

V25.0109: General Chemistry I: Honors

Mathematical Preliminaries

Chemistry, like physics, is a quantitative science rather than a descriptive one. Answers in chemistry are expressed as real numbers corresponding to observable properties of a system. In order to arrive at such numbers, you must be capable of translating a complicated situation into the appropriate mathematical language and then applying mathematical methods to solve the resulting equations. It should be clear, then, that mathematics is the underlying language of the quantitative sciences, and in order to use it, you must be reasonably fluent in its vocabulary. This document provides a brief overview of the mathematical concepts we will need throughout the course. It is strongly recommended that you familiarize yourself with each of the topics described herein. If you have difficulties with any of these subjects, come to office hours to have your questions answered.

I. VECTORS AND VECTOR ALGEBRA

Consider the real number line:

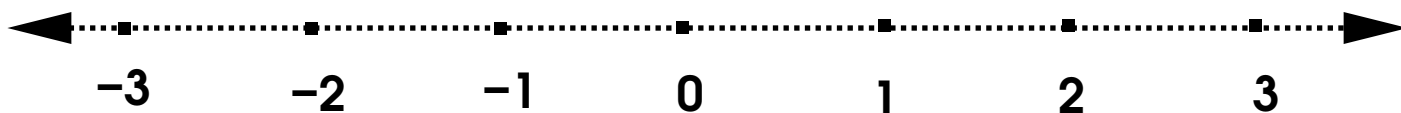


FIG. 1.

With zero in the middle, the line extends out in infinitely in both directions, with positive numbers to the right and negative numbers to the left. A number on the line, such as 2, can be located by starting at 0 (the “origin” of the line) and going to the right by 2 units. If it were generally agreed upon that all numbers should be specified in relation to the point at 0, then in order to describe to someone how to find the number 2 on the line, you could say “it is 2 units to the right.”

This statement has two parts. One is a number or *magnitude*, the number of units (2 in this case). The other is a *direction* (“to the right” in this example). Something that has both magnitude *and* a direction is called a *vector*. As this example shows, a vector describes the location of a point in some space (“2 units to the right” locates a point on the line). Vectors are generally designated by an arrow that points in the direction of the vector and whose length represents the magnitude associated with the vector. In the example we have been considering, the vector would appear as shown below:

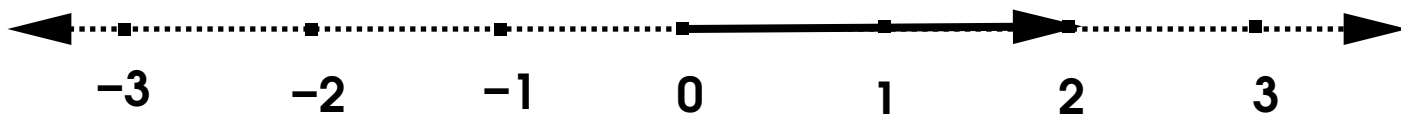


FIG. 2.

If we wanted to locate the point at -2, we should specify its vector as “2 units to the left.” The magnitude is the same, however, the direction has changed, and the vector would appear as:

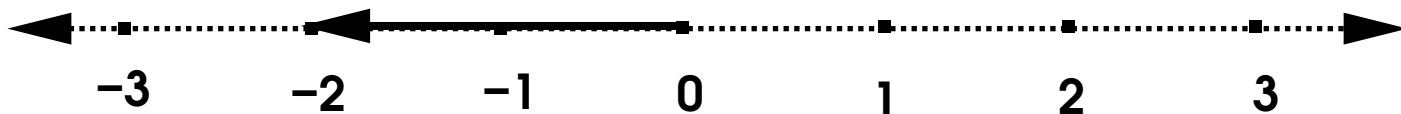


FIG. 3.

The numbers 2 and -2 corresponding to the two points we have considered are the coordinates of these points.

The statement $2+1 = 3$ can be expressed as a *vector addition*. The figure below shows the two vectors corresponding to the points at 2 and 1, respectively:

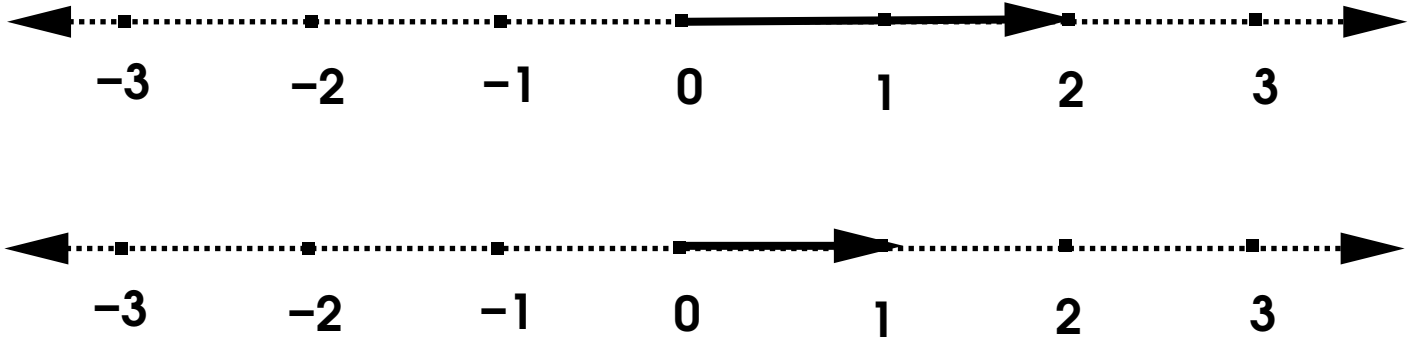
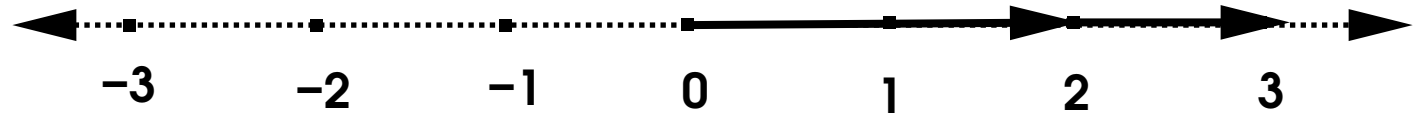


FIG. 4.

The sum can be expressed by placing the vectors together such that the head of one touches the tail of the other. The net length and direction is the answer as shown below:

Vector addition of 2+1



Answer: 3

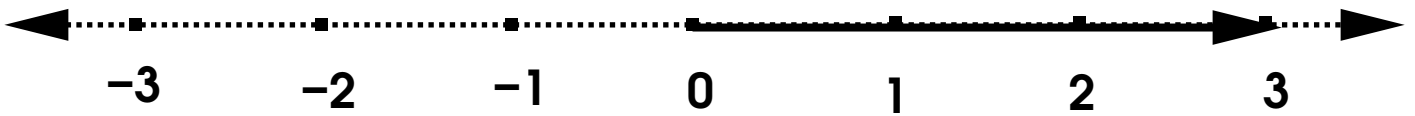


FIG. 5.

The result is vector “3 to the right.” Rather than continuing to specify vectors in this clumsy way, we introduce a notation for them. Generally, vectors are designated by a boldface symbol, e.g., \mathbf{v} . A standard notation is to specify a vector simply by the coordinates of the point it locates. Thus, if the vector \mathbf{v} is “2 to the right”, then we write $\mathbf{v} = 2$, and if it is “2 to the left” it becomes $\mathbf{v} = -2$. Here, the minus sign allows the direction to be determined unambiguously.

In one dimension, directions are easy, as there are only two choices, “to the right” or “to the left”. By contrast, in two dimensions, there is an infinite number of directions. Now two points are required to specify the location of

a point in a plane; its x and y coordinates. The coordinates are written as an ordered pair (x, y) . Thus, the point shown below:

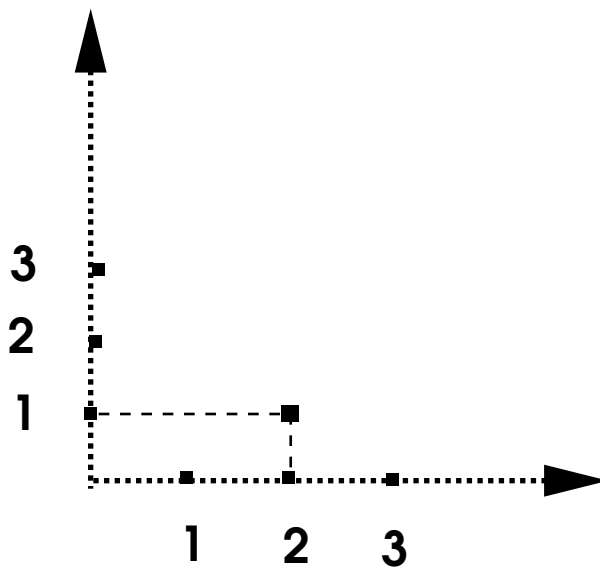


FIG. 6.

