EDMOND BROWN

Newton, Maxwell, Einstein

What Were They Thinking?

outskirtspress
DENVER, COLORADO

Newton, Maxwell, Einstein
What Were They Thinking?

V3.0

Outskirts Press, Inc.
http://www.outskirtspress.com

ISBN: 978-1-4787-5981-2

PRINTED IN THE UNITED STATES OF AMERICA

Table of Contents

Part 3-Einstein

Part I
Newton

Chapter 1
Gravity

In the year 2000 *Time* magazine chose Albert Einstein as the man of the century. It was a century with two world wars, great generals, great statesmen, and outstanding humanitarians. If *Time* had existed in the seventeenth century, there is little doubt that the editors would have chosen Sir Isaac Newton for similar praise. What was so significant about their theories? What bearing did they have on our lives? It seems as though only certain specialists can appreciate their contributions. Newton and Einstein were physicists. Is it possible for someone outside the field of physics to understand and appreciate their efforts?

These men were geniuses, and it is reasonable to assume that their theories are so complex that special training is required to understand them. Actually, their genius does not lie in the complexity of their thinking. It has more to do with their intellectual bravery. They were willing to give up certain ideas that the rest of the world took for granted. In the case of Einstein, it was the notion of time. If there were clocks all over the universe, it is generally thought that it would be possible to decide whether two events took place simultaneously. It turns out that this apparently obvious statement is not correct. A patient reader will see that the research of James Clerk Maxwell, in the nineteenth century, makes the commonsense notion about simultaneity suspect.

Before Newton's publication of his work, it was generally believed that the laws governing astronomical bodies were different from those dealing with earthly phenomena. Heavenly bodies, as they were called, were in perpetual motion. There must have been angels carrying the planets on their paths around the sun. Newton's intellectual bravery had to do with his willingness to dare think that there was a connection between falling bodies on the earth and the motion of the planets. He dared change the meaning of the words *velocity* and *acceleration* from their usual definitions to ones that would accommodate his theory. The simplicity of the equations that encapsulate his theories is amazing. These theories opened the doors to hundreds of physicists and astronomers, who applied them to all sorts of phenomena. The theory can be applied to the motion of gyroscopes, the propagation of sound waves, bridge building, and the weather, to name a few.

It is possible to get a good understanding of what these great men did by trying to imagine their thoughts as they tried to unravel some mystery. Their reasoning can be understood once prejudices are dropped. There are certain words that have specialized meanings to scientists, words like *velocity* and *acceleration,* for example. It is necessary to understand the reason for the change in order to understand the theory. Newton had a good reason for changing the meaning of these words. He was preoccupied with the motion of the planets. His laws of motion and of gravitation very likely resulted from his attempt to explain their orbits. From this it can be seen that physical laws are the outcome of an adventurous mind trying to solve a difficult puzzle. The connections between the research of Newton, Maxwell, and Einstein have led to an enormous cascade of physical thought. They also make an interesting story.

Newton's Laws

Isaac Newton was born in a small town in England, on Christmas Day in 1642. By that time, Magellan (1480–1521) had circumnavigated the earth, and Copernicus (1473–1543) had published his work indicating that the planets traveled in circular orbits about the sun. Subsequently, careful observations by the Danish astronomer Tycho Brahe (1546–1601), which were analyzed by his lab assistant, Johannes Kepler (1570–1630), indicated that planetary orbits were elliptical. Educated men understood that the apparent motion of the sun during the course of the day was a consequence of the earth's spin.

The scientist who very likely had the most influence on Newton was Galileo Galilei (1564–1642). Galileo, who died in the year of Newton's birth, was famous as a scholar and experimentalist. He is often credited with inventing the telescope, a doubtful claim, but he did develop it independently and used it to observe the moons of Jupiter. He publicized the idea that the earth was not the center of the universe, as was being claimed by the Catholic Church. It was just another planet, orbiting the sun, like the others. In fact, this idea led to his becoming a victim of the Inquisition. His later years were spent under house arrest by the church. The claims of Galileo were not essentially different from those of Copernicus, Tycho, and Kepler. However, the power of his influence was sufficient to convince many scholars of the validity of these claims, in spite of the attempts by the church to suppress them.

To Newton, Galileo's astronomical studies were less important than his work with gravitation. The ancient Greek philosopher Aristotle (384 BC–322 BC) had made the pronouncement that gravity was responsible for heavy bodies falling faster than light ones. Aristotle was such an authority on so many topics that almost two thousand years elapsed before anyone questioned the statement. What Galileo demonstrated was that gravity was not responsible for the differences in the motion of heavy and light bodies. If it weren't for the fact that air resistance has a bigger effect on a feather than a coin, they both would fall in precisely the same manner.

Try This Experiment

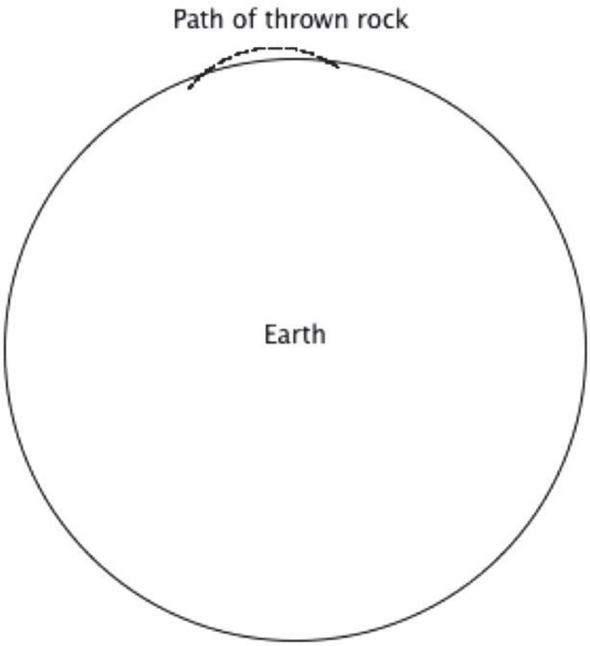
Galileo's conclusion concerning falling bodies can be made quite plausible. It isn't necessary to make use of a vacuum chamber or a pump. A bit of paper would fall as fast as a heavy book, if it weren't for the fact that air resistance has a bigger effect on the paper than it does on the book. To show this, try the following experiment: Take a book and hold it flat side up. Place a bit of paper on top of the book so that the book will run interference for the paper when they fall. Now drop the book. Hopefully, you've placed it over a bed or pillow for a safe landing. The paper will stay with the book. You will not notice any difference in the motion of the two. Why hadn't anyone done that in two thousand years? Perhaps most people don't question authority.

Galileo is one of the greatest scientists in history. He set a precedent for questioning authority and relying on carefully conceived experiments. His work set the stage for Newton. It is not known exactly whose influences shaped Newton's thinking the most, but he must have pondered about gravity and planetary orbits for a long time before he published his theories about motion.

Newton was also aware of the work of Rene Descartes (1596–1650), a mathematician, philosopher, and scientist who had stated what later came to be known as Newton's first law. The statement is that a body's natural state of motion is in a straight line at constant speed. Descartes had also introduced algebraic methods into geometry and introduced what are today known as Cartesian coordinates. These contributions will be shown to have played a role in Newton's deliberations.

