Integration

alexander

July 30, 2025

antiderivatives a function $F(x)$ is an antiderivative of $f(x)$ on (a, b) if $F'(x) = f(x)$ for all $x \in (a, b)$.

general antiderivative let $F(x)$ be an antiderivative of $f(x)$ on (a, b). then every other antiderivative on (a, b) is of the form $F(x) + C$ for some constant C.

indefinite integral the notion $\int f(x)\,dx = F(x) + C$ means that $F'(x) = f(x)$ we say that $F(x) + C$ is the antiderivative or indefinite integral of $f(x)$.

power rule for integrals $\int x^n\,dx = \frac{x^{n+1}}{n+1} + C$ for $n \neq -1$

linearity of the indefinite integral:

- $\int (f(x) + g(x))\,dx = \int f(x)\,dx + \int g(x)\,dx$

- $\int cf(x)\,dx = c \int f(x)\,dx$

basic trigonometric integrals:

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

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

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

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

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

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

for the rest of this section assume that $f(x)$ is continuous and positive, so that the graph of $f(x)$ lies above the x-axis. as a first step, we approximate the area using rectangles. first, choose a whole number N and divide $[a, b]$ into N subintervals of equal width, each subinterval has width $\Delta x = \frac{b-a}{N}$ since $[a, b]$ has width $b - a$. the right endpoints of the subintervals are $a + \Delta x, a + 2\Delta x, \ldots, a + (N-1)\Delta x, a + N\Delta x$. notice that the last right endpoint is b because $a + N\Delta x = a + N(\frac{b-a}{N}) = b$. next, above each subinterval, construct the rectangle whose height is the value of $f(x)$ at the right endpoint of the subinterval. each rectangle has width Δx. the height of the first rectangle is $f(a + \Delta x)$ and its area is $f(a + \Delta x)\Delta x$. similarly, the second rectangle

has a height $f(a + 2\Delta x)$ and area of $f(a + 2\Delta x)\Delta x$, ect. the sum of the areas of the rectangles is $f(a + \Delta x)\Delta x + f(a + 2\Delta x)\Delta x + \ldots + f(a + N\Delta x)\Delta x$. this sum is called the Nth right-endpoint approximation and is denoted R_N as $R_N = \Delta x[f(a + \Delta x) + f(a + 2\Delta x) + \ldots + f(a + N\Delta x)]$. in words, R_N is equal to Δx times the sum of the function values at the right endpoints of the subintervals.

linearity of summations

- $\sum_{j=m}^{n}(a_j + b_j) = \sum_{j=m}^{n} a_j + \sum_{j=m}^{n} b_j$

- $\sum_{j=m}^{n} Ca_j = C \sum_{j=m}^{n} a_j$ (C any constant)

- $\sum_{j=1}^{n} k = nk$ (k for any constant and $n \geq 1$)

$R_N = \Delta x[f(a + \Delta x) + f(a + 2\Delta x) + \ldots + f(a + N\Delta x)]$
$R_N = \Delta x \sum_{j=1}^{N} f(a + j\Delta x)$
$L_N = \Delta x[f(a) + f(a + \Delta x) + f(a + 2\Delta x) + \ldots + f(a + (N-1)\Delta x)]$
$L_N = \Delta x \sum_{j=0}^{N-1} f(a + j\Delta x)$
$M_N = \Delta x \sum_{j=1}^{N} f(a + (j - \frac{1}{2})\Delta x)$

if $f(x)$ is continuous on $[a, b]$, then the endpoint and midpoint approximations approach one and the same limit as as $N \to \infty$. in other words, there is a value L such that $\lim_{N \to \infty} R_N = \lim_{N \to \infty} L_N = \lim_{N \to \infty} M_N = L$

power sums
the kth power sum is the sum of the kth powers of the first N integers.

- $\sum_{j=1}^{N} j = 1 + 2 + \ldots + N = \frac{N(N+1)}{2} = \frac{N^2}{2} + \frac{N}{2}$

- $\sum_{j=1}^{N} j^2 = 1^2 + 2^2 + \ldots + N^2 = \frac{N(N+1)(2N+1)}{6} = \frac{N^3}{3} + \frac{N^2}{2} + \frac{N}{6}$

- $\sum_{j=1}^{N} j^3 = 1^3 + 2^3 + \ldots + N^3 = \frac{N^2(N+1)^2}{4} = \frac{N^4}{4} + \frac{N^3}{2} + \frac{N^2}{4}$

in the endpoint and midpoint approximations, we approximate the area under the graph by rectangles of equal width. the heights of the rectangles are the values of $f(x)$ at the endpoints or midpoints of the subintervals. in riemann sum approximations, we relax these requirements; the rectangles need not have equal width and height of a rectangle may be any value of $f(x)$ within the subinterval. more formally, to define a riemann sum, we choose a partition and a set of intermediate points. a partition P of length N is any choice of points in $[a, b]$ that divides the interval into N subintervals:

- partition P: $a = x_0 < x_1 < x_2 \ldots < x_N = b$

- subintervals: $[x_0, x_1], [x_1, x_2], \ldots [x_{N-1}, x_N]$

we denote the length of the ith interval by $\Delta x_i = x_i - x_{i-1}$

the maximum of the lengths Δx_i is called the norm of the partition P and is denoted $\|P\|$. a set of intermediate points is a set of points $C = c_1, \ldots, c_N$, where c_i belongs to $[x_{i-1}, x_i]$

now, over each subinterval $[x_{i-1}, x_i]$ we constuct the rectangle of height $f(c_i)$ and base Δx_i. this rectangle has area $f(c_i)\Delta x_i$. if $f(c_i) < 0$, the rectangle extends below the x-axis and the "area" $f(c_i)\Delta x_i$ is negative. the riemann sum is the sum of these positive or negative areas:

$$R(f, P, C) = \sum_{i=1}^{N} f(c_i)\Delta x_i = f(c_1)\Delta x_1 + f(c_2)\Delta x_2 + \ldots + f(c_N)\Delta x_N$$

the endpoint and midpoint approximations are examples of riemann sums. they correspond to the partition of $[a, b]$ into N subintervals of equal length $\Delta x = \frac{b-a}{N}$ and the choice of right endpoints, left endpoints, or midpoints as intermediate points.

having introduced riemann sums, we are ready to make the following definition: a function is integrable of $[a, b]$ if all of the riemann sums (not just the endpoint and midpoint approximations) approach one and the same limit L as the norm of the partition tends to zero. more formally, we write

$$L = \lim_{\|P\|\to 0} R(f, P, C) = \lim_{\|P\|\to 0} \sum_{i=1}^{N} f(c_i)\Delta x_i$$

if $|R(f, P, C) - L|$ gets arbitrarily small as the norm $\|P\|$ tends to zero, no matter how we choose the partition and intermediate points. note that as $\|P\| \to 0$, the number N of intervals tends to ∞. the limit L is called the definite integral of $f(x)$ over $[a, b]$.

definite integral
the definite integral of $f(x)$ over $[a, b]$ is the limit of riemann sums and is denoted by the integral sign:

$$\int_a^b f(x)\, dx = \lim_{\|P\|\to 0} R(f, P, C) = \lim_{\|P\|\to 0} \sum_{i=1}^{N} f(c_i)\Delta x_i$$

when this limit exists, we say that $f(x)$ is integrable over $[a, b]$.

remarks

- we often refer to the definite integral more simply as the integral of f over $[a, b]$.