has coordinates $(2,1)$. The vector locating this point appears as:

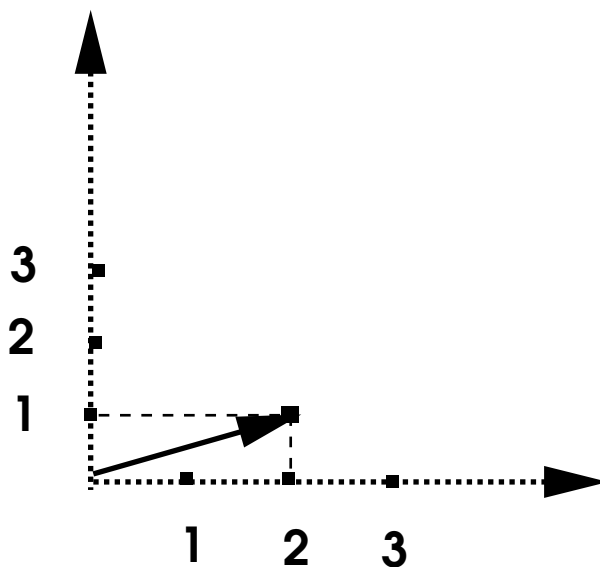


FIG. 7.

and is given simply by the coordinates $\mathbf{v} = (2, 1)$. Another way to express a vector is to introduce unit vectors $\hat{\mathbf{i}}$ and $\hat{\mathbf{j}}$ along the x and y axes and write

$$\mathbf{v} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}}$$

where $\hat{\mathbf{i}}$ is the vector $(1, 0)$ and $\hat{\mathbf{j}}$ is the vector $(0, 1)$. Although we will not use this notation in the course, you should be aware of it as many other sources employ it.

The magnitude of the vector, written as $|\mathbf{v}|$, is its length, given by the length of the hypotenuse of the right triangle: $|\mathbf{v}| = \sqrt{5}$. Its direction is, by convention, given by the angle that the vector makes with the x axis: $\theta = \tan^{-1}(1/2) = 26.57$ degrees. In general, if a vector is given by $\mathbf{v} = (a, b)$, then its length is

$$|\mathbf{v}| = \sqrt{a^2 + b^2}$$

and its direction is

$$\theta = \tan^{-1} \frac{b}{a}$$

Note that the length and direction of a vector are the same as the polar coordinate representation of the point (a, b) . The numbers a and b are called the *components* of \mathbf{v} .

The sum of two vectors, say $\mathbf{v}_1 = (2, 1)$ and $\mathbf{v}_2 = (1, 2)$ is obtained by adding the components separately: $\mathbf{v}_1 + \mathbf{v}_2 = (2 + 1, 1 + 2) = (3, 3)$. This is the same as placing the vectors end-to-end as in the one-dimensional case:

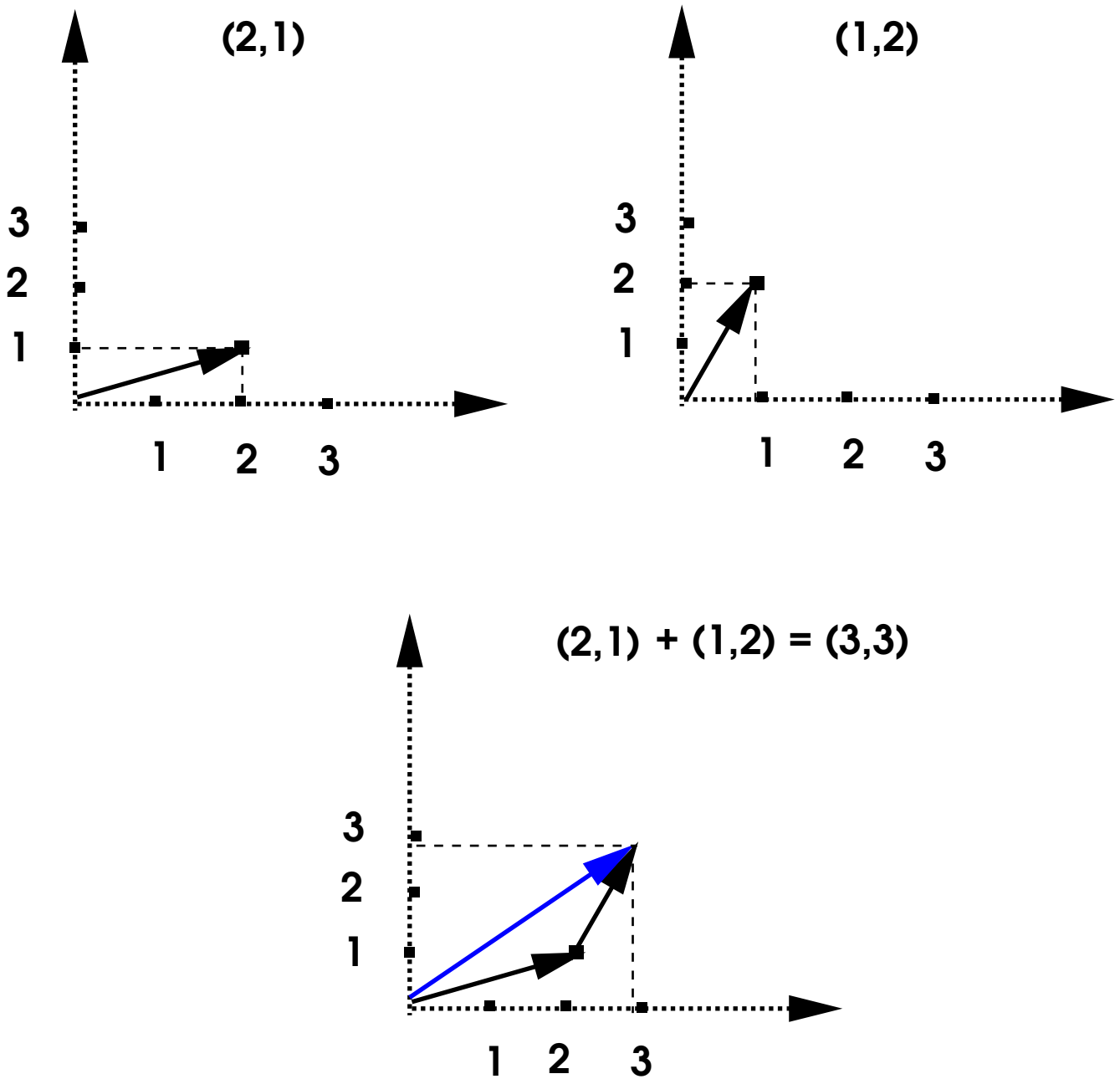


FIG. 8.

In general, if $\mathbf{v}_1 = (a_1, b_1)$ and $\mathbf{v}_2 = (a_2, b_2)$, then the sum $\mathbf{v}_3 = \mathbf{v}_1 + \mathbf{v}_2$ is given by

$$\mathbf{v}_3 = (a_1 + a_2, b_1 + b_2)$$

or if $\mathbf{v}_3 = \mathbf{v}_1 - \mathbf{v}_2$, then

$$\mathbf{v}_3 = (a_1 - a_2, b_1 - b_2)$$

In general, if \mathbf{V} is the sum of N vectors \mathbf{v}_i , $i = 1, \dots, N$, then \mathbf{V} will be given by

$$\mathbf{V} = \sum_{i=1}^N \mathbf{v}_i = \left(\sum_{i=1}^N a_i, \sum_{i=1}^N b_i \right)$$

Also, the product of a vector \mathbf{v} and a constant c is obtained simply by multiplying the components of \mathbf{v} by c :

$$c\mathbf{v} = (ca, cb)$$

One other important operation involves finding the angle between two vectors:

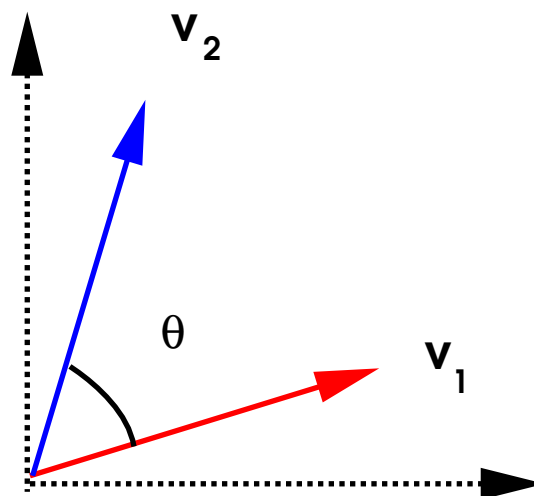


FIG. 9.

The angle θ between the two vectors \mathbf{v}_1 and \mathbf{v}_2 is given by

$$\theta = \cos^{-1} \left[\frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{|\mathbf{v}_1| |\mathbf{v}_2|} \right]$$

where the product $\mathbf{v}_1 \cdot \mathbf{v}_2$ is called the *dot product* between the vectors \mathbf{v}_1 and \mathbf{v}_2 . If $\mathbf{v}_1 = (a_1, b_1)$ and $\mathbf{v}_2 = (a_2, b_2)$, then

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = a_1 a_2 + b_1 b_2$$

Note that the dot product is a number, not a vector!

