$C^* = -F(a)$$

Hence,

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

Putting $x=b$, we obtain the Newton-Leibniz formula:

$$\int_a^b f(t) dt = F(b) - F(a)$$

or, replacing the notation of the variable of integration by x ,

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

It will be noted that the difference $F(b) - F(a)$ is independent of the choice of antiderivative F , since all antiderivatives differ by a constant quantity, which disappears upon subtraction anyway.

* It is necessary to point out that the name of formula (2) is not exact, since neither Newton nor Leibniz had any such formula in the exact meaning of the word. The important thing, however, is that namely Leibniz and Newton were the first to establish a relationship between integration and differentiation, thus making possible the rule for evaluating definite integrals.

If we introduce the notation *

$$F(b) - F(a) = F(x)|_a^b$$

then formula (2) may be rewritten as follows:

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

The Newton-Leibniz formula yields a practical and convenient method for computing definite integrals in cases where the anti-derivative of the integrand is known. Only when this formula was established did the definite integral acquire its present significance in mathematics. Although the ancients (Archimedes) were familiar with a process similar to the computation of a definite integral as the limit of an integral sum, the applications of this method were confined to the very simple cases where the limit of the sum could be computed directly. The Newton-Leibniz formula greatly expanded the field of application of the definite integral, because mathematics obtained a **general method** for solving various problems of a particular type and so could considerably extend the range of applications of the definite integral to technology, mechanics, astronomy, and so on.

Example 1.

$$\int_a^b x dx = \frac{x^2}{2} \Big|_a^b = \frac{b^2 - a^2}{2}$$

Example 2.

$$\int_a^b x^2 dx = \frac{x^3}{3} \Big|_a^b = \frac{b^3 - a^3}{3}$$

Example 3.

$$\int_a^b x^n dx = \frac{x^{n+1}}{n+1} \Big|_a^b = \frac{b^{n+1} - a^{n+1}}{n+1} \quad (n \neq -1)$$

Example 4.

$$\int_a^b e^x dx = e^x \Big|_a^b = e^b - e^a$$

* The expression $\Big|_a^b$ is called the sign of double substitution. In the literature we find two notations:

$$F(b) - F(a) = [F(x)]_a^b$$

or

$$F(b) - F(a) = F(x) \Big|_a^b$$

We shall use both notations.

Example 5.

$$\int_0^{2\pi} \sin x \, dx = -\cos x \Big|_0^{2\pi} = -(\cos 2\pi - \cos 0) = 0$$

Example 6.

$$\int_0^1 \frac{x \, dx}{\sqrt{1+x^2}} = \sqrt{1+x^2} \Big|_0^1 = \sqrt{2} - 1$$

11.5 CHANGE OF VARIABLE IN THE DEFINITE INTEGRAL

Theorem. *Given an integral*

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

where the function $f(x)$ is continuous on the interval $[a, b]$.

Introduce a new variable t using the formula

$$x = \varphi(t)$$

If

$$(1) \quad \varphi(\alpha) = a, \quad \varphi(\beta) = b,$$

$$(2) \quad \varphi(t) \text{ and } \varphi'(t) \text{ are continuous on } [\alpha, \beta],$$

$$(3) \quad f[\varphi(t)] \text{ is defined and is continuous on } [\alpha, \beta], \text{ then}$$

$$\int_a^b f(x) \, dx = \int_{\alpha}^{\beta} f[\varphi(t)] \varphi'(t) \, dt \quad (1)$$

Proof. If $F(x)$ is an antiderivative of the function $f(x)$, we can write the following equations:

$$\int f(x) \, dx = F(x) + C \quad (2)$$

$$\int f[\varphi(t)] \varphi'(t) \, dt = F[\varphi(t)] + C \quad (3)$$

The truth of the latter equation is checked by differentiation of both sides with respect to t . [It likewise follows from formula (2), Sec. 10.4]. From (2) we have

$$\int_a^b f(x) \, dx = F(x) \Big|_a^b = F(b) - F(a)$$

From (3) we have

$$\begin{aligned} \int_{\alpha}^{\beta} f[\varphi(t)] \varphi'(t) \, dt &= F[\varphi(t)] \Big|_{\alpha}^{\beta} \\ &= F[\varphi(\beta)] - F[\varphi(\alpha)] \\ &= F(b) - F(a) \end{aligned}$$

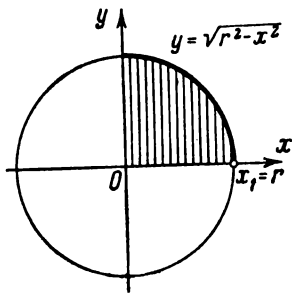


Fig. 221

The right sides of these expressions are equal, and so the left sides are equal as well, thus proving the theorem.

Note. It will be noted that when computing the definite integral from formula (1) we do not return to the old variable. If we compute the second of the definite integrals of (1), we get a certain number; the first integral is also equal to this number.