- the function $f(x)$ inside the integral sign is called integrand

- the numbers a and b representing the interval $[a, b]$ are called the limits of integration.

- any variable may be used as a variable of integration (this is a "dummy" variable). thus, the following three integrals all denote the same quantitiy: $\int_a^b f(x)\,dx$, $\int_a^b f(t)\,dt$, $\int_a^b f(u)\,du$

- recall that the indefinite integral $\int f(x)\,dx$ denotes the general antiderivative of $f(x)$

most functions that arise in practice are integrable. in particular, every continuous function is integrable, as stated in the following theorem.
if $f(x)$ is continuous on $[a, b]$, then $f(x)$ is integrable over $[a, b]$.

when $f(x)$ is continuous and positive, we interpret the definite integral as the area under the graph. if $f(x)$ is not positive, then the definite integral is not equal to an area in the usual sense, but we may interpret is as the signed area between the graph and the x-axis. signed area is defined as following:
signed area of a region = (area above x-axis) - (area below x-axis)

thus, signed area treats regions under the x-axis as "negative area". to see why the definit eintegral gives us signed area, suppose first that $f(x)$ is negative on $[a, b]$ and consider a riemann sum: $R(f, C, P) = f(c_1)\Delta x_1 + f(c_2)\Delta x_2 + \ldots + f(c_N)\Delta x_N$ since $f(c_i) < 0$, each term is equal to the negative of the area of a rectangle: $f(c_i)\Delta x_i = -(\text{area of rectangle of height } |f(c_i)|)$

therefore, the riemann sums $R(f, C, P)$ converge to the negative of the area between the graph and the x-axis. if $f(x)$ takes on both positive and negative values, then $f(c_i)\Delta x_i$ is positive or negative, depending on whether the corresponding rectangle lies above or below the x-axis. in this case, the riemann sums converge to the signed area. in summary,
$\int_a^b f(x)\,dx$ = signed area of region between the graph and x-axis over $[a, b]$

integral of a constant for any constant C, $\int_a^b C\,dx = C(b - a)$.

linearity of the definite integral if $f(x)$ and $g(x)$ are integrable over $[a, b]$, then

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

- $\int_a^b C f(x)\,dx = C \int_a^b f(x)\,dx$ for any constant C

reversing the limits of integration for $a < b$, we set $\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx$

additivity for adjacent intervals for $a \le b \le c$,
$\int_a^c f(x)\,dx = \int_a^b f(x)\,dx + \int_b^c f(x)\,dx$

comparison theorem if $g(x) \le f(x)$ on an interval $[a, b]$, then
$\int_a^b g(x)\,dx \le \int_a^b f(x)\,dx$

4

the fundamental theorem of calculus, part I
assume that $f(x)$ is continuous on $[a,b]$ and let $F(x)$ be an antiderivative of $f(x)$ on $[a,b]$. then $\int_a^b f(x)\,dx = F(b) - F(a)$

part I of the fundamental theorem shows that we can compute definite integrals using antiderivatives. part II turns this relationship around: it tells us that we can use the definite integral to construct antiderivatives. to state part II, we introduce the area function (or cumulative area function) associated to a function $f(x)$ and lower limit a: $A(x) = \int_a^x f(t)\,dx =$ signed area from a to x in essence, we turn the definite integral into a function by treating the upper limit x as a variable rather than a constant. notice that $A(a) = 0$ because $A(a) = \int_a^a f(t)\,dt = 0$. although $A(x)$ is defined by an integral, in many cases we can find an explicit formula it.

fundamental theorem of calculus, part II
let $f(x)$ be a continuous function on $[a,b]$. then $A(x) = \int_a^x f(t)\,dt$ is an antiderivative of $f(x)$, that is, $A'(x) = f(x)$, or equivalently,
$\frac{d}{dx}\int_a^x f(t)\,dt = f(x)$
furthermore, $A(x)$ satisfies the initial condition $A(a) = 0$.

integration (antidifferentiation) is generally more difficult than differentiation. there are no sure-fire methods for finding an antiderivative, and, in some cases, there is no elementary formula for the antiderivative. however, there are several general techniques that are widely applicable. one such technique is the substitution method, which uses the chain rule "in reverse".

substitution method if $F'(x) = f(x)$, then
$\int f(u(x))u'(x)\,dx = F(u(x)) + C$

say you're given an integral like this: $\int f(g(x)) \cdot g'(x)\,dx$ let $u = g(x)$ then $du = g'(x)dx$ the integral becomes: $\int f(u)\,du$

there are two ways of using substitution to compute definite integrals. one way is to carry out the substitution method fully to find an antiderivative in terms of the x-variable. however, it is often more efficient to evaluate the definite integral directly in terms of the u-variable. this can be done provided that the limits of integration are changed as indicated in the next theorem.

change of variables formula for definite integrals if $u(x)$ is differentiable on $[a,b]$ and $f(x)$ is integrable on the range of $u(x)$, then
$\int_a^b f(u(x))u'(x)\,dx = \int_{u(a)}^{u(b)} f(u)\,du$

area between the graphs $= \int_a^b (y_{\text{top}} - y_{\text{bot}})\,dx = \int_a^b (f(x) - g(x))\,dx$

area between the graphs $= \int_c^d (g_2(y) - g_1(y))\,dy = \int_c^d (x_{\text{right}} - x_{\text{left}})\,dy$

volume as the integral of cross-sectional area
suppose that a solid body extends from height $y = a$ to $y = b$. let $A(y)$ be the area of the horizontal cross section at height y. then
volume of the solid body $= \int_a^b A(y)\,dy$

recall that the average of N numbers $a_1, a_2, \ldots, a_N$ is the sum divided by N:
$\frac{a_1 + a_2 + \ldots a_N}{N} = \frac{1}{N} \sum_{j=1}^{N} a_j$

we cannot define the average value of a function $f(x)$ on an interval $[a, b]$ as a sum because there are infinitely many values of x to consider. however, the right-endpoint appproximation may be interpreted as an average value:
$R_N = \frac{b-a}{N}(f(x_1) + f(x_2) + \ldots + f(x_N))$
where $x_i = a + i(\frac{b-a}{N})$. dividing by $(b-a)$, we obtain the average of the function values $f(x_i)$:
$\frac{1}{b-a} R_N = \frac{f(x_1) + f(x_2) + \ldots + f(x_N)}{N}$
if N is large, it is reasonable to think of this quantity as an approximation to the average of $f(x)$ on $[a, b]$. therefore, we define the average value itself as the limit:

average value $= \lim_{N \to \infty} \frac{1}{b-a} R_N(f) = \frac{1}{b-a} \int_a^b f(x)\,dx$ the average value is also called the mean value.