Everything said here, generalizes to three dimensions, except that the vectors now have a third component, e.g., $\mathbf{v} = (a, b, c)$. Thus, for example, the dot product between two vectors becomes

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = a_1 a_2 + b_1 b_2 + c_1 c_2$$

and the length of a vector is given by

$$|\mathbf{v}| = \sqrt{a^2 + b^2 + c^2}$$

II. TRIGONOMETRIC IDENTITIES

Trigonometry and the identities obeyed by trigonometric functions arise in many of the calculations we will be doing this semester. Hence, it is important for you to be familiar with these. Let us begin by recalling the definitions of the basic functions $\cos(x)$ and $\sin(x)$. Here, we will assume $x \in [0, 2\pi]$, although since the functions are periodic, it does not matter if we extend the domain of x . However, by restricting x in this way, we can make the convenient connection of x to one of the angles in a right triangle (see figure).

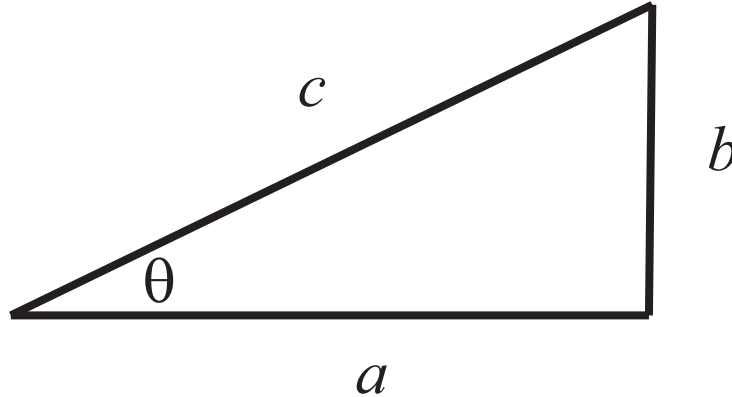


FIG. 10.

Let c be the length of the hypotenuse of the triangle, and let a and b be the sides adjacent and opposite to the angle x . Then, we have the definitions

$$\cos(x) = \frac{a}{c}, \quad \sin(x) = \frac{b}{c}$$

From the pythagorean theorem $a^2 + b^2 = c^2$, so that

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

The above relation gives us our first and most basic trigonometric identity satisfied by $\cos(x)$ and $\sin(x)$:

$$\cos^2(x) + \sin^2(x) = 1$$

Many other identities follow straightforwardly from this one. For example, the other basic trigonometric functions $\tan(x)$, $\cot(x)$, $\sec(x)$ and $\csc(x)$, defined as

$$\begin{aligned} \tan(x) &= \frac{\sin(x)}{\cos(x)}, & \cot(x) &= \frac{1}{\tan(x)} = \frac{\cos(x)}{\sin(x)} \\ \sec(x) &= \frac{1}{\cos(x)}, & \csc(x) &= \frac{1}{\sin(x)} \end{aligned}$$

satisfy a number of related identities. If we divide our first identity through by $\cos^2(x)$, for example, we obtain an identity satisfied by $\tan(x)$ and $\sec(x)$:

$$1 + \frac{\sin^2(x)}{\cos^2(x)} = \frac{1}{\cos^2(x)}$$

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

Similarly, if we divide through by $\sin^2(x)$, we obtain

$$\frac{\cos^2(x)}{\sin^2(x)} + 1 = \frac{1}{\sin^2(x)}$$

$$\cot^2(x) + 1 = \csc^2(x)$$

Note that the definitions of $\cos(x)$ and $\sin(x)$ imply that for certain values of x , $x = 0$, $x = \pi/2$, $x = \pi$, $x = 3\pi/2$ and $x = 2\pi$,

$$\begin{aligned}\cos(0) &= 1, & \sin(0) &= 0 \\ \cos(\pi/2) &= 0, & \sin(\pi/2) &= 1 \\ \cos(\pi) &= -1, & \sin(\pi) &= 0 \\ \cos(3\pi/2) &= 0, & \sin(3\pi/2) &= -1 \\ \cos(2\pi) &= \cos(0) = 1, & \sin(2\pi) &= \sin(0) = 0\end{aligned}$$

A number of other identities concern trigonometric functions whose arguments are sums or differences of angles $x \pm y$:

$$\begin{aligned}\cos(x + y) &= \cos(x)\cos(y) - \sin(x)\sin(y), & \cos(x - y) &= \cos(x)\cos(y) + \sin(x)\sin(y) \\ \sin(x + y) &= \sin(x)\cos(y) + \cos(x)\sin(y), & \sin(x - y) &= \sin(x)\cos(y) - \cos(x)\sin(y)\end{aligned}$$

In fact, the difference formulas can be derived from the sum formulas using the simple symmetry relations $\cos(-x) = \cos(x)$ and $\sin(-x) = -\sin(x)$. In the next section, we will introduce a very simple technique for proving these relations.

Now, if we set $x = y$, we can derive a useful identity for $\cos^2(x)$. Setting $x = y$ in the $\cos(x + y)$ and $\cos(x - y)$ identities, and adding the two identities together, we obtain

$$\begin{aligned}\cos(2x) &= \cos^2(x) - \sin^2(x) \\ \cos(0) &= 1 = \cos^2(x) + \sin^2(x) \\ 1 + \cos(2x) &= 2\cos^2(x) \\ \cos^2(x) &= \frac{1}{2}[1 + \cos(2x)]\end{aligned}$$

If we subtract the identities instead of adding them, we can derive a similar identity for $\sin^2(x)$:

$$\sin^2(x) = \frac{1}{2}[1 - \cos(2x)]$$

Finally, we note that $\cos(x)$ and $\sin(x)$ are expressible as infinite power series (also known as Taylor series) as

$$\begin{aligned}\cos(x) &= 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \frac{1}{6!}x^6 + \dots \\ &= \sum_{k=0}^{\infty} (-1)^k \frac{1}{(2k)!} x^{2k} \\ \sin(x) &= x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7 + \dots \\ &= \sum_{k=0}^{\infty} (-1)^k \frac{1}{(2k+1)!} x^{2k+1}\end{aligned}$$

These power series expressions are actually used in calculators and computer programs to evaluate $\cos(x)$ and $\sin(x)$.

III. COMPLEX NUMBERS

Complex numbers play a key role in quantum mechanics, even though anything that can be measured experimentally must ultimately be expressed as a real number. Without quantum theory, it is not possible to understand the nature of the chemical bond, and since quantum theory is expressed in terms of complex numbers, it is important to review some of the basic facts about complex quantities. Complex numbers are so named because they are composed of two types of more fundamental numbers: real numbers and imaginary numbers. Real numbers are very familiar and are the numbers we deal with on a daily basis. However, when one embarks on the study of nonlinear algebraic equations, it rapidly becomes clear that some equations do not have solutions within the set of real numbers. Take the example of the equation

$$x^2 + 1 = 0.$$

There are *no* real numbers x that satisfy this equation. The reason is clear if we just naively solve for x to find $x = \pm\sqrt{-1}$. The quantity $\sqrt{-1}$ is not defined within the real numbers. But mathematicians are never satisfied with the idea that something cannot be defined. In this case, the conundrum can be solved by simply extending the concept of a “number” to allow $\sqrt{-1}$ to be defined. This involves inventing a new set of numbers, called “imaginary numbers” (for no better reason than “imaginary” is one of the opposites to “real”), in which $\sqrt{-1}$ has a definite value, denoted i . Then, the solution to the above equation is simply $x = \pm i$. The number i is the fundamental “integer” unit in the set of imaginary numbers just as 1 is the fundamental integer unit in the real numbers. So, just as any real number a can always be expressed as $a \cdot 1$, any imaginary number can be expressed as a multiple of i . If β is an imaginary number, then we can write β as $b \cdot i$, where b is a real number. The number i , itself, is really $1 \cdot i$, and clearly, 1 is real.

Given imaginary numbers, the most general kind of number we can now imagine has both real and imaginary parts, and such a number is known as a *complex number*. A complex number z is expressed as the sum of its real and imaginary parts. If the real part of z is a and the imaginary part of z is bi , then we would write

$$z = a + bi$$

Note that this is similar to expressing a two-dimensional vector $\mathbf{v} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}}$. Indeed, we can exploit this analogy and use the ordered-pair notation to write a complex number as $z = (a, bi)$ as an alternate notation. As we will see, there are many such analogies between complex numbers and two-dimensional vectors. For instance, we can think of a complex number as a point in a plane in which the horizontal axis corresponds to the real parts of all complex numbers and the vertical axis corresponds to the imaginary parts. The complex number $z = a + bi$ would be the point (a, b) in this plane. In order to extract the real and imaginary parts of z , we introduce the functions Re and Im such that $\text{Re}(z) = a$ and $\text{Im}(z) = b$.