Example. Compute the integral

$$\int_0^r \sqrt{r^2 - x^2} dx$$

Solution. Make a change of variable:

$$x = r \sin t, \quad dx = r \cos t dt$$

Determine the new limits:

$$x = 0 \quad \text{for} \quad t = 0$$

$$x = r \quad \text{for} \quad t = \frac{\pi}{2}$$

Consequently,

$$\begin{aligned} \int_0^r \sqrt{r^2 - x^2} dx &= \int_0^{\frac{\pi}{2}} \sqrt{r^2 - r^2 \sin^2 t} r \cos t dt = r^2 \int_0^{\frac{\pi}{2}} \sqrt{1 - \sin^2 t} \cos t dt \\ &= r^2 \int_0^{\frac{\pi}{2}} \cos^2 t dt = r^2 \int_0^{\frac{\pi}{2}} \left(\frac{1}{2} + \frac{1}{2} \cos 2t \right) dt = r^2 \left[\frac{t}{2} + \frac{\sin 2t}{4} \right]_0^{\frac{\pi}{2}} = \frac{\pi r^2}{4} \end{aligned}$$

Geometrically, the computed integral is $\frac{1}{4}$ the area of the circle bounded by the circumference $x^2 + y^2 = r^2$ (Fig. 221).

11.6 INTEGRATION BY PARTS

Let u and v be differentiable functions of x . Then

$$(uv)' = u'v + uv'$$

Integrating both sides of the identity from a to b , we have

$$\int_a^b (uv)' dx = \int_a^b u'v dx + \int_a^b uv' dx \quad (1)$$

Since $\int (uv)' dx = uv + C$, we have $\int_a^b (uv)' dx = uv|_a^b$; for this reason, the equation can be written in the form

$$uv|_a^b = \int_a^b v du + \int_a^b u dv$$

or, finally,

$$\int_a^b u dv = uv|_a^b - \int_a^b v du$$

Example. Evaluate the integral $I_n = \int_0^{\frac{\pi}{2}} \sin^n x dx$.

$$\begin{aligned} I_n &= \int_0^{\frac{\pi}{2}} \sin^n x dx = \int_0^{\frac{\pi}{2}} \sin^{n-1} x \sin x dx = - \int_0^{\frac{\pi}{2}} \underbrace{\sin^{n-1} x}_u \underbrace{d \cos x}_{dv} \\ &= -\sin^{n-1} x \cos x \Big|_0^{\frac{\pi}{2}} + (n-1) \int_0^{\frac{\pi}{2}} \sin^{n-2} x \cos x \cos x dx \\ &= (n-1) \int_0^{\frac{\pi}{2}} \sin^{n-2} x \cos^2 x dx \\ &= (n-1) \int_0^{\frac{\pi}{2}} \sin^{n-2} x (1 - \sin^2 x) dx \\ &= (n-1) \int_0^{\frac{\pi}{2}} \sin^{n-2} x dx - (n-1) \int_0^{\frac{\pi}{2}} \sin^n x dx \end{aligned}$$

In the notation chosen we can write the latter equation as

$$I_n = (n-1) I_{n-2} - (n-1) I_n$$

whence we find

$$I_n = \frac{n-1}{n} I_{n-2} \quad (2)$$

Using the same technique, we find

$$I_{n-2} = \frac{n-3}{n-2} I_{n-4}$$

and so

$$I_n = \frac{n-1}{n} \frac{n-3}{n-2} I_{n-4}$$

Continuing in the same way, we arrive at I_0 or I_1 depending on whether the number n is even or odd.

Let us consider two cases:

(1) n is even, $n=2m$:

$$I_{2m} = \frac{2m-1}{2m} \cdot \frac{2m-3}{2m-2} \cdots \frac{3}{4} \cdot \frac{1}{2} I_0$$

(2) n is odd, $n=2m+1$:

$$I_{2m+1} = \frac{2m}{2m+1} \cdot \frac{2m-2}{2m-1} \cdots \frac{4}{5} \cdot \frac{2}{3} I_1$$

but since

$$I_0 = \int_0^{\frac{\pi}{2}} \sin^0 x \, dx = \int_0^{\frac{\pi}{2}} dx = \frac{\pi}{2}, \quad I_1 = \int_0^{\frac{\pi}{2}} \sin x \, dx = 1$$

we have

$$I_{2m} = \int_0^{\frac{\pi}{2}} \sin^{2m} x \, dx = \frac{2m-1}{2m} \cdot \frac{2m-3}{2m-2} \cdots \frac{5}{6} \cdot \frac{3}{4} \cdot \frac{1}{2} \cdot \frac{\pi}{2}$$

$$I_{2m+1} = \int_0^{\frac{\pi}{2}} \sin^{2m+1} x \, dx = \frac{2m}{2m+1} \cdot \frac{2m-2}{2m-1} \cdots \frac{6}{7} \cdot \frac{4}{5} \cdot \frac{2}{3}$$

From these formulas there follows the *Wallis formula*, which expresses the number $\frac{\pi}{2}$ in the form of an infinite product.

Indeed from the latter two equations we find, by means of termwise division,

$$\frac{\pi}{2} = \left(\frac{2 \cdot 4 \cdot 6 \cdots 2m}{3 \cdot 5 \cdots (2m-1)} \right)^2 \frac{1}{2m+1} \frac{I_{2m}}{I_{2m+1}} \quad (3)$$

We shall now prove that

$$\lim_{m \rightarrow \infty} \frac{I_{2m}}{I_{2m+1}} = 1$$

For all x of the interval $\left(0, \frac{\pi}{2}\right)$, the inequalities

$$\sin^{2m-1} x > \sin^{2m} x > \sin^{2m+1} x$$

hold.

Integrating from 0 to $\frac{\pi}{2}$, we get

$$I_{2m-1} \geq I_{2m} \geq I_{2m+1}$$

whence

$$\frac{I_{2m-1}}{I_{2m+1}} \geq \frac{I_{2m}}{I_{2m+1}} \geq 1 \quad (4)$$

From (2) it follows that

$$\frac{I_{2m-1}}{I_{2m+1}} = \frac{2m+1}{2m}$$