average value the average value of an integrable function $f(x)$ on $[a, b]$ is the quantity
average value $= \frac{1}{b-a} \int_a^b f(x)\,dx$

mean value theorem for integrals if $f(x)$ is continuous on $[a, b]$, then there exists a value $c \in [a, b]$ such that
$f(c) = \frac{1}{b-a} \int_a^b f(x)\,dx$

a solid of revolution is a solid obtained by rotating a region in the plane about an axis. the sphere and right circular cone are familiar examples of such solids.

area of a circle $= \pi r^2$
volume $= \int_a^b (\text{area of slice})(\text{thickness}) = \pi \int_a^b [f(x)]^2\,dx$

volume of a solid of revolution: disk method if $f(x)$ is continuous and $f(x) \geq 0$ on $[a, b]$, then the volume V obtained by rotating the region under the graph about the x-axis is [with $R = f(x)$]:
$V = \pi \int_a^b R^2\,dx = \pi \int_a^b f(x)^2\,dx$

we now consider some variations on the formula for a volume of revolution. first suppose that the region between two curves $y = f(x)$ and $y = g(x)$, where $f(x) \geq g(x) \geq 0$, is rotated about the x-axis. then the vertical cross section of the solid at x is generated as the segment $\overline{AB}$ revolves around the x-axis. this

cross section is a washer of outer radius $R = f(x)$ and inner radius $r = g(x)$. the area of this washer is $\pi R^2 - \pi r^2$ or $\pi(f(x)^2 - g(x)^2)$, and we obtain the volume of the solid as the integral of the cross-section area:
$V = \pi \int_a^b (R^2 - r^2)\, dx = \pi \int_a^b (f(x)^2 - g(x)^2)\, dx$

we computed volumes by integrating cross-section area. the shell method is based on a different idea and is more convenient in some cases. in the shell method, we use cylindrical shells to approximate a volume of revolution. to this end, let us first derive an approximation to the volume of a cylindrical shell of height h, outer radius R, and inner radius r. since the shell is obtained by removing a cylinder of radius r from the wider cylinder of radius R, its volume is equal to $\pi R^2 h - \pi r^2 h = \pi h(R^2 - r^2) = \pi h(R + r)(R - r) = \pi h(R + r)\Delta r$ where $\Delta r = R - r$ is the shell's thickness. if the shell is very thin, then R and r are nearly equal and we may replace $(R + r)$ by $2R$ in the equation above to obtain the approximation volume of shell $\approx 2\pi R h \Delta r$

now consider a solid obtained by rotating the region under $y = f(x)$ from $x = a$ to $x = b$ around the y-axis. the idea is to divide the solid into this concentric shells. more precisely, we divide $[a, b]$ into N subintervals of length $\Delta x = \frac{b-a}{N}$ with endpoints $x_0, x_1, \ldots, x_N$. when we rotate the thin strip of area above $[x_{i-1}, x_i]$ about the y-axis, we obtain a thin shell whose volume we denote by V_i. the total volume V of the solid is equal to $V = \sum_{i=1}^{N} V_i$.

the top rim of the ith shell is curved. however, when Δx is small we may approximate this thin shell by the cylindrical shell (with a flat rim) of height $f(x_i)$. $V_i \approx (\text{circumference})(\text{height})(\text{thickness}) = 2\pi x_i f(x_i)\Delta x$ therefore,
$V = \sum_{i=1}^{N} V_i \approx \sum_{i=1}^{N} 2\pi x_i f(x_i)\Delta x$

volume of a solid of revolution: the shell method the volume V of the solid obtained by rotating the region under the graph of $y = f(x)$ over the interval $[a, b]$ about the y-axis is equal to $V = 2\pi \int_a^b x f(x)\, dx$

an exponential function is a function of the form $f(x) = b^x$, where $b > 0$ and $b \neq 1$. the number b is called the base. some examples are 2^x, $(1.4)^x$, and 10^x. we exclude the case $b = 1$ because $f(x) = 1^x$ is a constant function. calculators give good decimal approximations to values of exponential functions: $2^4 = 16$, $2^{-3} = 0.125$, $(1.4)^3 = 2.744$, $10^{4.6} \approx 39,810.717$. three properties of exponential functions should be singled out from the start:

- exponential functions are positive: $b^x > 0$ for all x

- $f(x) = b^x$ is increasing if $b > 1$ and decreasing if $0 < b < 1$

- the range of $f(x) = b^x$ is the set of all positive real numbers

another important feature of $f(x) = b^x$ (for $b > 1$) is that it increases rapidly. although "rapid increase" is a subjective term, the following precise statement is true: $f(x) = b^x$ increases more rapidly than every polynomial function. for example, $f(x) = 3^x$ eventually overtakes and increases faster than the power functions x^3, x^4, and x^5.

laws of exponents $(b > 0)$

- $b^0 = 1$

- $b^x b^y = b^{x+y}$

- $\frac{b^x}{b^y} = b^{x-y}$

- $b^{-x} = \frac{1}{b^x}$

- $(b^x)^y = b^{xy}$

- $b^{\frac{1}{n}} = \sqrt[n]{b}$

inverse
let $f(x)$ have a domain D and range R. the inverse function $f^{-1}(x)$ (if it exists) is the function with domain R such that $f^{-1}(f(x)) = x$ for $x \in D$ and $f(f^{-1}(x)) = x$ for $x \in R$ if f^{-1} exists, then f is called invertible.

one-to-one functions a function $f(x)$ is one-to-one (on its domain D) if for every value c, the equation $f(x) = c$ has at most one solution for $x \in D$

existence of inverses if $f(x)$ is one-to-one on its domain D, then f is invertible. furthermore,

- domain of f = range of f^{-1}

- range of f = domain of f^{-1}

horizontal line test a function $f(x)$ is one-to-one if and only if every horizontal line intersects the graph of f in at most one point.

derivative of the inverse assume that $f(x)$ is differentiable and one-to-one with inverse $g(x) = f^{-1}(x)$. if b belongs to the domain of $g(x)$ and $f'(g(b)) \neq 0$, then $g'(b)$ exists and $g'(b) = \frac{1}{f'(g(b))}$

the logarithm to the base b (where $b > 0$ and $b \neq 1$) is the inverse of the exponential function $f(x) = b^x$ and is denoted $\log_b x$. by definition, $b^{\log_b x} = x$ and $\log_b b^x = x$.