We noted above that experimental measurements ultimately yield real numbers, so we need a way of going from a complex number to a real one. For this, we exploit the analogy with two-dimensional vectors even more and define the “measure” or *magnitude* of a complex number $z = a + ib$, which we denote as $|z|$, as

$$|z| = \sqrt{a^2 + b^2}$$

which is the same as the magnitude of a two-dimensional vector $\mathbf{v} = (a, b) = a\hat{\mathbf{i}} + b\hat{\mathbf{j}}$. Another important definition is the so-called *complex conjugate* of z , denoted z^* , which is defined as follows: If $z = a + ib$, then $z^* = a - ib$. In general, all we have to do to obtain the complex conjugate of a number is replace i by $-i$. Note that the magnitude of z is

$$\begin{aligned} |z| &= \sqrt{z^*z} \\ &= \sqrt{(a - ib)(a + ib)} \\ &= \sqrt{a^2 + iab - iab + b^2} \\ &= \sqrt{a^2 + b^2} \end{aligned}$$

where we have used the fact that $i(-i) = -i^2 = 1$. Thus, a complex number times its complex conjugate is a real number.

What happens if we are given a complex number, but it is not expressed in the simple form $z = a + ib$? Can we manipulate it so that it finally does have this form? The answer is 'yes' with a little mathematical creativity. Consider, first, a simple product $z = (u + iv)(x + iy)$. This can be expressed in the canonical form by simply multiplying it out algebraically:

$$\begin{aligned} z &= ux + ivx + iuy - vy \\ &= (ux - vy) + i(vx + uy) \end{aligned}$$

Next, consider a ratio $z = (u + iv)/(x + iy)$. In order to simplify this, we use the fact that $(x + iy)(x - iy)$ is real. Thus, we multiply z by 1 in the form of $(x - iy)/(x - iy)$:

$$\begin{aligned} z &= \frac{u + iv}{x + iy} \frac{x - iy}{x - iy} \\ &= \frac{(u + iv)(x - iy)}{(x + iy)(x - iy)} \\ &= \frac{ux + yv + i(vx - uy)}{x^2 + y^2} \\ &= \frac{ux + yv}{x^2 + y^2} + i \frac{vx - uy}{x^2 + y^2} \end{aligned}$$

One of the most frequently occurring complex number forms is

$$z = e^{i\theta}$$

where θ is a real number. The complex conjugate of z is obtained simply by replacing i by $-i$:

$$z^* = e^{-i\theta}$$

and so that magnitude of z is

$$|z|^2 = z^*z = e^{-i\theta}e^{i\theta} = 1$$

So, z is a complex number whose square magnitude is 1. On the other hand, if we seek to represent z in canonical form as $z = a + ib$, then it follows that

$$|z|^2 = a^2 + b^2 = 1$$

Since a and b are actually functions of θ , we can ask what two functions of θ are related by $a^2 + b^2 = 1$. We have already seen what these are – they are $\cos(\theta)$ and $\sin(\theta)$ since $\cos^2(\theta) + \sin^2(\theta) = 1$. Thus, in canonical form, we can let $a = \cos(\theta)$ and $b = \sin(\theta)$ and write

$$z = e^{i\theta} = \cos(\theta) + i \sin(\theta)$$

which is known as *Euler's formula*. The above argument is not a very rigorous one, as there are many other functions a and b of θ we could have chosen that satisfy $a^2 + b^2 = 1$. However, we can actually prove rigorously that a and b must be, respectively, $\cos(\theta)$ and $\sin(\theta)$ using the fact that the function $\exp(x)$ has the following Taylor series expression:

$$\begin{aligned} e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \frac{x^6}{6!} + \dots \\ &= \sum_{k=0}^{\infty} \frac{x^k}{k!} \end{aligned}$$

Applying this to $\exp(i\theta)$, we obtain

$$\begin{aligned}
e^{i\theta} &= 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \frac{(i\theta)^5}{5!} + \dots \\
&= \sum_{k=0}^{\infty} \frac{(i\theta)^k}{k!}
\end{aligned}$$

Now the even power terms are purely real, and the odd powers are purely imaginary, so if we wish to put the above power series into the canonical form, we obtain

$$\begin{aligned}
e^{i\theta} &= 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \dots + i \left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} + \dots \right) \\
&= \cos(\theta) + i \sin(\theta)
\end{aligned}$$

which proves Euler's formula.

Euler's formula can be used to simplify proofs of trigonometric identities. For example, we know that $z^*z = 1$. But from Euler's formula, $z^*z = \cos^2(\theta) + \sin^2(\theta)$, so $\cos^2(\theta) + \sin^2(\theta) = 1$. If we use the fact that

$$z^* = e^{-i\theta} = \cos(\theta) - i \sin(\theta)$$

Thus, adding z to z^* , we obtain

$$\cos(\theta) = \frac{1}{2} (e^{i\theta} + e^{-i\theta})$$

Subtracting z from z^* yields

$$\sin(\theta) = \frac{1}{2i} (e^{i\theta} - e^{-i\theta})$$

Now consider $\cos^2(\theta)$. From the above formula, we have

$$\begin{aligned}
\cos^2(\theta) &= \frac{1}{4} (e^{i\theta} + e^{-i\theta})^2 \\
&= \frac{1}{4} (e^{2i\theta} + 2e^{i\theta}e^{-i\theta} + e^{-2i\theta}) \\
&= \frac{1}{4} (2 + e^{2i\theta} + e^{-2i\theta}) \\
&= \frac{1}{2} + \frac{1}{4} (e^{2i\theta} + e^{-2i\theta}) \\
&= \frac{1}{2} [1 + \cos(2\theta)]
\end{aligned}$$

Finally, note that if we multiply Euler's formula by a real number A , we have a complex number of the form

$$z = Ae^{i\theta}$$

In canonical form

$$z = A \cos(\theta) + iA \sin(\theta)$$

and $|z|^2 = A \exp(-i\theta) * A \exp(i\theta) = A^2$, so $|z| = A$. When we write $z = A \exp(i\theta)$, The angle θ is called the *phase* of z .

IV. DERIVATIVE OF A FUNCTION

A line, described by the equation

$$y = mx + b$$

has a constant slope m defined by

$$m = \frac{\Delta y}{\Delta x}$$

where the above equations means compute the change, Δy in y for a given change, Δx in x as shown in the figure:

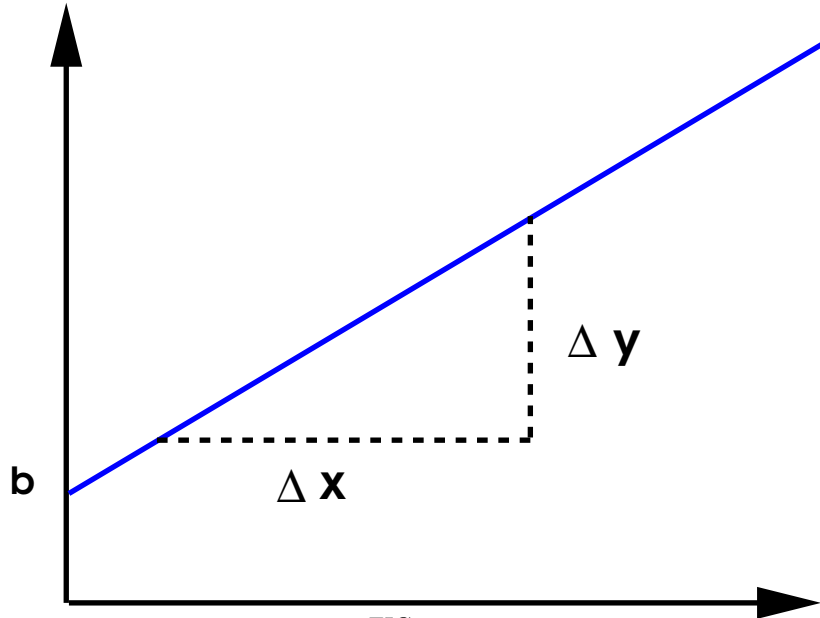


FIG. 11.