Newton might have been thinking about planetary motion when he came up with his three laws of motion. Let's try to imagine what his thoughts were in light of what he eventually published. The story is told that he was influenced by the fall of an apple to think about gravity. Is it possible that a gravitational force toward the sun could cause a planet to go around it in a circle, or an ellipse? What did Newton already know about gravitational pulls? Galileo's work made it appear that all objects would fall in precisely the same way in a vacuum, regardless of size or weight. Kepler had found that the time it took for a planet to complete an orbit about the sun depended only on the planet's distance from the sun. Planets come in all sizes, and there didn't seem to be any dependence of planetary motion on size or mass. This behavior is consistent with Galileo's observations about gravity. No other mechanism suggested itself.

Galileo's work indicated that when an object falls, its speed increases at a regular rate. This is equivalent to motion with constant acceleration. Let's get in Newton's head once more. Is the acceleration constant because the force on it is constant? It is likely that the acceleration would be greater if the force were greater. Is there also a connection between the way things fall on earth and the way the planets move about the sun? If you throw a rock, it travels in a curved path. It hits the ground eventually. The harder you throw, the farther it goes before it hits the ground. Is it possible, if you threw it hard enough, that the curvature of its path would be less than the curvature of the spherical earth? If so, it would be in orbit about the earth! It was an exciting idea! It was not obvious how to deal quantitatively with this possibility.



How to deal with the curved trajectory of a rock in flight? Galileo had studied the vertical motion of an object when dropped. Is it possible that the downward motion of a rock traveling in an arc is the same as it is when dropped? What would the rock's motion look like to someone who tried to run directly below it while it was in flight? Would that person be running at constant speed in a straight line? From our vantage point today, that seems reasonable. If you were on an ocean liner traveling at 30 knots and decided to throw a rock straight up, it would return to your hands (if it were sheltered from the wind). The motion doesn't seem to be influenced at all by the speed of the boat. It would be straight up and down from the view of passengers on the boat but not from those on the shore. The possibility of separating out motion in two dimensions, vertical and horizontal, and treating each one separately, must have been appealing to a mathematician of Newton's caliber.

The idea is illustrated in the photograph below of a plane dropping bombs while in steady horizontal flight. Before the wind has a chance to affect the motion of the bombs, they remain vertically below the plane. From the pilot's viewpoint, the bombs have no horizontal motion at all.

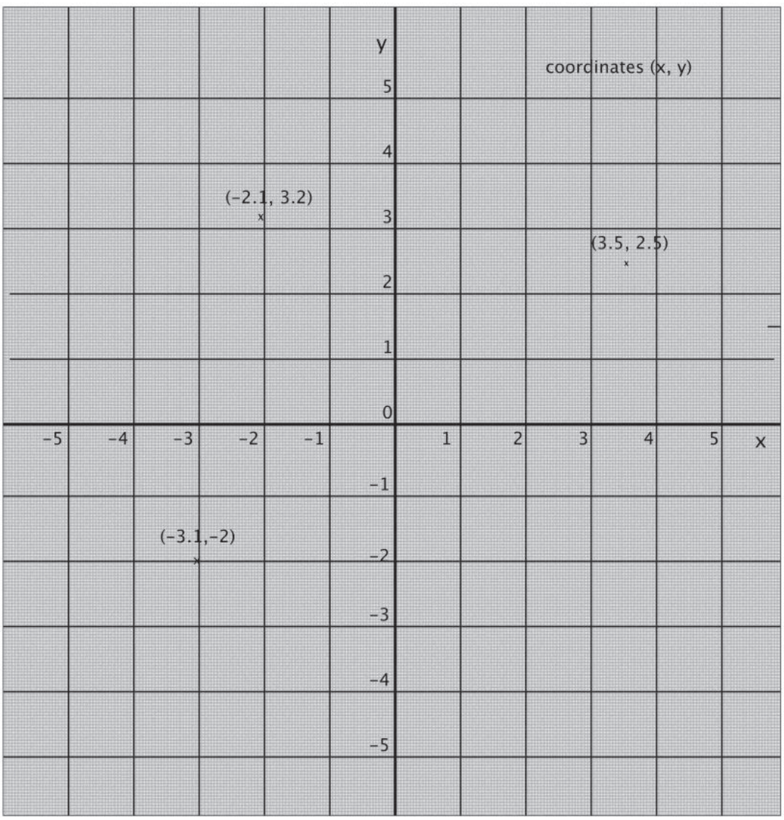
Coordinate Systems and Reference Frames

Newton knew the earth was spinning about an axis. Planetary motion is very complicated as viewed from the earth. Astronomers prefer to describe motion with respect to the stars. If you look up on a clear night, you will see constellations that seem to be fixed in place, as though they are embedded in a celestial dome. From our point of view, here on Earth, the dome is rotating. Planetary motion is easier to understand if the dome is considered as being still, while the earth is spinning beneath it. From the vantage point of this dome, planetary orbits look flat, approximating a circle, with the sun at the center. According to Kepler, it would be more precise to call the orbits elliptical, but it is simpler to deal with circles first.

Specifying planetary motion with respect to the earth presents other problems. The usual words that are used to describe direction, such as *northward*, *up*, or *down* can be misleading. Two people standing near, but on opposite sides of, the North Pole will be looking in opposite directions when both are facing due north. The same is true for two people facing east. Down to an Australian is not even nearly in the same direction as down to a Canadian. The words *horizontal* and *vertical* may have meaning locally but are misleading on a global scale. A person looking at the North Star from a point on the equator is looking horizontally. Another person viewing the same star from the North Pole is looking vertically. They are looking, almost exactly, in the same direction.

Because of the fact that the words *horizontal* and *vertical* don't correspond to fixed directions in space, the description of the motion of the bombs, in the photograph above, is not precise. If the plane were to fly around the earth, it would be clear it was not flying in a fixed direction. Moreover, if you were seated on the celestial dome, you would see that the plane was not traveling in a straight line, even if it were not moving with respect to the earth at all.

If an exact formula is wanted for describing motion, it would be simplest to express the motion with respect to a coordinate system embedded in the celestial dome. There are many possible alternatives for the choice of coordinates. On the earth, the position of an object is usually specified in terms of latitude, longitude, and altitude. The motion of an object can then be expressed by specifying these position coordinates in terms of the time. Descartes invented a more useful coordinate system. These Cartesian position coordinates are shown in the figure below. In actual usage, it would be necessary to provide units, such as feet, meters, and miles. The advantage of such a system, for Newton, was that the simplest motion of all, the motion of a body subject to no forces, could be expressed very simply. By a suitable orientation of the coordinate system, such motion could be expressed in such a way that only one coordinate depended on the time. The others would be constant.



Cartesian Coordinate System
The origin is at the intersection of the axes (coordinates 0, 0).