laws of logarithms

- $\log_b(1) = 0$

- $\log_b(b) = 1$

- $\log_b(xy) = \log_b x + \log_b y$

- $\log_b \frac{x}{y} = \log_b x - \log_b y$

- $\log_b(\frac{1}{x}) = -\log_b(x)$

- $\log_b(x^n) = n \log_b x$

we also note that all logarithm functions are proprtional. more precisely, the following change of base formula holds: $\log_b x = \frac{\log_a x}{\log_a b}$, $\log_b x = \frac{\ln(x)}{\ln(b)}$

derivative of $f(x) = \ln(x)$ $\frac{d}{dx} \ln(x) = \frac{1}{x}$ for $x > 0$

antiderivative of $y = \frac{1}{x}$ the function $f(x) = \ln|x|$ is an antiderivative of $y = \frac{1}{x}$ in the domain $x : x \neq 0$, that is, $\int \frac{dx}{x} = \ln|x| + C$

in this section we study quantities $P(t)$ that depend exponentially on time:

$$P(t) = P_0 e^{kt}$$

where P_0 and k are constants. if $k > 0$, then $P(t)$ grows exponentially with growth constant k, and if $k < 0$, then $P(t)$ decays exponentially with decay constant k. the coefficient P_0 is the initial size since $P(0) = P_0 e^{k \cdot 0} = P_0$. exponential growth or decay may also be expressed by the formula $P(t) = P_0 b^t$ with $b = e^k$, since $P_0 b^t = P_0 (e^k)^t = P_0 e^{kt}$.

the exponential functions $y = P_0 e^{kt}$ are the only functions that satisfy this differential equation. theorem 1 if $y(t)$ is differentiable function satisfying the differential equation $y' = ky$ then $y(t) = P_0 e^{kt}$, where P_0 is the initial value $P_0 = y(0)$

doubling time if $P(t) = P_0 e^{kt}$ with $k > 0$, then the doubling time of P is: doubling time $= \frac{\ln 2}{k}$ if $P(t) = P_0 e^{-kt}$ then this is the halfing time.

1. $P(t) = 2P_0$

2. $P(t) = P_0 e^{kt}$

3. $2P_0 = P_0 e^{kt}$

4. $2 = e^{kt}$

5. $\ln(2) = kt$

6. $t = \frac{\ln(2)}{k}$

exponential functions play an important role in financial calculations. in this section,we use these functions to compute compound interest and present value. the initial sum of money P_0 in an account or investment is called the principal. when a principal of P_0 dollars is deposited into a bank account, the amount or balance in the account at time t depends on two factors: the interest rate r and how often interest is compounded. if interest is paid out once a year at the end of the year, we say that the interest is compounded annually. if the principal and interest are retained in the account, then the balance grows exponentially.

compound interest if P_0 dollars are deposited into an account earning interest at an annual rate r, compounded M times yearly, then the value of the account after t years is $P(t) = P_0(1 + \frac{r}{M})^{Mt}$.

limit formula for e and e^x
$e = \lim n \to \infty (1 + \frac{1}{n})^n$ and $e^x = \lim_{n \to \infty}(1 + \frac{x}{n})^n$ for all x

continuously compounded interest if P_0 dollars are deposited into an account earning interest at an annual rate r, compounded continuously, then the value of the account after t years is $P(t) = P_0 e^{rt}$ the concept of present value (PV) is used in finance to compare the values of payments made at different times. assume that there is an interest rate r at which any investor can lend or borrow money and that interest is compounded continuously. then the PV of P dollars to be received t years in the future is defined to be Pe^{-rt}: The PV of P dollars received at time t is Pe^{-rt} what is the reasoning behind this definition? when you invest at the interest rate r for t years, your principal increases by a factor e^{rt}, so you invest Pe^{-rt}, dollars, your principal grows to $Pe^{-rt}e^{rt} = P$ dollars at time t. thus, the present value Pe^{-rt} is the amount you would have to invest today in order to have P dollars at time t

PV of an income stream the present value at interest rate r of an income stream paying out $R(t)$ dollars/year continuously for T years is $PV =$

$\int_0^T R(t)e^{-rt}\,dt$

we have seen that a quantity grows or decays exponentially if its rate of change is proportional to the amount present. this characteristic property is expressed by the differential equation $y' = ky$. we now study a closely related differential equation: $\frac{dy}{dt} = k(y - b)$ where k and b are constants and $k \neq 0$. this differential equation describes a quantity y whose rate of change is proportional to the size of the difference $y - b$. the general solution is $y(t) = b + Ce^{kt}$ where C is a constant determined by the initial condition. to very this, we note that $(y - b)' = y'$ since b is a constant, so the first equation may be rewritten $\frac{d}{dt}(y - b) = k(y - b)$ in other words, $y - b$ satisfies the differential equation of an exponential function. thus, $y - b = Ce^{kt}$ or $y = b + Ce^{kt}$, as claimed.

L'Hôpital's rule assume that $f(x)$ and $g(x)$ are differentiable on an open interval containing a and that $f(a) = g(a) = 0$ also assume that $g'(x) \neq 0$ for x near but not equal to a. then

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

provided that the limit on the right exists. this conclusion also holds if $f(x)$ and $g(x)$ are differentiable for x near (but not equal to) a and
$\lim_{x \to a} f(x) = \pm\infty$ and $\lim_{x \to a} f(x) = \pm\infty$ furthermore, these limits may be replaced by one-sided limits.

sometimes, we are interested in determining which of two functions, $f(x)$ and $g(x)$, grows faster. for example, there are two standard computer algorithms for sorting data (alphabetizing, ordering according to rank, ect.): quick sort and bubble sort. the average time equired to sort a list of size n has order of magnitude $n \ln(n)$ for quick sort and n^2 for bubble sort. which algorithm is faster when the size n is large? although n is a whole number, this problem amounts to comparing the growth of $f(x) = x \ln(x)$ and $g(x) = x^2$ as $x \to \infty$. we say that $f(x)$ grows faster than $g(x)$ if $\lim x \to \infty \frac{f(x)}{g(x)} = \infty$ or equivalently,
$\lim x \to \infty \frac{g(x)}{f(x)} = 0$ to compare the growth of functions, we need a version of L'Hôpital's rule that applies to limits as $x \to \infty$.

L'Hôpital's rule for limits as $x \to \infty$ assume that $f(x)$ and $g(x)$ are differentiable in an interval (b, ∞) and that $g'(x) \neq 0$ for $x > b$. if $\lim_{x \to \infty} f(x)$ and $\lim_{x \to \infty} g(x)$ both exists and either both are zero or both are infinite, then $\lim_{x \to \infty} \frac{f(x)}{g(x)} = \lim_{x \to \infty} \frac{f'(x)}{g'(x)}$ provided that the limit on the right exists. a similar result holds for limits as $x \to -\infty$.