For a general function, such as that shown below:

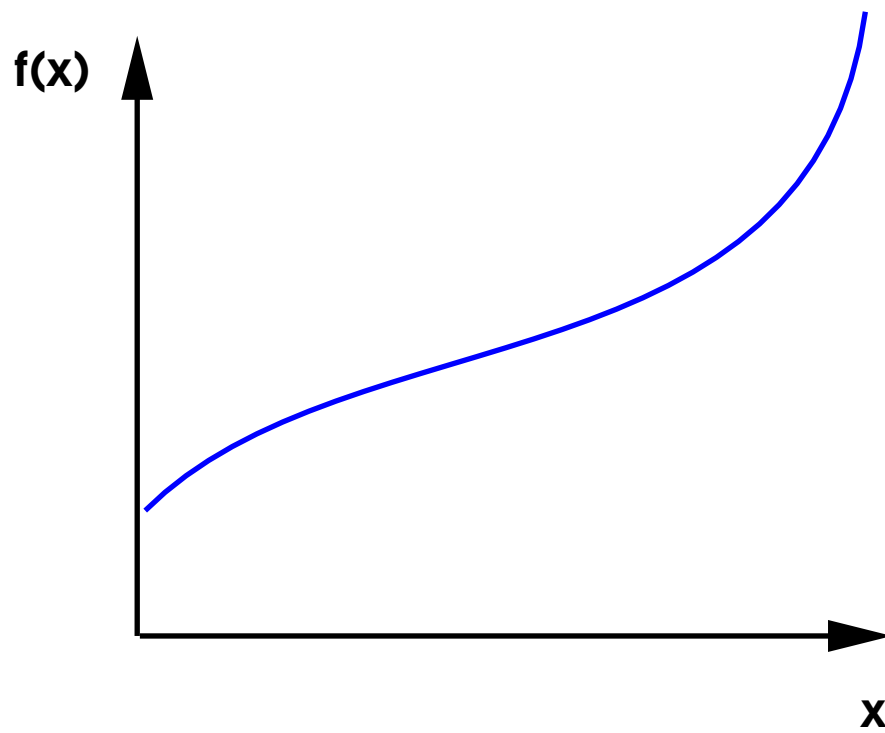


FIG. 12.

the slope is not a constant. However, it is possible to define the slope of the curve at each point by taking the slope of the line tangent to the curve at each point. The slope will also be a function of x and is called the derivative of the function $f(x)$. To see how the slope is computed, consider a line connecting two points of the function corresponding to two points x and $x + \Delta x$, where Δx is a small number:

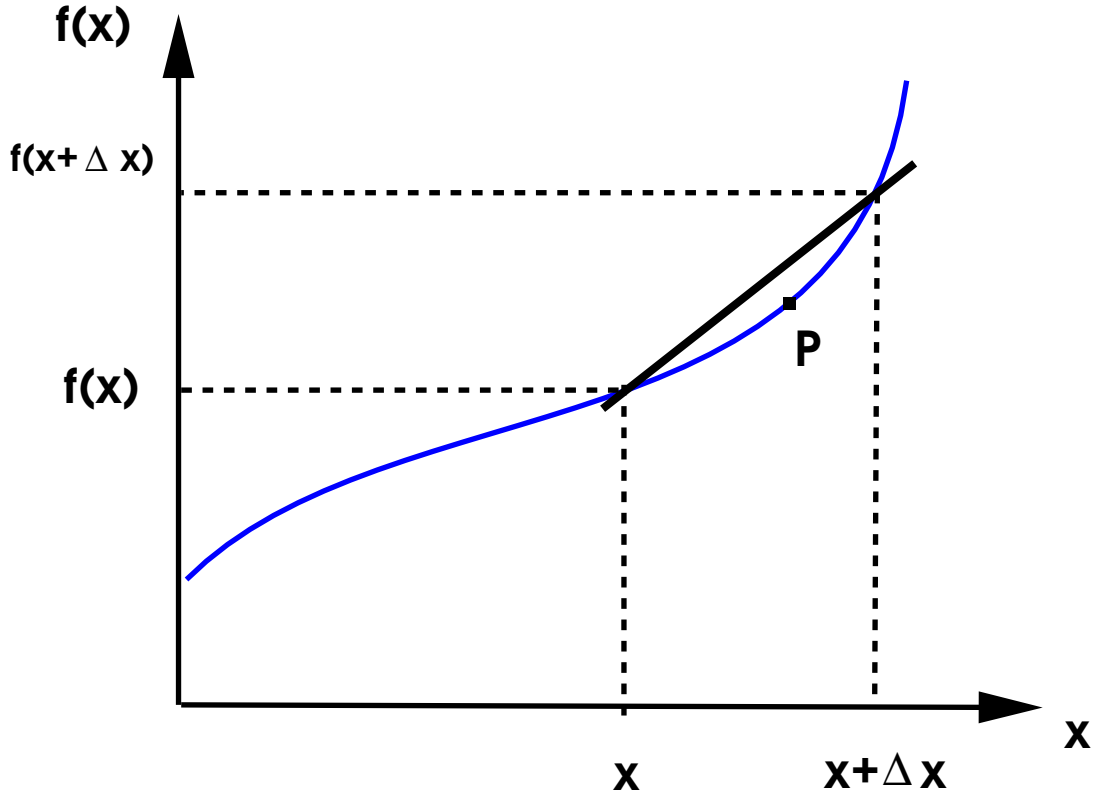


FIG. 13.

The slope of this line is

$$m = \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

To find the slope at the single point P , we need to let Δx go to 0, so that the line intersects the curve at the point P , where it will be tangent to the curve. Thus, by taking the *limit* as $\Delta x \rightarrow 0$, we obtain the derivative of $f(x)$ at the point P . But since the point P is arbitrary, we may perform this operation at any point and, thus, obtain the derivative of $f(x)$ at any point. The general expression is

$$f'(x) = \frac{df}{dx} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

where $f'(x)$ and df/dx are standard notations for the derivative.

As an example, consider the function $f(x) = x^2$. According to the formula:

$$\begin{aligned} f'(x) &= \lim_{\Delta x \rightarrow 0} \frac{(x + \Delta x)^2 - x^2}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \frac{x^2 + 2x\Delta x + \Delta x^2 - x^2}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} (2x + \Delta x) = 2x \end{aligned}$$

Thus, the derivative of x^2 is $2x$. In fact, in general, if $f(x) = x^n$, $f'(x) = nx^{n-1}$. A few other functions and their derivatives are listed below:

$$\begin{aligned} f(x) &= e^x & f'(x) &= e^x \\ f(x) &= \ln x & f'(x) &= \frac{1}{x} \\ f(x) &= \sin x & f'(x) &= \cos x \\ f(x) &= \cos x & f'(x) &= -\sin x \end{aligned}$$

Other important rules of differentiation are the *product rule*:

if $h(x) = f(x)g(x)$ then

$$h'(x) = f'(x)g(x) + f(x)g'(x)$$

and the *chain rule*:

if $h(x) = f(g(x))$, then

$$h'(x) = f'(g(x))g'(x)$$

V. INTEGRAL OF A FUNCTION

It is often necessary to calculate the area under a function, $f(x)$ between two points x_1 and x_2 . Such an operation is necessary, for example, in mechanics to compute the work performed by an agent exerting a force on an object. An analogous situation arises in thermodynamics in computing the work done by an agent exerting an external pressure on a system.

Consider the area under the curve designated below:

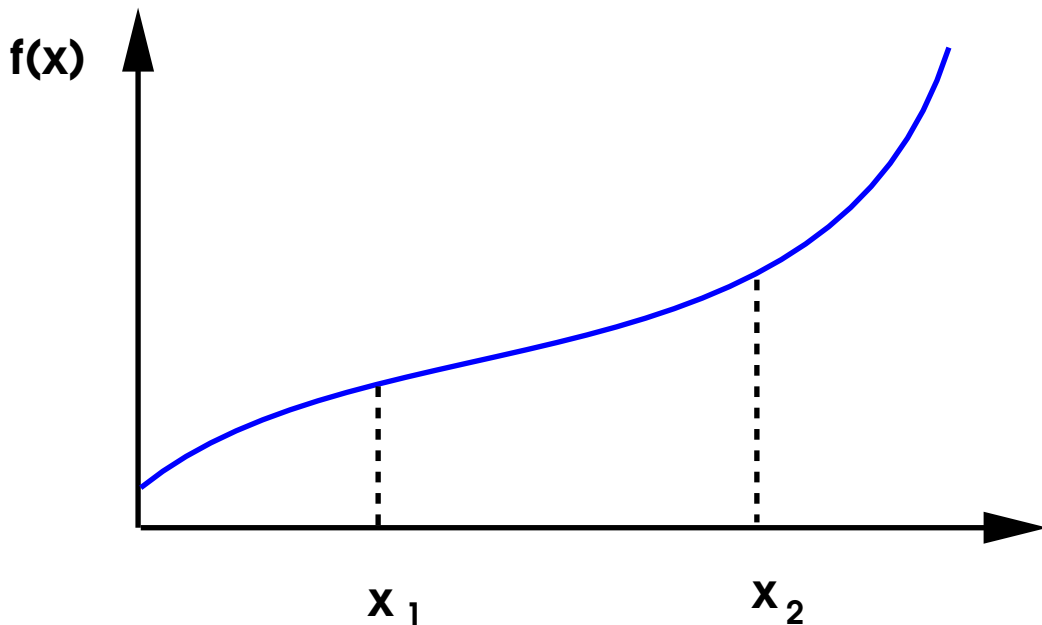


FIG. 14.