The use of negative numbers in Cartesian coordinates allows us to extend the coordinates as far as necessary. It is similar to the use of negative numbers to specify temperatures colder than zero degrees. For the general case of motion in three dimensions, a third coordinate perpendicular to the plane of the paper is necessary. A point in three-dimensional space could then be specified by the coordinates (x, y, z). Newton could imagine such a coordinate system embedded in some rigid framework. Such a structure could be used to specify the position of any point in the universe. The system is called a frame of reference. A body may be stationary as seen from one reference frame but be moving as seen from another. For example, if a coordinate system were embedded in a subway car, people seated in the car would be considered stationary while the subway stations could be moving.

Today, scientists are leery of extending coordinate systems outwardly without limit. Such an extension makes assumptions about the nature of space. It assumes that space is flat and infinite, in some sense. Scientists don't want to make the kind of mistake that people made when they assumed the earth was flat. To make this point a little clearer, try to imagine a straight-line journey through space in which you eventually come back to your starting point. If you find that difficult to do, you might be suffering from the same lack of imagination that plagued the believers in a flat earth. However, you cannot blame Newton for making the assumption that space is flat, at least in our neighborhood in the solar system.

Readers should be warned that mathematical equations are used throughout this book. They are included despite the warning by Stephen Hawking that an author loses half the potential readership for every equation included in a book. Physics and mathematics are closely intertwined, and to tell the story without math is to cheat a little. However, let the reader be reassured that it is possible to skim over the equations, and even ignore some of them, without much harm. In no case is the reader required to do any calculating, although a few exercises are suggested. Mathematical proofs are included for those readers who like to see how everything fits together. You should feel free to ignore these.

The Problem with Defining a Rate

(A Prelude to the Calculus)

Newton's thoughts about motion were completely new. Galileo's gravitational studies of motion had to do with motion in a straight line. They suggested that the speed of a freely falling object increased at a constant rate. How did Newton deal with the rate of change of speed? The notion of speed itself presents a mathematical problem. It is a rate, and rates involve a ratio, or a quotient, like miles divided by hours. If you are in a car going 60 mph (miles per hour), you expect to cover a distance of one mile in one minute. It is necessary to measure both a distance and a time to check whether your speedometer is accurate. How do you know if the speed has varied during that time? You would have to check by taking shorter time intervals and measuring shorter distances and computing the ratio again. But you still don't know for sure how fast you are going at any instant. An instant has no duration. You travel zero distance in zero time; division of zero by zero is a no-no to a mathematician. It is not determinate.

Newton wanted to deal with the rate of change of a rate. He asked himself, "Suppose I had a record of the position of a particle (an object small enough to be approximated by a single point) at all times. Is it possible to assign a speed to it at every instant along its path?" Newton was willing to consider motion in a straight line, since he believed he could focus on one Cartesian coordinate at a time. "If I could get a record of the speed at all times by some procedure, then it should be possible to use a similar procedure to get the rate of change of speed." Perhaps Newton was thinking of Galileo's conclusion that the distance covered by an object dropped from rest is proportional to the square of the time. This can be put in the form of an equation, as

$$x = At^2,$$

where x represents the distance fallen in time t. The symbols just mean that this distance is the product of some constant with the square of the time of fall. The square of the time, in turn, is just the product of t with itself. For example, when t takes on the values 1, 2, 3, 4, 5 for the time of fall in seconds, say, then t^2 takes on the values 1, 4, 9, 16, 25. The choice of the constant A depends on the units chosen for time and distance. If t is measured in seconds and x is measured in feet, A is approximately 16. It will be different if x is measured in meters. You can make it whatever number you want by inventing your own system of units. For example, A could have the value 1.00 by choosing the unit of distance to be the distance that the object falls in the first second. The unit of distance is then approximately 16 feet. Newton might have tried to get an estimate of the speed at a given instant, as well as its rate of change, by tabulating this information. He would have used the equation,

$$x = t^2,$$

in which t is measured in seconds, and distance is measured in units approximately 16 feet long. Suppose he tabulated this information and made use of an approximate procedure to get the speed and the acceleration as shown below.

t	0	1	2	3	4	5	6	7	8
x	0	1	4	9	16	25	36	49	64
v	1	3	5	7	9	11	13	15	
a	2	2	2	2	2	2	2		

Explanation: The top row indicates the value of the time of fall in seconds after being dropped. The next row is the distance fallen, according to the equation $x = t^2$ in the assumed unit of distance. Thus, the distance fallen at t = 5 seconds is 25 units, or approximately 25 times 16 = 400 feet. The third row, labeled v, is a crude estimate of the velocity at the given time. The estimate for v at t = 3, for example, is 7 units of distance per second. It is probably a better estimate for the speed at t = 3.5 because it is obtained by taking the distance covered between t = 3 and t = 4, namely 16 - 9 = 7, divided by the elapsed time of one second. Do not be concerned at this point that the value of v is not completely meaningful. The best that can be said for v is that it is a good estimate for the speed at a time between t = 3 and t = 4. The last row is labeled a for acceleration. Again it is an estimate. The value at t = 3 is obtained by finding the change in speed between t =3 and t = 4, then dividing by the elapsed time. The fact that it is constant suggests that it might be an accurate value in spite of the fact that the values of v are approximate. Thus, the table suggests that the acceleration is constant at a value of 2 units per second per second. In other words, the speed goes up by 2 units per second each second. Since 2 units is approximately 32 feet, this can be written as a = 32ft/sec^2. This is an abbreviated notation. Although it is read as 32 feet per second squared, it just means that the speed increases by 32 ft/sec each second.

Newton recognized the inadequacy of the tabular method for obtaining the speed at a given moment. All that has been obtained is the average speed in a given time interval, and then that average is used as the speed at the beginning of the interval. Clearly the error would have been smaller if shorter time intervals had been used. For example, if the speed at t =3 had been estimated by using the average speed in the interval between 3 and 3.5 instead of the interval between 3 and 4, it would have been better. The table below is constructed by the same method as before, except that the time is given at every half second.

t	0	.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5
x	0	.25	1.0	2.25	4.0	6.25	9.0	12.25	16.0	20.25
v	0.5	1.5	2.5	3.5	4.5	5.5	6.5	7.5	8.5	
a	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0		

The value for v at t = 3 is obtained by noting that in the interval between 3 and 3.5, it travels 12.25 - 9 = 3.25 units in one half second.

This yields an average of 6.5 units per second instead of 7. No matter how short an interval that is chosen, the tabular method is not reliable. Newton found a better method than the tabular one.