$\theta = \sin^{-1}(x)$ is the unique angle in $[-\frac{\pi}{2}, \frac{\pi}{2}]$ such that $\sin(\theta) = x$
$\theta = \cos^{-1}(x)$ is the unique angle in $[0, \pi]$ such that $\cos(\theta) = x$

derivatives of arcsine and arccosine

$\frac{d}{dx}\sin^{-1}(x) = \frac{1}{\sqrt{1-x^2}}$, $\frac{d}{dx}\cos^{-1}(x) = -\frac{1}{\sqrt{1-x^2}}$

derivatives of inverse trigonometric functions

- $\frac{d}{dx}\tan^{-1}(x) = \frac{1}{\sqrt{1+x^2}}$

- $\frac{d}{dx}\cot^{-1}(x) = -\frac{1}{\sqrt{1+x^2}}$

- $\frac{d}{dx}\sec^{-1}(x) = \frac{1}{|x|\sqrt{1-x^2}}$

- $\frac{d}{dx}\csc^{-1}(x) = -\frac{1}{|x|\sqrt{1-x^2}}$

the hyperbolic functions are certain special combinations of e^x and e^{-x} that play a role in engineering and physics. the hyperbolic sine and cosine, often called "cinch" and "cosh", are the functions defined as follows:

$\sinh(x) = \frac{e^x - e^{-x}}{2}$, $\cosh(x) = \frac{e^x + e^{-x}}{2}$

as the terminology suggest, there are similarities between the hyperbolic and trigonometric functions. the trigonometric functions and their hyperbolic analogs have the same parity., thus, sin x and sinh x are both odd, and cos x and cosh x are both even. the basic trigonometric identity $\sin^2(x) + \cos^2(x) = 1$ has a hyperbolic analog: $\sinh^2(x) - \cosh^2(x) = 1$. the addition formulas satisfied by $\sin(\theta)$ and $\cos(\theta)$ have hyperbolic analogs:

$\sinh(A + B) = \sinh(A)\cosh(B) + \cosh(A)\sinh(B)$
$\cosh(A + B) = \cosh(A)\cosh(B) + \sinh(A)\sinh(B)$

it follows from the definitions that $\cosh(x)+\sinh(x) = e^x$, $\cosh(x)-\sinh(x) = e^x$
we obtain the following by multiplying these two equation together:
$\cosh^2(x) - \sinh^2(x) = (\cosh(x) + \sinh(x))(\cosh(x) - \sinh(x)) = e^x \cdot e^{-x} = 1$

hyperbola instead of the circle: the identity $\sin^2(t) + \cos^2(t) = 1$ tells us that the point $(\cos(t), \sin(t))$ lies on the unit circle $x^2 + y^2 = 1$. similarly, the identity $\cosh^2(t) - \sinh^2(t) = 1$ says that the point $(\cosh(t), \sinh(t))$ lies on the hyperbola $x^2 - y^2 = 1$

derivative formulas: the formulas for derivatives of the hyperbolic functions are similar to the formulas for the corresponsing trigonometric function, differing at most by a minus sign: $\frac{d}{dx}\sinh(x) = \cosh(x)$, $\frac{d}{dx}\cosh(x) = \sinh(x)$

the terms "hyperbolic sine" and "hyperbolic cosine" suggest a connection between the hyperbolic and trigonometric functions. this excursion explored the source of the connection, which leads us to complex numbers and Euler's famous formula.

recall that $y = e^t$ satisfies the differential equation $y' = y$ and every solution is of the form $y = Ce^t$ for some constant C. on the other hand, we can check that both $y = e^t$ and $y = e^{-t}$ satisfy the second-order differential equation $y'' = y$ indeed, $(e^t)'' = e^t$ and $(e^{-t})'' = (-e^{-t}) = e^{-t}$. it can be shown that every solution of this equation has the form $y = Ae^t + Be^{-t}$ for some constants A and B. now let's examine what happens when we change the equation by a minus sign: $y'' = -y$ in this case, $y = \sin(t)$ and $y = \cos(t)$ are solutions because $(\sin(t))'' = (\cos(t))' = -\sin(t)$, $(\cos(t))'' = (-\sin(t))' = -\cos(t)$ furthermore, it can be proved as before that every solution of the equation has the form $y = A\cos(t) + B\sin(t)$ this might seem to be the end of the story. however, there is another way to write down solutions to the equation. consider the exponential functions $y = e^{it}$ and $y = e^{-it}$, where $i = \sqrt{-1}$

the number i is an imaginary complex number satisfying $i^2 = -1$. since i is not a real number, the exponential e^{it} is not defined without further explanation. but if we accept that e^{it} can be defined and that it is legitimate to apply the usual rules of calculus to it, we obtain $\frac{d}{dt}e^{it} = ie^{it}$ differentiating a second time yields $(e^{it})'' = (ie^{it})' = i^2 e^{it} = -e^{it}$ in other words, $y = e^{it}$ is a solution to the equation $y'' = -y$ and must therefore be expressible in terms of $\sin(t)$ and $\cos(t)$. that is, for some constant A and B, $e^{it} = A\cos(t) + B\sin(t)$ we determine A and B by considering initial conditions. first, set $t = 0$ in the equation $1 = e^{i0} = \frac{d}{dt}e^{it} = A\cos(0) + B\sin(0) = A$ then take the derivative of the equation and set $t = 0$: $ie^{it} = \frac{d}{dx}e^{it} = A\cos'(t) + B\sin'(t) = -A\sin(t) + B\cos(t)$ $i = ie^{i0} = -A\sin(0) + B\cos(0) = B$ thus, $A = 1$ and $B = i$, and the equation yields Euler's formula:

$$e^{it} = \cos(t) + i\sin(t)$$

Euler derived this result using power series, which allows us to define e^{it} in a rigorous fashion. At $t = \pi$, Euler's formula yields $e^{i\pi} = -1$ here we have a simple but surpring relation between the three important numbers e, π, and i. euler's formula also reveals the source of the analogy between hyperbolic and trigonometric functions. let us calculate the hyperbolic cosine at $x = it$: $\cosh(it) = \frac{e^{it}+e^{-it}}{2} = \frac{\cos(t)+\sin(t)}{2} + \frac{\cos(-t)+i\sin(-t)}{2} = \cos(t)$ a similar calculation shows that $\sinh(it) = i\sin(t)$. in other words, the hyperbolic and trigonometric functions are not merly analogous - once we introduce complex numbers, we see that they are very nearly the same functions.

numerical integration is the process of approximating a definite integral using well-chosen sums of function values. we use numerical integration when we cannot find a formula for an antiderivative and thus cannot apply the FTC. for instance, the gaussian function $f(x) = e^{\frac{-x^2}{2}}$, whose graph is a bell-shaped curve, does not have an elementary antiderivative. numerical methods are needed to approximate $\int_a^b e^{\frac{-x^2}{2}}\, dx$.