One way to calculate the area approximately would be to cover the area with narrow rectangles or width Δx :

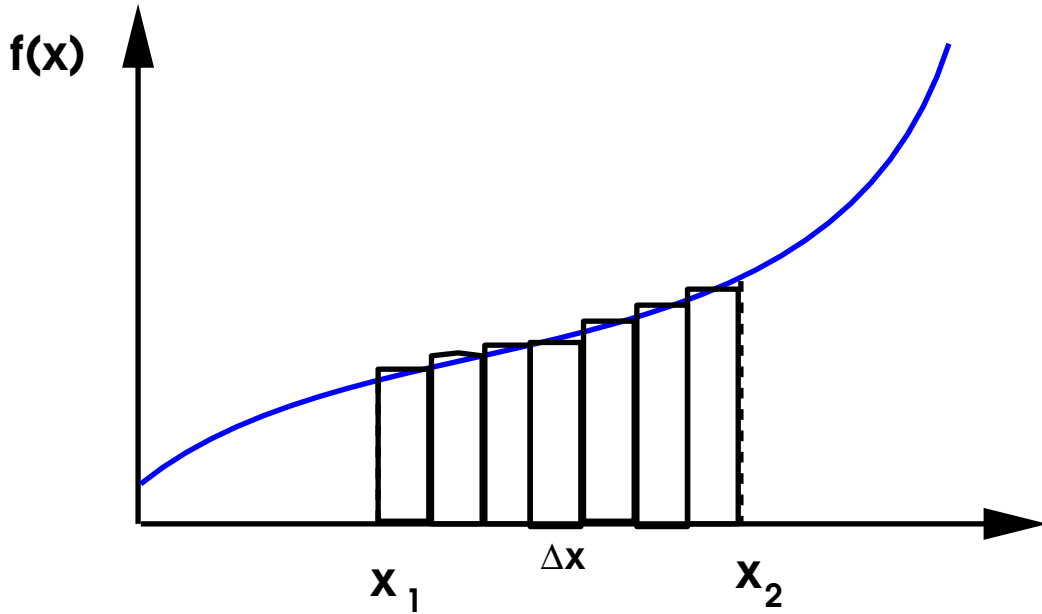


FIG. 15.

The height of the rectangle between x_1 and $x_1 + \Delta x$ will be $f(x_1^*)$, where x_1^* is the midpoint of the interval. The area of the rectangle will be $\Delta x f(x_1^*)$. If there are N such rectangles, then $\Delta x = (x_2 - x_1)/N$, and the area will be

$$\text{area} \approx \Delta x \sum_{i=1}^N f(x_i^*)$$

In the limit that $\Delta x \rightarrow 0$ and $N \rightarrow \infty$, this expression gives the area exactly. This limit is expressed as the *integral* of the function:

$$\text{area} = \int_{x_1}^{x_2} f(x) dx$$

The fundamental theorem of calculus states that

$$\int_{x_1}^{x_2} f(x) dx = F(x_2) - F(x_1)$$

where $F(x)$ is the *antiderivative* of $f(x)$, i.e., it satisfies

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

Some common integrals are:

$$\begin{aligned} f(x) = x & & F(x) = \frac{x^2}{2} + C \\ f(x) = \frac{1}{x} & & F(x) = \ln x + C \\ f(x) = e^x & & F(x) = e^x + C \\ f(x) = x^n & & F(x) = \frac{x^{n+1}}{n+1} + C \\ f(x) = \cos(x) & & F(x) = \sin(x) + C \\ f(x) = \sin(x) & & F(x) = -\cos(x) + C \end{aligned}$$

Here, C is an arbitrary constant.

A. Integration by parts

Consider the integral of a function $f(x)$

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

that can be expressed in the form

$$\int_a^b u(x)dv(x)$$

Integration by parts is based on the product rule for differentiation $d(uv)/dx = u(x)dv/dx + v(x)du/dx$. Using the product rule, we can express udv as $d(uv) - vdu$ and write the integral as

$$\int_a^b u(x)dv(x) = u(x)v(x)|_a^b - \int_a^b v(x)du(x)$$

Integration by parts can be used to simplify integrals that are complicated products of functions. Consider the following example:

$$\int_0^1 xe^{-ax}dx$$

where a is a constant. If we choose $u(x) = x$ and $dv(x) = e^{-ax}dx$, then we can apply integration by parts to simplify the integral. According to the formula, we have $du(x) = dx$ and $v(x) = -(1/a)e^{-ax}$, leading to

$$\begin{aligned}\int_0^1 xe^{-ax}dx &= -\frac{x}{a}e^{-ax}\Big|_0^1 + \frac{1}{a}\int_0^1 e^{-ax}dx \\ &= -\frac{1}{a}e^{-a} - \frac{1}{a^2}e^{-ax}\Big|_0^1 \\ &= -\frac{1}{a}e^{-a} + \frac{1}{a^2}(1 - e^{-a})\end{aligned}$$

B. Integration by differentiation with respect to a constant

Another handy trick for simplifying integrals is to express a complicated integrand as the derivative of a simpler integrand with respect to a constant. Consider again the example

$$\int_0^1 xe^{-ax}dx$$

Notice that if we differentiate e^{-ax} with respect to a , we have

$$\frac{d}{da}e^{-ax} = -xe^{-ax}$$

Thus, we can express the integral above as

$$\int_0^1 xe^{-ax}dx = -\int_0^1 \frac{d}{da}e^{-ax}dx$$

and we are free to take the derivative outside of the integral since the final answer will only depend on a anyway:

$$\int_0^1 x e^{-ax} = -\frac{d}{da} \int_0^1 e^{-ax} dx$$

The above expression tells us that we can evaluate the integral first, which is now a very simple integral to do, and then take the derivative of the resulting expression with respect to a . Thus, we have

$$\begin{aligned} \int_0^1 x e^{-ax} dx &= -\frac{d}{da} \left[-\frac{1}{a} e^{-ax} \Big|_0^1 \right] \\ &= -\frac{d}{da} \left[\frac{1}{a} (1 - e^{-a}) \right] \\ &= \frac{1}{a^2} (1 - e^{-a}) - \frac{1}{a} e^{-a} \end{aligned}$$

which agrees with the result obtained by integrating by parts.

VI. PROBABILITIES AND PROBABILITY DISTRIBUTION FUNCTIONS

When an event occurs in a random way, such as the toss of a coin or the roll of a die, we cannot predict the exact outcome of any single occurrence of the event, but we can often predict the *probability* that a given outcome will result. Probability is, therefore, the mathematical expression of chance. Suppose an event E has N possible outcomes with probabilities p_k , $k = 1, \dots, N$. First, each p_k must be a number between 0 and 1. If E is a coin toss, then $N = 2$ and $p_1 = p_2 = 1/2$. If E is the roll of a die, then $N = 6$, and $p_1 = p_2 = \dots = p_6 = 1/6$. Note that the probabilities do not all need to be equal! Probabilities are important in quantum theory, as one of the theories fundamental tenets is that the precise outcome of a given experiment cannot be determined, we can only know the probability of a given outcome.

The probabilities for the roll of a single die tell us that for any individual roll, the probability of rolling 1 is $1/6$, the probability of rolling 2 is $1/6$,... Suppose we ask what the probability is that a single roll turns up 1 *or* 2. Of all possible outcomes of the roll, the condition 1 *or* 2 constitutes $1/3$ of the possibilities, which tells us that in order to obtain the probability for one outcome *or* another, we should add the corresponding probabilities. In this case, we add $1/6 + 1/6$ and obtain $1/3$ for the probability of rolling 1 or 2. Extending this idea, we can ask what the probability is for rolling 1 or 2 or 3 or 4 or 5 or 6. Since this exhausts all possibilities, the answer is obviously 1. However, we could just naively add up the 6 probabilities and obtain the answer by direct calculation $1/6 + 1/6 + 1/6 + 1/6 + 1/6 + 1/6 = 1$. Thus, it is clear that the sum of the probabilities of *all possible* outcomes must be 1 because one of the outcomes must be obtained:

$$\sum_{k=1}^N p_k = 1$$

In order to determine, in general, the probabilities associated with different outcomes of an event E , we need to determine the number of ways n_k in which the k th outcome can occur and divide it by the total number \mathcal{N} of ways all outcomes can occur

$$p_k = \frac{n_k}{\mathcal{N}}$$

Since $\sum_k n_k = \mathcal{N}$, these probabilities add to 1. For example, if our event E is the toss of two identical coins, then there are three possible outcomes: two heads, heads and tails, and two tails. For the outcome “two heads”, this can occur in just one way – both coins turn up heads. The same is true for “two tails”. But for “heads and tails”, this can occur in two ways: coin 1 turns up heads and coin 2 turns up tails, or vice versa. Thus, there are four possible combinations, so the probabilities are $1/4$, $1/2$, and $1/4$ for “two heads”, “heads and tails” and “two tails”, respectively. If the coins were not identical, then we might distinguish the outcomes [coin 1 = heads, coin 2 = tails] and [coin 1 = tails, coin 2 = heads], in which case there would be four independent outcomes, each with probability $1/4$. This example illustrates an important point. Suppose we roll a die twice and ask what the probability is that the first roll turns up 1 and the second roll turns up 6. Since there are 36 possible combinations that we can obtain,

the probability of this single outcome is $1/36$, which is the product $(1/6)(1/6)$ for each individual outcome. The probability of obtaining outcome k in trial 1 and outcome l in trial 2 is the product $p_k p_l$ *only* when the two trials are completely independent of each other.