Differential Calculus

Consider the same motion, namely

$$x = t^2,$$

and focus your attention on one particular time, t. This is equivalent to consideration of one column of the tabular method. Suppose the symbol, s, denotes the number of seconds between columns in the table. In other words, an algebraic symbol replaces the numbers 1 and 0.5 that were used previously. The column in the table next to t is t + s. The entry for x in this column of the table is $(t + s)^2$. A formula is needed for this, which turns out to be

$$(t+s)^2 = t^2 + 2st + s^2.$$

The proof of this formula is simple, as will be shown shortly. The procedure Newton followed is the same as that of the tabular method. The distance traveled in the time between the time t and the time t + s is

$$(t^2 + 2st + s^2) - t^2 = 2st + s^2 = s(2t + s).$$

In order to get the average speed in this interval, it is necessary to divide by s. The result is

$$\mathrm{v} = 2t + s.$$

This result can be checked for the two tables considered earlier. In the first table, the formula is v = 2t + 1. In the second table, it is 2t + 0.5. You can check back to verify that this is indeed what was found earlier. For t = 3, the formula yields 7 and 6.5. The difference between the algebraic and the tabular methods is that the algebraic method specifies the result for any conceivable table. If an interval of 1 millionth of a second had been taken the formula would have been, v = 2t + 0.000001. It should be obvious that the formula v = 2t is an accurate measure of the velocity of the body at time t. The letter v will no longer denote an average value. If you have followed the arguments given here, you know what it means to differentiate with respect to time (alternatively: to take a time derivative.) The idea used by Newton, of taking a limit, avoided the problem of dividing zero by zero. It turned out to be an extremely useful idea in much of mathematics.*

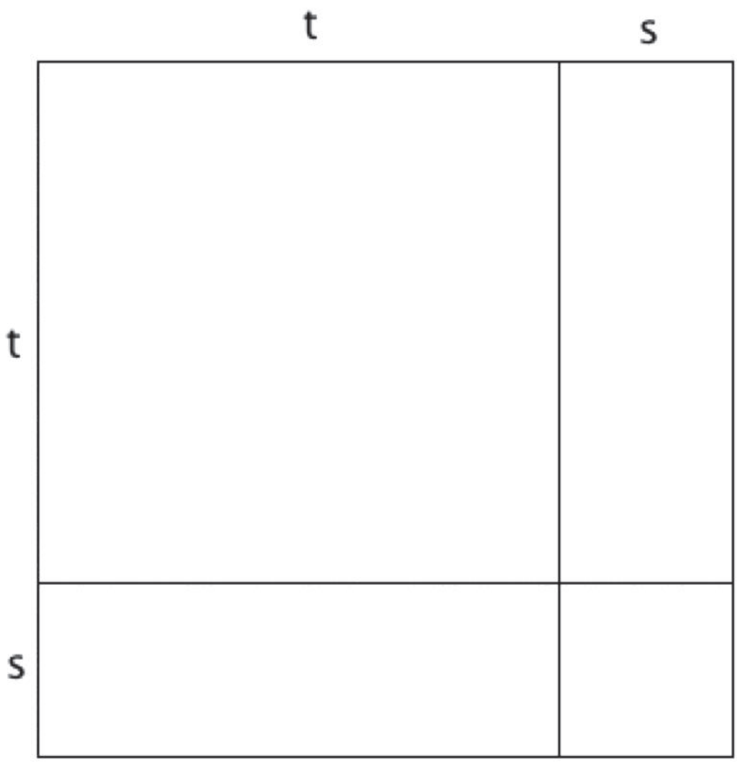
---

* Newton shouldn't be given sole credit for the invention of calculus. The mathematician Gottfried Leibniz (1646–1716) invented this technique independently. He made valuable contributions to the subject. It is his notation that is in most common use today. It is the one used in this book.

In order to prove the formula used earlier for the calculation of the square of a sum of two quantities, consider the diagram below. Each algebraic symbol denotes a length rather than a time. They are just numbers, as long as the units have been assumed (inches or feet, perhaps). The large square has a side of length $(t + s)$. Its area is thus $(t + s)(t + s)$, which can be written as $(t + s)^2$. This area is made up of two smaller squares and two rectangles. The areas of the enclosed squares are t^2 and s^2. The area of each rectangle is st. The area of the whole figure is equal to the sum of the areas of its parts, so that

$$(t + s)^2 = t^2 + s^2 + 2st\ .$$

Question: Can you draw a figure to show that $t^2 + 2st = t(t + 2s)$?



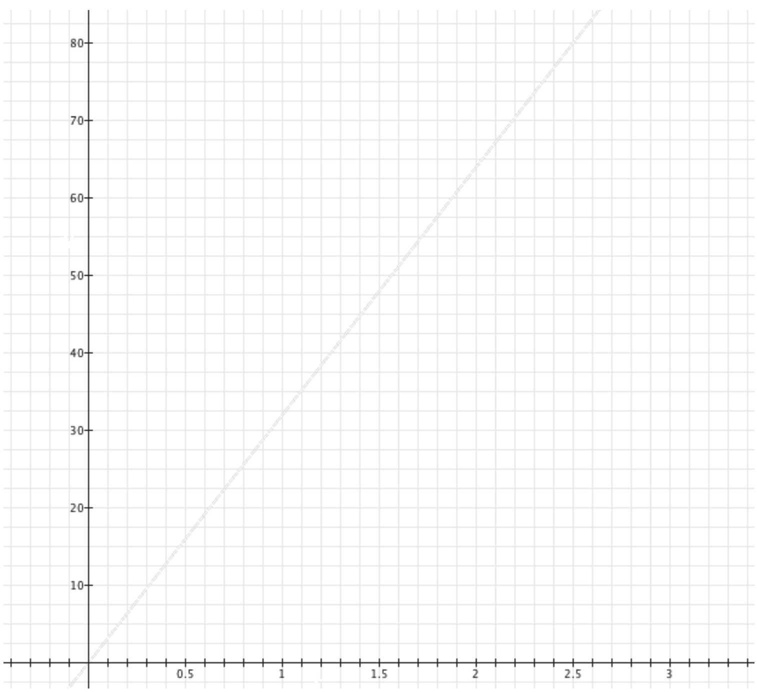
Newton found a method for calculating the velocity of an object at any instant of time. It assumed that a record of its position at all times was available. The same method could be used to calculate the body's acceleration from its speed. It involved a derivative with respect to time. In our present-day notation, the formula is written as

$$a = \frac{d\text{v}}{dt}.$$

The letter *d* designates a differential quantity, similar to the differences used in the tabular method. Thus, dv designates the change in velocity in the short time interval dt. These differential quantities are assumed to approach 0. The formula yields the same result that was found by the tabular method, namely a = 2. This is equivalent to an acceleration of 32 ft/sec^2.

Newton must have realized that Galileo's experiments with falling bodies was limited to motion in a straight line and restricted to a region near the earth's surface. It wasn't directly connected to the kind of motion he wanted to consider, but it must have made him aware that it was the acceleration that was the quantity of interest. Perhaps the acceleration of a falling body was constant because the force on the falling body was constant. After all, the weight of a body didn’t change from point to point, at least not noticeably in Newton's day. The problem, however, was that the acceleration didn't seem to work for planetary motion. The planets were traveling at nearly constant speed. If acceleration had to do with the rate of change of speed, the acceleration of the planets would be nearly zero.