let N be a whole number. the trapezodial rule T_N consists of approximating $\int_a^b f(x)\, dx$ by the average of the left-and right-endpoint approximations: $T_N = \frac{1}{2}(R_N + L_N)$ this approximation was used informally in previous chapters, where we observed that the average was likely to give a better approximation than either of R_N or L_N alone. to derive a formula for T_N, let us recall the formulas for the endpoint approximations. we divide $[a, b]$ into N subintervals of length $\Delta x = \frac{b-a}{N}$ with endpoints $x_0 = a, x_1 = a + \Delta x, x_2 = a + 2\Delta x, \ldots, x_N = b$ for convenience, we denote the values of $f(x)$ at the endpoints by y_j: $y_j = f(x_j) = f(a + j\Delta x)(j = 0, 1, 2, \ldots, N)$ in particular, $y_0 = f(a)$ and $y_N = f(b)$. then $R_N = \Delta x \sum_{j=1}^{N} y_j$, $L_N = \Delta x \sum_{j=0}^{N-1} y_j$ when we form the sum $R_N + L_N$, each value y_j occurs twice except for the first y_0 (which only occursd in L_N) and the last y_N (which only occurs in R_N). therefore, $R_N + L_N = \Delta x(y_0 + 2y_1 + 2y_2 + \ldots + 2y_{N-1} + y_N)$ dividing by 2 we obtain the formula for T_N

trapezodial rule the Nth trapezodial approximation to $\int_a^b f(x)\, dx$ is
$T_N = \frac{1}{2}\Delta x(y_0 + 2y_1 + \ldots + 2y_{N-1} + y_N)$
where $\Delta x = \frac{b-a}{N}$ and $y_j = f(a + j\Delta x)$.

midpoint rule the Nth midpoint approximation to $\int_a^b f(x)\, dx$ is
$M_N = \Delta x(f(c_1) + f(c_2) + \ldots + f(c_N))$
where $\Delta x = \frac{b-a}{N}$ and $c_j = a + (j - \frac{1}{2})\Delta x)$ is the midpoint of the jth interval $[x_{j-1}, x_j]$.

in applications, it is important to know the accuracy of a numerical approximation. we define the error in T_N and M_N by
$\text{Error}(T_N) = |T_N - \int_a^b f(x)\, dx|$, $\text{Error}(M_N) = |M_N - \int_a^b f(x)\, dx|$
accoring to the following theorem, the magnitudes of these errors are related to the size of the second derivative $f''(x)$-here we assume $f''(x)$ exists and is continuous.

error bound for T_N and M_N let K_2 be a number such that $|f''(x)| \leq K_2$ for all $x \in [a, b]$. then
$\text{Error}(T_N) \leq \frac{K_2(b-a)^3}{12N^2}$, $\text{Error}(M_N) \leq \frac{K_2(b-a)^3}{24N^2}$

can we improve on the trapezoidal and midpoint rules? one clue is that when $f(x)$ is concave (up or down), then the exact value of the integral lies between

T_N and M_N. in fact, we see geometrically that

- if $f(x)$ is concave down $\Rightarrow T_N \le \int_a^b f(x)\,dx \le M_N$

- if $f(x)$ is concave up $\Rightarrow M_N \le \int_a^b f(x)\,dx \le T_N$

this suggests that the average of T_N and M_N may be more accurate than either T_N or M_N alone.

simpson's rule S_N exploits this idea, but it takes into account that M_N is roughly twice as accurate as T_N. therefore, S_N uses a weighted average, which is more accurate than the ordinary average. for N even, we define

$S_N = \frac{1}{3}T_{\frac{N}{2}} + \frac{2}{3}M_{\frac{N}{2}}$

to derive a formula for S_N, let $\Delta x = \frac{b-a}{N}$ and, as before, set $x_j = a + j\Delta x$ and $y_j = f(a + j\Delta x)$. if we use just the even-numbered endpoints, we obtain a partition of $[a,b]$ into $\frac{N}{2}$ subintervals (keep in mind that N is even): $[x_0, x_2], [x_2, x_4], \ldots, [x_{N-2}, x_N]$
the endpoints of these intervals are used to compute $T_{\frac{N}{2}}$, and we use the midpoints $x_1, x_3, \ldots, x_{N-1}$ to compute $M_{\frac{N}{2}}$

$N_{\frac{N}{2}} = (\frac{1}{2})\frac{b-a}{\frac{N}{2}}(2y_0 + 2y_2 + 2y_4 + \ldots + 2y_{N-2} + y_N)$

$M_{\frac{N}{2}} = \frac{b-a}{\frac{N}{2}}(y_1 + y_3 + y_5 + \ldots + y_{N-1}) = \frac{b-a}{N}(2y_1 + 2y_3 + 2y_5 + \ldots + 2y_{N-1})$

thus,

$S_N = \frac{1}{3}T_{\frac{N}{2}} + \frac{2}{3}M_{\frac{N}{2}} = \frac{1}{3}\Delta x(y_0 + 2y_2 + 2y_4 + \ldots + 2y_{N-2} + y_N) + \frac{1}{3}\Delta x(4y_1 + 4y_3 + 4y_5 + \ldots + 4y_{n-1})$

simpson's rule assume that N is even. let $\Delta x = \frac{b-a}{N}$ and $y_j = f(a + j\Delta x)$. the Nth approximation to $\int_a^b f(x)\,dx$ by simpson's rule the quantity
$S_N = \frac{1}{3}\Delta x[y_0 + 4y_1 + 2y_2 + \ldots + 4y_{N-3} + 2y_{N-2} + 4y_{N-1} + y_N]$

we now state (without proof) a bound for the error in simpson's rule, where
$\text{Error}(S_N) = |S_N(f) - \int_a^b f(x)\,dx$
the error involves the fourth derivative, which we assume exists and is continuous.

error bound for S_N let K_4 be a number such that $|f^{(4)}(x)| \le K_4$ for all $x \in [a,b]$. then
$\text{Error}(S_N) \le \frac{K_4(b-a)^5}{180N^4}$

integration by parts formula

1. $\frac{d}{dx}(u \cdot v) = u \cdot \frac{dv}{dx} + v \cdot \frac{du}{dx}$

2. $\int \frac{d}{dx}(u \cdot v)\,dx = \int u \cdot \frac{dv}{dx}\,dx + \int v \cdot \frac{du}{dx}\,dx$

3. $uv = \int u\,dv + \int v\,du$

4. $\int u\,dv = uv - \int v\,du$

summary in terms of notation:

- $u = u(x)$

- $v = \int dv = v(x)$

- then $du = \frac{du}{dx}dx$, and $dv = \frac{dv}{dx}dx$

integration by parts can be used to derive reduction formulas for integrals that depend on a positive integer n such as $\int x^n e^x\,dx$ or $\int \ln^n x\,dx$. a reduction formula (also called a recursive formula) expresses the integral for a given value of n in terms of a similar integral for a smaller value of n. the desired integral is evaluated by applying the reduction formula repeatedly.