So far, we have been discussing probabilities associated with discrete outcomes. Now imagine random tosses of a ball such that the ball lands on the ground along a line that begins at $x = a$ and ends at $x = b$. If the ground is even and level in this region, then the probability that the ball finally comes to rest in an interval Δx ($\Delta x \ll b - a$) somewhere between $x = a$ and $x = b$ will not depend on where this interval is, i.e., the probability will be uniform. However, if the ground in this region is uneven and bumpy, then the probability of the ball coming to rest in some intervals will be greater than other intervals. That is, the probability should be described by a function $f(x)$, $x \in [a, b]$ that characterizes the high and low probability regions. The function $f(x)$ actually gives the probability per unit length for the ball coming to rest at various points in the interval $[a, b]$. $f(x)$ is, therefore, known as a *probability density* or *probability distribution*, and it must satisfy $f(x) \geq 0$ for all x . Thus, the probability that the ball comes to rest in a *small* interval Δx centered on the point x is $f(x)\Delta x$. More generally, the probability that the ball comes to rest in *any* interval $x \in [x_1, x_2]$, with $x_2 - x_1 < b - a$ is

$$\text{Probability} = \int_{x_1}^{x_2} f(x) dx$$

That is, we simply sum the probabilities $f(x)dx$ associated with all intervals dx in the region $[x_1, x_2]$ because we are calculating the probability of the ball coming to rest in the interval $[x_1, x_1 + dx]$ or in the interval $[x_1 + dx, x_1 + 2dx]$, or,.... Obviously, if the ball is restricted to end up somewhere in the interval $[a, b]$, then we have the condition

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

VII. GEOMETRIC SERIES

Let r be a real number such that $0 < r < 1$. A geometric series S_N of N terms is defined to be

$$S_N = \sum_{k=0}^{N-1} r^k$$

We will prove that the result of performing this sum is

$$S_N = \frac{1 - r^N}{1 - r}$$

In order to prove the result, we will use the technique of mathematical induction. That is, we assume the result is true for any N and prove that it is true for a series of $N + 1$ terms. Such a series is given by

$$\begin{aligned} S_{N+1} &= \sum_{k=0}^N r^k \\ &= \sum_{k=0}^{N-1} r^k + r^N \\ &= S_N + r^N \end{aligned}$$

Assuming the result for S_N , S_{N+1} becomes

$$\begin{aligned} S_{N+1} &= \frac{1 - r^N}{1 - r} + r^N \\ &= \frac{1 - r^N}{1 - r} + \frac{r^N(1 - r)}{1 - r} \end{aligned}$$

$$\begin{aligned}
&= \frac{1 - r^N + r^N - r^{N+1}}{1 - r} \\
&= \frac{1 - r^{N+1}}{1 - r}
\end{aligned}$$

which is the correct result for S_{N+1} . This proves the sum formula for S_N .

What happens if we let the series become infinite, i.e., we let $N \rightarrow \infty$? Since $0 < r < 1$, the limit as $N \rightarrow \infty$ of r^N is 0. Thus, the numerator in the sum formula reverts to 1, and we obtain

$$S_\infty = \frac{1}{1 - r}$$

VIII. DIFFERENTIAL EQUATIONS

A *differential equation* is an equation that expresses a relationship between a function and its derivatives. Differential equations are at the heart of physics and much of chemistry. The importance of a differential equation as a technique for determining a function is that if we know the function and possibly some of its derivatives at a particular point, then this information, together with the differential equation, can be used to determine the function over its entire domain.

As an example, consider the differential equation

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

for a function $f(x)$ defined for all $x > 0$. This is an example of a *first-order* differential equation, since it involves only the first derivative of f . Suppose we know that $f(0) = A$, where A is some number. This information, together with the differential equation, is sufficient to determine $f(x)$ for all x .

In fact, the equation can be solved essentially by inspection, since there is really only one function whose first derivative is equal to the function itself, and that is $f(x) = e^x$. However, it is worth reviewing the systematic procedure for solving the equation. The method is known as *separation of variables*. The method regards the function f and the variable x as independent objects that can be individually integrated, provided all of the f -dependent quantities be put on one side of the equation and all of the x -dependent quantities can be put on the other side. This can be easily accomplished by multiplying both sides by dx and dividing by f , which yields

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

Now, we are free to integrate both sides using the relations from the previous section:

$$\begin{aligned}
\int \frac{1}{f} df &= \int dx \\
\ln f + C &= x + C'
\end{aligned}$$

where C and C' are two constants of integration. These can be combined into a third constant C'' so that

$$\ln f = x + C''$$

Solving for f yields

$$f(x) = e^{C''} e^x = K e^x$$

where $K = e^{C''}$. We see that the solution is proportion to e^x as expected. However, we still have an undetermined constant of integration K . In order to determine K , we use the given information that $f(0) = A$. Evaluating the

solution at $x = 0$, we have $f(0) = K$. But since $f(0) = A$, we immediately see that the undetermined constant is $K = A$, and so the solution we seek that satisfies the given information is

$$f(x) = Ae^x$$

Next, we will consider an example of a *second-order* differential equation. A second-order equation contains the second derivative of f and possibly also the first derivative of f . As an example, consider the simple differential equation

$$\frac{d^2 f}{dx^2} = g$$

where g is a constant. Suppose we also know that $f(0) = A$ and $f'(0) = B$. Then, the equation can be solved uniquely for the required function $f(x)$. The solution of such an equation is simple, since it only involves $f''(x)$. All we need to do is integrate twice with respect to x on both sides. Integrating twice will, of course, bring in two unknown constants of integration, which is why both the function and its first derivative must be known at a given point x . Integrating once gives

$$\frac{df}{dx} = gx + C$$

and integrating again gives

$$f(x) = \frac{1}{2}gx^2 + Cx + C'$$

where C and C' are the constants of integration. Now, using the given information $f(0) = A$. But from the solution $f(0) = C'$, so we immediately see that $C' = A$. From the first derivative, we have $f'(0) = C$, but we also know that $f'(0) = B$, so we immediately have that $C = B$. Using this information, the solution becomes

$$f(x) = \frac{1}{2}gx^2 + Bx + A$$

As one final example, consider the second-order equation

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

with $f(0) = A$ and $f'(0) = B$. The solution procedure for such an equation is not as simple as for the first-order equation, so we will not present it here. However, fortunately this is an equation that can also be solved by inspection. There are two functions that satisfy the property that differentiating them twice returns the function with a minus sign. These functions are $\cos(x)$ and $\sin(x)$. In fact, since the equation is linear, we can construct a solution using *both* $\cos(x)$ and $\sin(x)$. Consider the solution

$$f(x) = C \cos(x) + C' \sin(x)$$

If we differentiate this twice with respect to x , we obtain $-C \cos(x) - C' \sin(x) = -f(x)$, so the equation is solved. However, we have two constants of integration C and C' to determine. Again, we use the given information about $f(0)$ and $f'(0)$. Since $\cos(0) = 1$ and $\sin(0) = 0$, $f(0) = C$. But since $f(0) = A$ is given, we see that $C = A$. Similarly, $f'(x) = -C \sin(x) + C' \cos(x)$, hence $f'(0) = C'$. But $f'(0) = B$, so we see that $C' = B$. Thus, the solution we seek is

$$f(x) = A \cos(x) + B \sin(x)$$

IX. FUNCTIONS OF SEVERAL VARIABLES

A function $f(x, y)$ of two variables x and y maps a point (x, y) in the xy plane onto a single number. Similarly, a function $f(x, y, z)$ maps a point in three-dimensional space onto a single number. Derivatives of functions of several variables must be performed with respect to one of the dependent variables. Such a derivative is called a *partial*

derivative. For a function $f(x, y)$ two partial derivatives can be defined with respect to x or with respect to y , and these are denoted

$$f_x \equiv \frac{\partial f}{\partial x}, \quad f_y \equiv \frac{\partial f}{\partial y}$$

The partial derivative f_x , for example, is performed by differentiating $f(x, y)$ with respect to x holding y fixed. Similarly, the partial derivative f_y is performed by differentiating with respect to y holding x fixed.