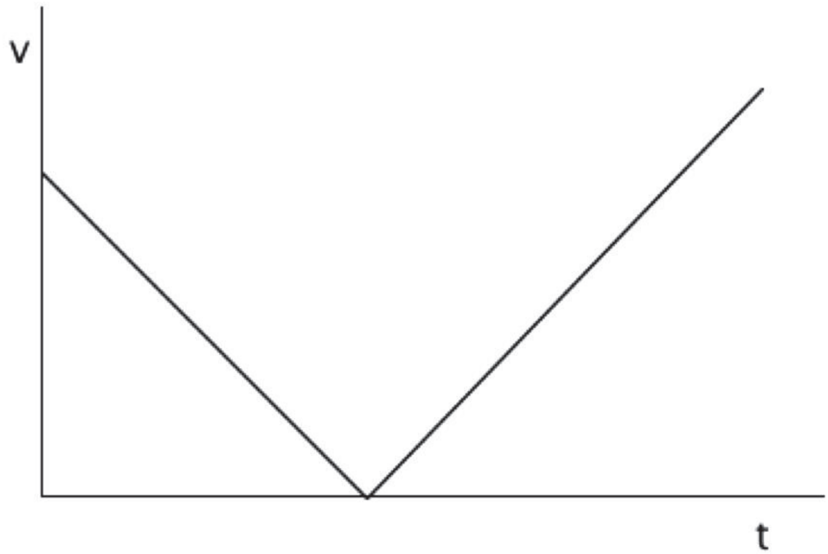
A graphical plot of v versus t, for the motion just considered, is a straight line, as shown below. The slope of the line is a measure of the rate at which v is changing and is, thus, proportional to the acceleration.



Velocity (or Speed) vs. Time in Free Fall

The Difference between Velocity and Speed

What would a plot of speed versus time look like if the object was thrown upward instead of falling from rest? Clearly its speed would be going down on the way up. It would be decelerating. A graph of speed versus time might look like the one sketched below. The speed is dropping as the particle approaches the top of its path. After that, it is expected to rise. There is a sudden change in the slope of the graph when v reaches 0. This corresponds to the top of the path, at the time when the body is turning around. A layman might say that the acceleration had changed abruptly. To a mathematician of the caliber of Newton, this abrupt change in slope is misleading.



Graph of Speed (not Velocity) vs. Time

The reason for the sudden change comes from the fact that acceleration is defined as the rate of change of speed. Speed, on the other hand, is not really a rate of change of anything. It is never negative. The speedometer in a car starts at 0. It specifies the magnitude (the size) of a rate of change but not the direction. A rate of change can be negative as well as positive. The difference between the two can be important. Consider the rate of change of your bank account, for example. If you are told it is ten thousand dollars a year, wouldn't you want to know whether this rate is positive or negative? If the altimeter reading on your plane is changing very rapidly, you certainly would want to know if the rate is positive.

The words *velocity* and *speed* have been treated as synonymous. The physicist makes a distinction between these two words. If the upward direction is treated as being positive, for example, motion in the downward direction can be specified by a negative velocity. Thus, if the velocity had been plotted, rather than the speed, in the previous graph, there would be no break in the slope of the line. It would be one continuous straight line. The right-hand part of the graph would be replaced by its mirror image, extending below the x-axis. The acceleration for this motion is negative, or downward, throughout the motion. Such a motion has constant acceleration. There is no need to use the word *deceleration* to describe a motion, since a given acceleration can be associated with speeding up or slowing down. Since Newton's time, physicists have defined acceleration as the rate of change of velocity, not the rate of change of speed.

The change in the meaning of the word a*cceleration* must have opened up a world of possibilities for Newton. It was now possible for the planets to be accelerating even though their speed was nearly constant. He needed a definition of acceleration that was valid for all sorts of motion. He also needed to examine the implications of his new way of thinking. If acceleration was the rate of change of velocity, not speed, he needed a definition of velocity that was valid for motion in three dimensions.

It is necessary to reconsider the earlier motion, in which x denoted the distance an object fell. Like the word *speed*, the word *distance* is restrictive. It cannot take on negative values. It is more appropriate to think of x as a measure of position. In the one-dimensional motion of a vertical fall, it could represent the altitude of a body, measured from some reference point. If g were used to designate the acceleration due to gravity, the equation of motion would be as follows:

$$x = x_0 + \mathrm{v}_0 t - \frac{1}{2} g t^2 .$$

Here, x_0 is the starting position coordinate; v_0 is the initial velocity (positive if upward). The last term is negative because g is conventionally listed as a positive number, whereas the acceleration is negative (downward). In this case, the velocity is specified by the formula

$$\mathrm{v} = \frac{dx}{dt} = \mathrm{v}_0 - g t .$$

Newton realized these equations were approximate. The earth is spinning, and g is not really a constant. He needed to use a frame of reference in which the celestial dome was fixed in which to embed his coordinate system. He relied on the work of Rene Descartes and chose three Cartesian coordinates to specify the position of a planet. It was reasonable to ignore the motion of the sun and to choose the origin of his coordinate system at the solar position. Three numbers are needed to specify the velocity once the time dependence of the position coordinates are known. They are given by

$$\mathrm{v}_x = \frac{dx}{dt}, \quad \mathrm{v}_y = \frac{dy}{dt}, \quad \mathrm{v}_z = \frac{dz}{dt} .$$

In a similar way the acceleration has three components given by

$$a_x = \frac{d\mathrm{v}_x}{dt}, \quad a_y = \frac{d\mathrm{v}_y}{dt}, \quad a_z = \frac{d\mathrm{v}_z}{dt} .$$

Once Newton had changed the meaning of the words *velocity* and *acceleration*, it was necessary to find rules that described the motion of bodies. He was guided by a statement made earlier by Descartes.

Newton's First Law of Motion

Newton thought of gravity as a pull, a force. For the apple, it was the earth that was doing the pulling; for the planets, it was the sun. What is the relation between force and motion? A falling object had almost constant acceleration in Galileo's experiments because the force on it was almost constant near the surface of the earth. A massive object like a bull or an elephant needs a much greater force, for a given acceleration, than a bird or a feather. In this context, mass denotes the inertial properties of a body (the tendency to resist a change in velocity). Perhaps the acceleration is inversely proportional to the mass. If motion under gravitational forces doesn't depend on mass, maybe it is because the gravitational force is proportional to the mass. In a sense, the gravitational property of mass and the inertial property of mass cancel, leading to motion that doesn't depend on mass when the only force on a body is gravitational.

What would motion look like if there were no forces on an object at all? Here on Earth, friction is encountered everywhere. Moving objects can be slowed by the underbrush, the air, or a rough terrain. Wouldn't motion look different in a region where there was no friction and far from any gravitational affects? Newton could easily imagine that the planets were traveling through a vacuum, a region devoid of friction. Would a planet travel in a curved path if there were no sun? Which way would it curve? It would very likely move in a straight line and not slow up. The idea wasn't original. Rene Descartes said the same thing years earlier. Nevertheless, it came to be known as Newton's first law. A fairly precise statement might read as follows:

The natural state of motion for a body is to travel in a straight line at constant speed. This is equivalent to the following: The natural state of motion is of constant velocity, or zero acceleration.

Although not explicitly stated, this law implies that you are considering motion as viewed from a specific vantage point, not on the spinning earth. In fact, physicists interpret the first law as implying that there exist reference frames in which bodies subject to no forces will coast in straight lines at constant speed. If one such frame of reference can be found, it is possible to find another. It would be a frame that is coasting at constant velocity with respect to the first and not spinning with respect to it. All these special frames in which the first law is valid are called inertial reference frames. If you were on such a frame of reference, the constellations would not seem to be moving at all. You would have to imagine the sun was very far away so that it didn't affect your coasting through the cosmos at a steady speed in a straight line.