a reduction formula verify the reduction formula
$\int x^n e^x\,dx = x^n e^x - n \int x^{n-1} e^x\,dx$
then use the reduction formula to evaluate $\int x^3 e^x\,dx$. **solution** we derive the reduction formula using integration by parts once:
$u = x^n, v' = e^x, u' = nx^{n-1}, v = e^x$
$\int x^n e^x\,dx = uv - \int u'v\,dx = x^n e^x - n \int x^{n-1} e^x\,dx$
to evaluate $\int x^3 e^x\,dx$, we will need to use the reduction formula for $n = 3, 2, 1$:

1. $\int x^3 e^x\,dx = x^3 e^x - 3 \int x^2 e^x\,dx$

2. $= x^3 e^x - 3(x^2 e^x - 2 \int xe^x\,dx) = x^3 e^x - 3x^2 e^x + 6 \int xe^x\,dx$

3. $= x^3 e^x - 3x^2 e^x + 6(xe^x - \int e^x\,dx) = x^3 e^x - 3x^2 e^x + 6xe^x - 6e^x + C$

4. $= (x^3 - 3x^2 + 6x - 6)e^x + C$

in this section, we combine substitution and integration by parts with the appropriate identities to integrate various trigonometirc functions. first, consider $\int \sin^m(x) \cos^n(x)\,dx$
where n, m are positive integers. the easier case occurs when one or both of the exponents m, n is odd.

in general, if n is odd, we may use the identity $\cos^2(x) = 1 - \sin^2(x)$ to write $\cos^n(x)$ as a power of $(1 - \sin^2(x))$ times $\cos(x)$ and subsitute $u = \sin(x), du = \cos(x)dx$. alternatively, if m is odd, we write $\sin^m(x)$ as a power of $(1 - \cos^2(x))$ times $\sin(x)$.

reduction formulas for sine and cosine

- $\int \sin^n(x)\,dx = -\frac{1}{n} \sin^{n-1} x \cos(x) + \frac{n-1}{n} \int \sin^{n-2}(x)\,dx$

- $\int \cos^n(x)\,dx = -\frac{1}{n} \cos^{n-1} x \sin(x) + \frac{n-1}{n} \int \cos^{n-2}(x)\,dx$

evaluate $\int \sin^4(x)\, dx$

by reduction formula with $n = 4$, we obtain

$\int \sin^4(x)\, dx = -\frac{1}{4}\sin^3(x)\cos(x) + \frac{3}{4}\int \sin^2(x)\, dx$

to evaluate the integral on the right, we apply the reduction formula again with $n = 2$:

$\int \sin^2(x)\, dx = -\frac{1}{2}\sin(x)\cos(x) + \frac{1}{2}\int dx = -\frac{1}{2}\sin(x)\cos(x) + \frac{1}{2}x + C$

$\int \sin^4(x)\, dx = -\frac{1}{4}\sin^3(x)\cos(x) - \frac{3}{8}\sin(x)\cos(x) + \frac{3}{8}x + C$

trigonometric integrals can be expressed in more than one way because of the large number of trigonometric identities. for example, a computer algebra system may yield the following evaluation of the integral in the previous example:

$\int \sin^4(x)\, dx = \frac{1}{32}(x - 8\sin(2x) + \sin(4x)) + C$

you can check that this agrees with the result. the following formulas are verified using some of those double angle identities.

- $\int \sin^2(x)\, dx = \frac{x}{2} - \frac{\sin(2x)}{4} + C = \frac{x}{2} - \frac{1}{2}\sin(x)\cos(x) + C$

- $\int \cos^2(x)\, dx = \frac{x}{2} + \frac{\sin(2x)}{4} + C = \frac{x}{2} + \frac{1}{2}\sin(x)\cos(x) + C$

more work is required to integrate $\sin^m(x)\cos^n(x)$ when both n and m are even:

if $m \leq n$, use the identity $\sin^2(x) = 1 - \cos^2(x)$ to replace $\sin^m(x)$ by $(1 - \cos^2(x))^{\frac{m}{2}}$:

$\int \sin^m(x)\cos^n(x)\, dx = \int (1 - \cos^2(x))^{\frac{m}{2}}\cos^n(x)\, dx$

expand the integral on the right to obtain a sum of integrals of powers of $\cos(x)$ and use reduction formula.

if $m \geq n$, replace $\cos^n(x)$ by $(1 - \sin^2(x))^{\frac{n}{2}}$:

$\int \sin^m(x)\cos^n(x)\, dx = \int (\sin^m(x))(1 - \sin^2(x))^{\frac{n}{2}}\, dx$

expand the integral on the right to obtain a sum of integrals of powers of $\sin(x)$ and again evaluate using reduction formula.

our next goal is to integrate functions involving one of the square root expressions:

$\sqrt{a^2 - x^2}$, $\sqrt{x^2 + a^2}$, $\sqrt{x^2 - a^2}$

in each case, a substitution transforms the integral into a trigonometric integral. for example, the substitution $x = a\sin(\theta)$ may be used when the integrand involves $\sqrt{a^2 - x^2}$.

integrals involving $\sqrt{a^2 - x^2}$ if $\sqrt{a^2 - x^2}$ occurs in an integral where $a > 0$, try the substitution

$x = a\sin(\theta)$, $dx = a\cos(\theta)d\theta$, $\sqrt{a^2 - x^2} = a\cos(\theta)$

integrals involving $\sqrt{a^2 + x^2}$ if $\sqrt{a^2 + x^2}$ occurs in an integral where $a > 0$, try the substitution

$x = a\tan(\theta)$, $dx = a\sec^2(\theta)d\theta$, $\sqrt{a^2 + x^2} = a\sec(\theta)$

integrals involving $\sqrt{x^2 - a^2}$ if $\sqrt{x^2 - a^2}$ occurs in an integral where $a > 0$, try the substitution

$x = a\sec(\theta)$, $dx = a\sec(\theta)\tan(\theta)d\theta$, $\sqrt{x^2 - a^2} = a\tan(\theta)$

in this section, we introduce the methods of partial fractions, used to integrate rational functions $\frac{P(x)}{Q(x)}$, where $P(x)$ and $Q(x)$ are polynomials. as we will see the integral of a rational function can be expressed as a sum of three types of terms: rational functions, arctangents of linear or quadratic functions, and logarithms of polynomials. here is a typical example:

$\int \frac{(2x^3 + x^2 - 2x)dx}{(x^2+1)^2} = \frac{x+4}{2x^2+2} + \frac{3}{2}\tan^{-1}x + \ln(x^2+1) + C$

the method of partial fractions yields an explicit antiderivative of this type whenever we can factor the denominator $Q(x)$ into a product of linear and quadratic factors.

a rational function $\frac{P(x)}{Q(x)}$ is called proper if the degree of $P(x)$ is less than the degree of $Q(x)$.

suppose first that $\frac{P(x)}{Q(x)}$ is proper and that the denominator $Q(x)$ factors as a product of distinct linear factors. in other words,