As an example, suppose $f(x, y) = x^2y^3$. The partial derivatives f_x and f_y are, therefore,

$$f_x = \frac{\partial f}{\partial x} = 2xy^3, \quad f_y = \frac{\partial f}{\partial y} = 3x^2y^2$$

In a similar manner, integrals of functions of several variables are defined by integrating with respect to each variable holding the other fixed. As an example, we take the function $f(x, y)$ above and ask what the value is of the following integral:

$$I = \int_0^1 dx \int_0^1 dy f(x, y) = \int_0^1 dx \int_0^1 dy x^2 y^3$$

We can perform the integrals in either order, so let's first integrate with respect to y holding x fixed. This gives

$$I = \int_0^1 dx x^2 \left. \frac{y^4}{4} \right|_0^1 = \int_0^1 dx x^2 \left(\frac{1}{4} - 0 \right) = \frac{1}{4} \int_0^1 dx x^2$$

Now we perform the x integral in the usual way:

$$I = \frac{1}{4} \int_0^1 dx x^2 = \frac{1}{4} \left. \frac{x^3}{3} \right|_0^1 = \frac{1}{4} \left(\frac{1}{3} - 0 \right) = \frac{1}{12}$$

X. POLAR AND SPHERICAL COORDINATES

The location of a point P in a plane is determined by specifying the *coordinates* of the point, as noted above. The simplest set of coordinates are the usual Cartesian coordinates (x, y) as shown in the figure below.

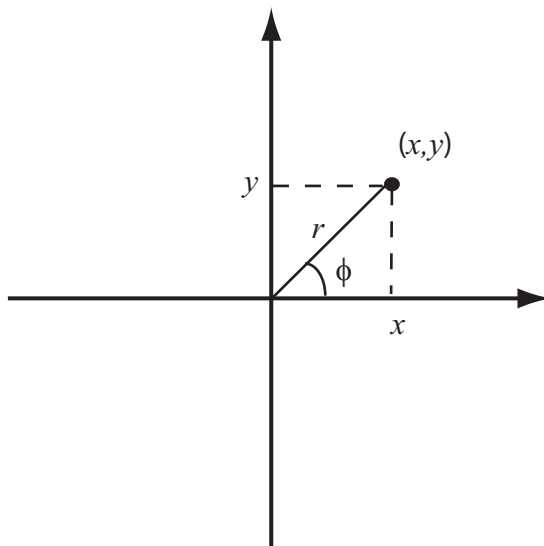


FIG. 16. Polar coordinates.

In many cases, it is convenient to represent the location of P in an alternate set of coordinates, an example of which are the so-called *polar coordinates*. In polar coordinates, the point is located uniquely by specifying the distance r of the point from the origin of a given coordinate system and the angle ϕ of the vector from the origin to the point from the positive x -axis. This representation is also shown in the figure above.

If we know the location of the point in one set of coordinates, other coordinates can be determined via a *coordinate transformation*. For polar coordinates, the transformation that determines r and ϕ given x and y is

$$r = \sqrt{x^2 + y^2}$$

$$\phi = \tan^{-1} \left(\frac{y}{x} \right)$$

Alternatively, if we are given r and ϕ , the Cartesian coordinates can be calculated from

$$x = r \cos \phi$$

$$y = r \sin \phi$$

Suppose we are given a function $f(x, y)$ of the two variables x and y . If the function is expressed in terms of r and ϕ , its form will be different, so we denote the transformed function as $\tilde{f}(r, \phi)$. From the transformation equations, it can be shown that the two dimensional integral

$$I = \int \int f(x, y) dx dy$$

can be re-expressed as an integral over the polar coordinates as

$$I = \int \int \tilde{f}(r, \phi) r dr d\phi$$

If the integral is taken over the entire x - y plane, then limits are added and

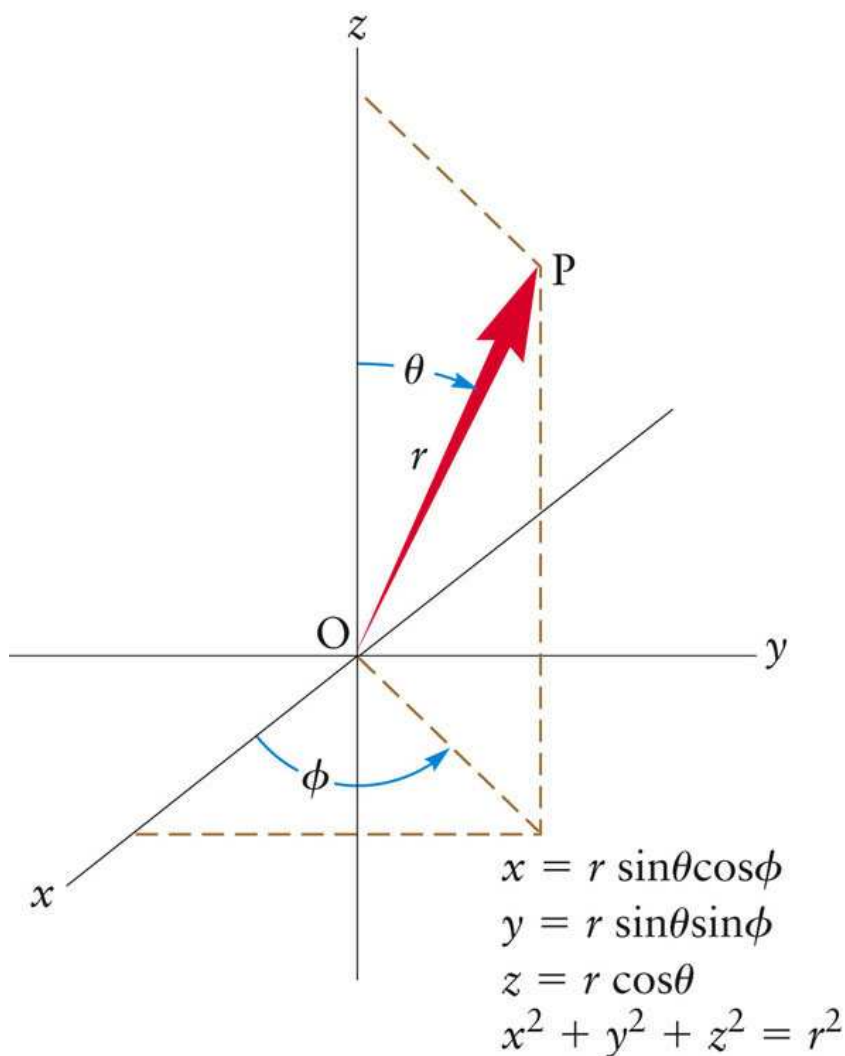
$$I = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy$$

In terms of polar coordinates the integral becomes

$$I = \int_0^{2\pi} \int_0^{\infty} \tilde{f}(r, \phi) r dr d\phi$$

That is, r is integrated from 0 to ∞ and ϕ is integrated from 0 to 2π .

In a similar manner, a point P in three-dimensional space can be located by specifying the three Cartesian coordinates (x, y, z) . Alternatively, the location can be specified by a generalization of polar coordinates to three dimensions called *spherical coordinates* or *spherical polar coordinates*. These are shown in the figure below.



© 2007 Thomson Higher Education

FIG. 17. Spherical Coordinates.

In three dimensions, the point P is located by specifying the distance of P from the origin of a given coordinate system, the angle θ of the vector from the origin to the point from the positive z -axis, and the angle ϕ of the projection of this vector into the x - y plane from the positive x -axis. The transformation from (r, ϕ, θ) to (x, y, z) is

$$\begin{aligned} x &= r \sin \theta \cos \phi \\ y &= r \sin \theta \sin \phi \\ z &= r \cos \theta \end{aligned}$$

and the transformation from (x, y, z) to (r, ϕ, θ) is

$$\begin{aligned} r &= \sqrt{x^2 + y^2 + z^2} \\ \phi &= \tan^{-1} \frac{y}{x} \\ \theta &= \tan^{-1} \frac{\sqrt{x^2 + y^2}}{z} \end{aligned}$$

From the transformation equations, it can be shown that an integral of a function $f(x, y, z)$

$$I = \int \int \int f(x, y, z) \, dx \, dy \, dz$$

in spherical coordinates becomes

$$I = \int \int \int \tilde{f}(r, \phi, \theta) r^2 dr \sin \theta d\theta d\phi$$

Here $\tilde{f}(r, \phi, \theta)$ is the transformed form of the function expressed in spherical polar coordinates. If the integral is over all space, then the limits in Cartesian coordinates are

$$I = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y, z) dx dy dz$$

and in spherical polar coordinates, this becomes

$$I = \int_0^{2\pi} \int_0^{\pi} \int_0^{\infty} \tilde{f}(r, \phi, \theta) r^2 dr \sin \theta d\theta d\phi$$

Note that θ is integrated from 0 to π only, which covers only the positive z -axis, but ϕ is integrated from 0 to 2π . Many of the functions $\tilde{f}(r, \phi, \theta)$ we will have to deal with are simple products of functions of one variable, i.e.

$$\tilde{f}(r, \phi, \theta) = g(r)h(\phi)v(\theta)$$

for this type of function, the three-dimensional integral in polar coordinates becomes a product of ordinary one-dimensional integrals:

$$I = \left[\int_0^{2\pi} d\phi h(\phi) \right] \left[\int_0^{\pi} d\theta \sin \theta v(\theta) \right] \left[\int_0^{\infty} dr r^2 g(r) \right]$$