Up to this point, an attempt has been made to get into Newton's head and imagine the kind of problems that were occupying his thoughts. His laws of motion are elegant attempts to answer the questions he was asking himself. However, these laws are deceptive in that every term he used has to be examined carefully in order to be understood.

The Second Law

Newton's second law is one of the most famous laws in all of physics. It is usually written as F = ma and stated as "force equals mass times acceleration." However, this formulation is misleading. It is better to state it in the form

$$a = F/m.$$

In words, it becomes “the acceleration of a body equals the net force on it divided by its mass.” This is mathematically equivalent to the first way of writing it, but it helps us realize that it is the acceleration of a body that is determined by the net force acting on it, not the other way around. The forces acting on the body are determined by other factors.

A remark about the second law is in order. A constant, k, should be inserted into the equation. Thus, the equation becomes a = kF/m. This constant can be eliminated by the proper choice of units. One such system, in common use by scientists, is the MKS system. In this system, the unit of length is the meter (39.37 in), the unit of mass is the kilogram, and the unit of time is the second. In this system, the unit of force is called the newton. A force of one newton is that force that will give a 1 kilogram mass an acceleration of 1 meter/sec^2. A 1 kg mass weighs about 2.2 lbs on the surface of the earth. A 2.2 lb. force will give a 1 kg mass an acceleration of 9.8 meter/sec^2 (32 ft/sec^2). From these facts you should be able to compare the force of 1 pound with that of 1 newton. Hint: The pound is bigger.

It is hard to imagine an equation that looks any simpler than the second law. It turns out to be amazingly subtle and amazingly useful. Let us look first at the case where there are no forces acting on a body. It then follows that the acceleration is zero. It seems as though the first law isn't necessary since the second law tells us the same thing. Remember, however, that the first law deals with special frames of reference called inertial frames. The second law is only valid with respect to these special reference frames. If you use it in any other case, you are either making an approximation or a mistake. Thus, strictly speaking, Newton's Law isn't valid with respect to a reference frame fixed on the earth. It is often a good approximation, however, to use it on the earth for phenomena that last only a matter of seconds. The earth doesn't turn very much in that short a time.

It is important to examine the meaning of each symbol in the second law. They are taken up in detail here.

Force: Of the three terms involved in the second law, force is the most complicated. In Newtonian physics, all forces on a body arise out of interactions with other bodies. There are no exceptions. You probably have heard explanations of why water doesn't spill out of a bucket that is being whirled in a vertical circle in terms of centrifugal force. It is advisable to forget these explanations, as they are almost always wrong. (To be sure, if you want to take the bucket as a frame of reference, it is possible to do so by introducing fictitious forces, such as centrifugal force. For the time being, accept the fact that these forces are fictitious. They have been introduced in order to make use of non-inertial frames of reference.)

Physicists today tell us of four basic forces: strong, weak, electromagnetic, and gravitational. Of these, Newton was only familiar with the last. For our present purposes, it is most useful to divide forces into two categories: 1. contact forces 2. action-at-a distance forces. The only one in the latter category with which Newton concerned himself is that of the gravitational force. The rest of the forces come about because of deformations of bodies in contact with one another. Air resistance can even be considered as such a force. It is caused by large numbers of gas molecules hitting a body each second and deforming as they make contact.

It is important to recognize that force is a vector quantity, which means that it has a direction as well as a magnitude. When a number of bodies are exerting forces on a body, it is necessary to sum these vectors to determine the net force. Forces can be specified in terms of their Cartesian components. The x-component of a sum of forces, for example, is just the sum of the x-components of the individual forces.

Mass: This is a property of a body. It doesn't depend on a body's location. Newton assumed that a bag containing a dozen baseballs would have twice the mass of a bag containing six baseballs. If the accelerations of the two bags are the same, the force on the bigger bag must be twice as great as the force on the other. (Granted, Newton never saw a baseball.) Since the two bags fall in precisely the same way in a vacuum, the weight force on the bigger bag must be double the weight of the smaller one. In fact, the weight of a body must be proportional to the mass, since, according to Galileo, the acceleration of free fall, g, is the same for both bodies. According to the second law, when applied to a freely falling body, then

$$W = mg.$$

This equation is valid whether the body is falling or not. The mass of a body doesn't change unless a body loses or gains material, such as by evaporation or chemical reaction. The weight, which is the magnitude of the earth's attractive force on the body, does change with position, and so does g. Note that the weight force is truly a vector (pointing downward), and so is g. Mass, on the other hand, has no directional properties. It is a scalar.

There are two distinct properties associated with mass. On the one hand, it is a measure of a body's inertia. This means that for a given force, a body with twice the inertial mass of another body will have half the acceleration. On the other hand, it is also a measure of the gravitational force on a body in a given location. A body with twice the gravitational mass of a second body will have twice the weight at a given location. Newton didn't distinguish between the two aspects. Gravitational mass and inertial mass were one and the same. Today, some scientists wonder why such a coincidence should occur. Why doesn't mass matter (no pun intended) when something falls? Experiments have been carried out to tremendous accuracy in recent years. As of this writing, no measurable difference between the two types of mass has been found. No distinction will be made between the two meanings. A scale can be used to compare masses. The kilogram mass has been defined as the mass of a platinum-iridium cylinder stored in Sevres, France. Secondary standards are in use in different countries. There is some concern that the materials aren't stable enough, and some better standards are necessary. Scientists are willing to revise standards when something better becomes available. Such a revision will be encountered when the unit of time, the second, is defined.

Acceleration: Acceleration has been defined as the rate of change of velocity with time. It is assumed that these are measured in some inertial frame. Velocity, in turn, is defined as the rate of change of position with respect to time.

"So, what good are Newton's laws to us, here on the earth, if they don't work because of the earth's rotation?" you could ask. One answer might be that they work well enough for many purposes, because our planet turns so slowly. Another answer might be "If a body's motion with respect to an inertial frame can be measured, it is then possible to specify its motion with respect to the earth." This is the more important answer. It is a straightforward matter to adapt Newton's laws so that they apply to bodies on the earth. Far from treating the earth as a special place at the center of the universe, Newton took us into a new world. You might say that the space age began with him.

Why the Need for Particles in the Second Law

It is necessary to make a slight alteration in what has been written so far. Newton realized that bodies have spatial extent and that all the parts might not be moving in the same way. In a moving automobile, for example, the different parts of the wheels are traveling at different velocities with respect to the ground. The velocity of a body, or its acceleration, has no meaning if different parts of the body are moving differently. In order to deal with such a situation, Newton imagined that it is possible to divide the body into such small parts that variations of velocity within each part are negligible. Newton referred to these bits of matter as particles and applied his laws to each of them. Newton's laws then contain the word *particle* instead of *body.* In Chapter 4, it will be shown that Newton's laws can be applied to an extended body as long as reference is made to the acceleration of a particular point, known as the center of mass of the body. This is useful in showing that the orbits of planets are not affected by planetary rotation. Since a planet is spherically symmetrical, its center of mass is at its geometrical center, regardless of planetary rotation. Planets can be treated as though they were particles, as far as their orbits are concerned.