$\frac{P(x)}{Q(x)} = \frac{P(x)}{(x-a_1)(x-a_2)\ldots(x-a_n)}$

where the roots $a_1, a_2, \ldots, a_n$ are all distinct and $\deg(P) < n$. then there is a partial fraction decomposition:

$\frac{P(x)}{Q(x)} = \frac{A_1}{(x-a_1)} + \frac{A_2}{x-a_2} + \ldots + \frac{A_n}{x-a_n}$

a quadratic polynomial $ax^2 + bx + c$ is called irreducible if it cannot be written as a product of two linear factors (without using complex numbers). a power of an irreducible quadratic factor $(ax^2 + bx + c)^M$ contributes a sum of the following type to a partial fraction decomposition:

$\frac{A_1 x + B_1}{ax^2+bx+c} + \frac{A_2 x + B_2}{(ax^2+bx+c)^2} + \ldots + \frac{A_M x + B_M}{(ax^2+bx+c)^M}$

for example,

$\frac{4-x}{x(x^2+4x+2)^2} = \frac{1}{x} - \frac{x+4}{x^2+4x+2} - \frac{2x+9}{(x^2+4x+2)^2}$

you may need to use trigonometric substitution to integrate these terms. in particular, the following result may be useful.

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

sometimes, we are interested in integrating a function over an infinite interval. for example, in statistics, the probability integral

$\int_{\infty}^{\infty} e^{\frac{-x^2}{2}}\,dx$

represents the area under the bell-shaped curve for $-\infty < x < \infty$. this is an example of an improper integral. although the region under the curve extends infinitely to the left and right, the total area is finite.

we consider three types of improper integrals, depending on whether one or both endpoints are infinite:

$\int_{-\infty}^{b} f(x)\,dx$, $\int_{a}^{\infty} f(x)\,dx$, $\int_{-\infty}^{\infty} f(x)\,dx$

improper integral fix a number a and assume that $f(x)$ is integrable over $[a, b]$ for all $b \geq a$. the improper integral of $f(x)$ over $[a, \infty)$ is defined as the limit (if it exists)

18

$\int_a^\infty f(x)\,dx = \lim_{R\to\infty} f(x)\,dx$

when the limit exists, we say that the improper integral converges. in a similar fashion we define $\int_{-\infty}^a f(x)\,dx = \lim_{R\to-\infty} \int_R^a f(x)\,dx$

a doubly infinite improper integral is defined as a sum (provided that both integrals on the right converge):
$\int_{-\infty}^\infty f(x)\,dx = \int_{-\infty}^0 f(x)\,dx + \int_0^\infty f(x)\,dx$

show that $\int_2^\infty \frac{dx}{x^2}$ converges and compute its value.

first we evaluate the definite integral of a finite interval $[2, R]$: $\int_2^R \frac{dx}{x^2} = -x^{-1}\Big|_2^R = \frac{1}{2} - \frac{1}{R}$

the limit as $R \to \infty$ exists, so the improper integral converges:
$\int_2^\infty \frac{dx}{x^2} = \lim_{R\to\infty} \int_2^R \frac{dx}{x^2} = \lim_{R\to\infty}(\frac{1}{2} - \frac{1}{R}) = \frac{1}{2}$

when the limit defining an improper integral is either infinite or does not exists, we say that the improper integral diverges. for instance, the are under the parabola $y = x^2$ appears infinite and indeed, the following integral diverges:
$\int_0^\infty x^2\,dx = \lim_{R\to\infty} \int_0^R x^2\,dx = \lim_{R\to\infty} \frac{1}{3}R^3 = \infty$

in general, however, we cannot tell if an improper integral converges or diverges merely by looking at the graph of the function. the function $f(x) = x^{-2}$ and $g(x) = x^{-1}$ are both decreasing, but $f(x) = x^{-2}$ decreases more rapidly as $x \to \infty$, and the are under its graph over $[2, \infty)$ is finite, as we saw in the previous example. the next example shows that $g(x) = x^{-1}$ does not tend to zero quickly enough for the improper integral to converge.

determine if $\int_{-\infty}^{-1} \frac{dx}{x}$ converges.

$\int_R^{-1} \frac{dx}{x} = \ln(|x|)\Big|_R^{-1} = \ln(|-1|) - \ln(|R|) = -\ln(|R|)$

the improper integral diverges because the limit as $R \to -\infty$ is infinite:
$\lim_{R\to-\infty} \int_R^{-1} \frac{dx}{x} = \lim_{R\to-\infty}(-\ln(|R|)) = -\infty$

improper integral of x^{-P} over $[1, \infty]$

$\int_1^\infty \frac{dx}{x^P} = \begin{cases} \frac{1}{p-1} & \text{if } p > 1 \\ \infty & \text{if } p \leq 1 \end{cases}$

another type of improper integral occurs when the integrand $f(x)$ becomes infinite at one or both of the endpoints of the interval of integration. for example, $\int_0^1 \frac{dx}{\sqrt{x}}$ is an improper integral because the integrand $x^{\frac{-1}{2}}$ tends to ∞ as $x \to 0+$. this integral represents the area of a region that is infinite in the vertical direction. improper integrals of this type are defined as one-side limits.

integrands with infinite discontinuities if $f(x)$ is continuous on $[a, b)$ but discontinuous at $x = b$, we define
$\int_a^b f(x)\,dx = \lim_{R\to b-} \int_a^R f(x)\,dx$
similarly, if $f(x)$ is continuous on $(a, b]$ but discontinuous at $x = a$,

$\int_a^b f(x)\,dx = \lim_{R \to a+} \int_R^b f(x)\,dx$
if, in either case, the limit does not exist, the integral is said to diverge.

improper integral of x^{-p} over $[0,1]$

$\int_0^1 \frac{dx}{x^P} = \begin{cases} \frac{1}{1-p} & \text{if } p < 1 \\ \infty & \text{if } p \geq 1 \end{cases}$

sometimes we are interested in determining whether an improper integral converges, even if we cannot find its exact value. for instance, the following integral cannot be evaluated explicitly:

$\int_1^\infty \frac{e^{-x}}{x}\,dx$

however, we have the inequalities

$0 \leq \frac{e^{-x}}{x} \leq e^{-x}$ for $x \geq 1$

thus, the graph of $\frac{e^{-x}}{x}$ lies underneath the graph of e^{-x} to the right of $x = 1$ and we should be able to conclude that our integral converges:

$\int_1^\infty \frac{e^{-x}}{x}$ this converges by comparison

$dx \leq \int_1^\infty e^{-x}\,dx = e^{-1}$ converges by direct computation

the next theorem summarizes this type of argument.

comparison test for improper integrals
assume that $f(x) \geq g(x) \geq 0$ for $x \geq a$,

- if $\int_a^\infty f(x)\,dx$ converges, then $\int_a^\infty g(x)\,dx$ also converges

- if $\int_a^\infty g(x)\,dx$ diverges, then $\int_a^\infty f(x)\,dx$ also diverges