The Third Law

Newton's third law is simple to state, but it is counterintuitive to many (perhaps most) people. A fairly precise statement might read as follows: "Whenever body A exerts a force on body B, then body B exerts a force on body A that is equal in magnitude but in the opposite direction." Remember, the first word is *whenever*. There is nothing implied about equilibrium. There are no delays. There are no statements to the effect that one body causes the interaction. The two bodies are on an equal footing. There are two and only two. A shorthand statement of the third law is "action equals reaction." However, since the two bodies involved are on an equal footing, it is arbitrary as to which force is the action. The shorthand statement also doesn't emphasize the important fact that the two forces act on (and are caused by) different bodies.

What led Newton to think of the third law? It is interesting to speculate on the type of reasoning he might have used. The motion of the planets, as described by Kepler's laws, is extremely regular, dependent only on the distance from the sun. If it is assumed that planets are subject to internal forces, like those on the earth, associated with storms, avalanches, volcanoes, and earthquakes, they have no discernible effect on the overall motion. As will be seen in Chapter 5, the motion of the center of mass of a body is unaffected by internal forces. This result is a consequence of Newton's third law. If one part of a body could push on another part, without feeling an equal size reaction force, it would be possible to change the motion of the center of mass of the body, without the need for external forces. When a young child attempts to lift herself off the ground by pulling upward on her feet, she is unaware of Newton's third law.

Scalars and Vectors

Physical quantities require numbers and units. A quantity like that of mass requires only one of each, but three numbers have to be specified in the three-dimensional world to specify position, velocity, and acceleration. Actually, these three numbers depend on the frame of reference and the choice of coordinate system. Such quantities are called vectors, while those quantities, like mass, that don't depend on the coordinate system are called scalars. Quite frequently the motion of a particle is confined to a plane. In such a case, it is convenient to choose a coordinate system embedded in that plane. The z coordinate can be ignored, for example, and the vector can be expressed in terms of the x and y components.

Today, physicists often display vectors as arrows. For example, if the components of velocity are given as

$$\mathrm{v}_x = -4\,ft/\sec, \quad \mathrm{v}_y = -3\,ft/\sec,$$

the velocity vector can be specified by means of an arrow, as in the diagram below. The position vector can be specified by an arrow, from the origin of our coordinate system to the coordinates of the point in question. Boldfaced symbols are used in the text to distinguish vectors from scalars. Thus, the velocity vector is denoted by the letter **v.** The speed of the particle, v, is just the magnitude of this vector. The Pythagorean theorem, (proven in the next section) can be used to obtain the result

$$\mathrm{v} = \sqrt{\mathrm{v}_x{}^2 + \mathrm{v}_y^2} = \sqrt{16+9} = \sqrt{25} = 5\,ft/\sec.$$

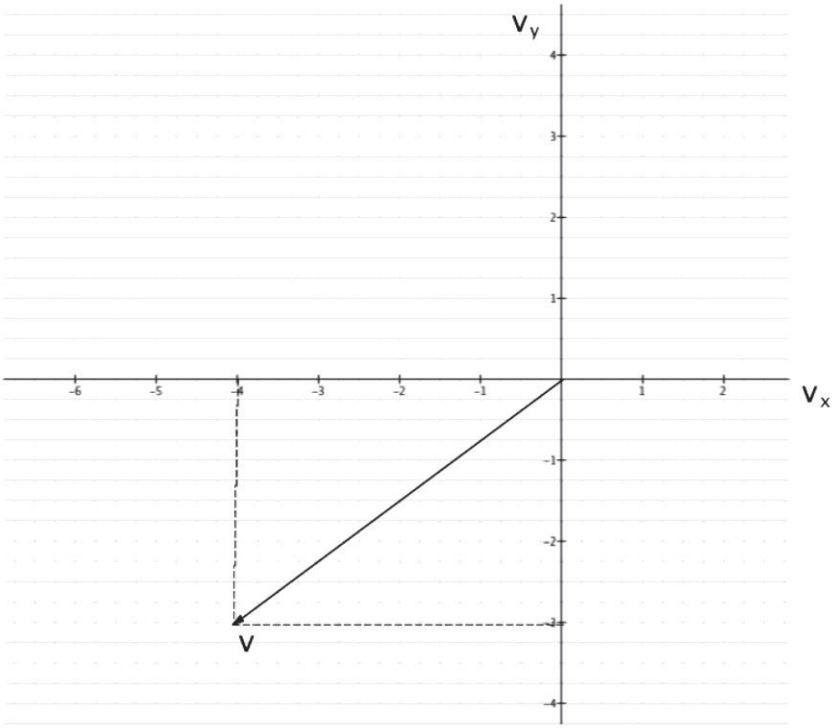
In the most general case, the formula is

$$\mathrm{v} = \sqrt{\mathrm{v}_x^2 + \mathrm{v}_y^2 + \mathrm{v}_z^2}.$$

Note that Newton's second law actually consists of three equations, namely

$$a_x = F_x/m, \quad a_y = F_y/m, \quad a_z = F_z/m,$$

with reference to some Cartesian coordinate system embedded in an inertial frame of reference.



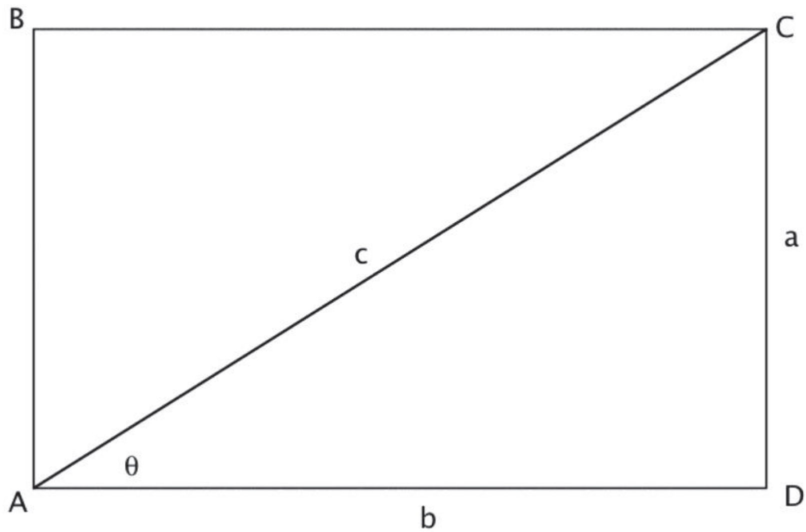
Velocity Vector as an Arrow

Right Triangles

This section is devoted to a few definitions and a few mathematical derivations. You might find these helpful in dealing with the more interesting subject matter of the next chapter. If you are not interested in derivations, you might prefer to skim over the equations that define certain trigonometric quantities.

The figure below consists of a rectangle, ABCD, that has been divided into two identical triangles, ABC and ADC. Let's focus our attention on the triangle ADC and note that the Greek letter *theta* (θ) denotes the angle at the vertex A of this triangle. (Physicists often use Greek letters to denote angles.) The side opposite to vertex A, CD, is of length a, and the side adjacent to it, AD, is of length b.

The angle at the vertex D is a right angle. (The angles at all the corners of a rectangle are right angles.) The side opposite this, AC, of length c, is called the hypotenuse of the triangle. Every triangle that contains a right angle is called a right triangle. If this figure were scaled up or down, each length would change by the same factor. However, the angles wouldn't change. Since the ratios of these lengths are unaffected by scaling, the fraction a/c depends only on the angle θ.



The common unit of angular measure is the degree. It is defined in such a manner that a right angle is 90 degrees, written as 90^0. It follows that the sum of the angles in the rectangle is 360^0, and the sum of the angles in the triangle is 180^0. (Many readers probably already know that the sum of the angles doesn't depend on the presence of right angles.) The fact that the sum of the angles in the two triangles is the same has been used.

There are certain ratios that are used so frequently that they are given names. The ratio a/c, which depends only on θ, is called the sine of the angle θ. The letter *e* is omitted from the word *sine* when writing this in an equation. Thus, the equation is written as

$$\sin\theta = \frac{a}{c}.$$

The equation for the cosine is written as

$$\cos\theta = \frac{b}{c}.$$

One more ratio is in common use. It is called the tangent and is defined as the opposite over the adjacent, so that

$$\tan\theta = \frac{a}{b}.$$

The definitions given by the previous three equations only work for angles less than 90^o. The definitions will be extended in the next chapter so as to be applicable to any angle. The three trigonometric functions are related to one another. Knowledge of one of them determines the other two. In particular, it will be shown that

$$\cos\theta = \sqrt{1-\sin^2\theta}, \qquad \tan\theta = \frac{\sin\theta}{\cos\theta}.$$

Note that the superscript 2 is placed where it is in order to denote that the sine of the angle is to be squared, not the angle.

An alert reader might be able to show that the last equation can be derived from the previous equations in this section. The first depends on the Pythagorean theorem.

Exercise: Show from the fact that the sum of the angles in a triangle is 180^0 the following:

$$1. \sin(90^0 - \theta) = \cos\theta, \text{ and } 2. \ \cos(90^0 - \theta) = \sin\theta \ .$$

The Pythagorean Theorem

The mathematical theorem that will be proved here dates back to the interesting Greek mathematician and scholar Pythagoras (570 BC–495 BC). The theorem can be stated very simply in mathematical terms. Referring back to the triangle of the last section, the statement of the law is as follows:

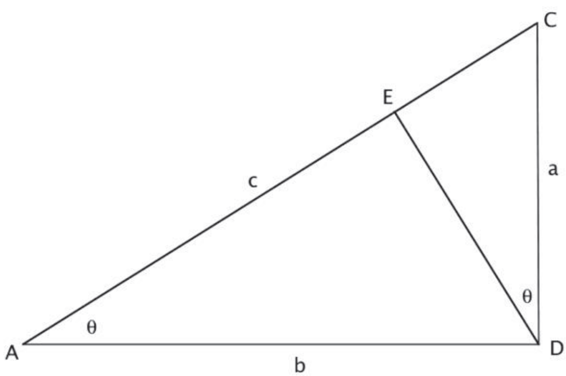
$$a^2 + b^2 = c^2 .$$

Numerous proofs have been given, some of which are quite complicated. Two proofs are given here. If one seems too difficult, try the other.

Proof #1.

In the figure below, the original triangle, ADC, is redrawn. A line, DE, is drawn that makes a right angle with the hypotenuse of the original triangle, AC. Let us show the reasoning that allows us to conclude that the angle specified by EDC is the same as the angle CAD, namely θ. Consider triangle ADE. From the fact that the sum of the angles is 180^0, it is found that the angle ADE + θ = 90^0. However, angle ADE and angle EDC add up to 90^0 also, because angle ADC is a right angle. Note that all three triangles, ADE, EDC, and the original one, ADC, contain the same set of angles.

The proof of the Pythagorean theorem now consists of two parts. Note that the area of the original triangle is the sum of the areas of the other two triangles.



The second part consists of expressing these areas in terms of the hypotenuse and the angle θ. Let's recall that the triangle under consideration was obtained by dividing a rectangle into two equal parts. Since the area of a rectangle is obtained by multiplication of the lengths of two adjacent sides, it follows that

Area of ADC = ab/2.

This can be expressed in terms of the hypotenuse by use of the relations

$$a = c \sin \theta, \qquad b = c \cos \theta,$$

so that

$$\text{Area of Large Triangle} = c^2 \sin\theta \cos\theta / 2.$$

Since there is nothing special about the large right triangle, a similar formula holds for any right triangle in which one of the angles is θ. All such triangles are similar, differing from one another only in scale. This result is applicable to the three right triangles in the figure, since they all contain the same three angles. In other words, the area of each triangle is given by a formula of the form area = k times the square of the hypotenuse, where $k = \sin\theta \cos\theta / 2$. The rest of the proof is simple:

Area of ADC = Area of DCE + Area of ADE

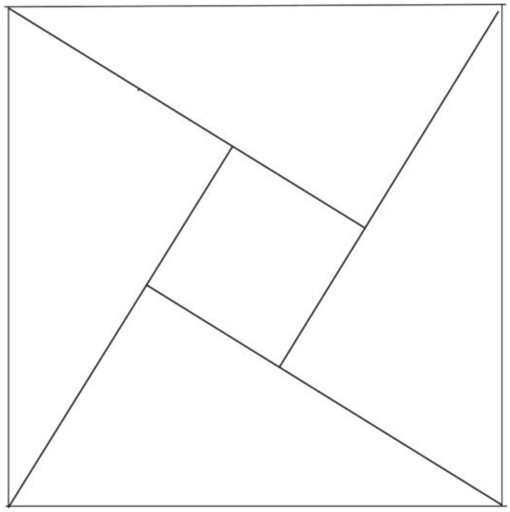
$$kc^2 = ka^2 + kb^2.$$

Division by k leads to the Pythagorean theorem;

$$c^2 = a^2 + b^2.$$

Proof #2.

Consider the original triangle, ADC, and imagine that you have cut out four pieces of cardboard exactly like it. Assume you have assembled these four triangles as in the figure below.



Readers should be able to convince themselves that the four triangles assemble into a square with a square hole. The larger square has an area given by c^2. The square hole has a side of length (b - a), under the assumption that b is larger than a. The hole has an area given by $(b - a)^2$. Recall the formula for $(t + s)^2$ and replace t with b and s with -a. Thus:

$$\text{Area of hole} = b^2 + a^2 - 2ab.$$

The area of the four triangles, which make up the rest of the large square, is just 2ab. Once again it is found that

$$c^2 = a^2 + b^2.$$

Consider the given triangle, ADC, once more and note that

$$\sin\theta = \frac{a}{c}, \qquad \cos\theta = \frac{b}{c}.$$

Squaring both sides and summing leads to

$$\sin^2\theta + \cos^2\theta = \frac{a^2 + b^2}{c^2} = 1,$$

so that

$$\cos\theta = \sqrt{1 - \sin^2\theta}.$$