Calculus is Everywhere # 1
Graphs Tell It All

The graph of a function conveys a great deal of information quickly. Here are four examples, all based on numerical data.

The Hybrid Car

A friend of ours bought a hybrid car that runs on a fuel cell at low speeds and on gasoline at higher speeds and a combination of the two power supplies in between. He also purchased the gadget that exhibits "miles-per-gallon" at any instant. With the driver glancing at the speedometer and the passenger watching the gadget, we collected data on fuel consumption (miles-per-gallon) as a function of speed. Figure C.1.1 displays what we observed.

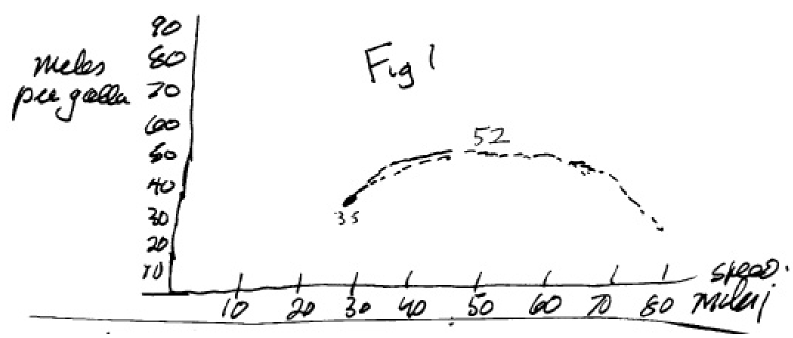


Figure C.1.1: ARTIST: Please extend the vertical axis to include 100.

The straight-line part is misleading, for at low speeds no gasoline is used. So 100 plays the role of infinity. The "sweet spot," the speed that maximizes fuel efficiency (as determined by miles-per-gallon), is about 55 mph, while speeds in the range from 40 mph to 70 mph are almost as efficient. However, at 80 mph the car gets only about 30 mpg.

To avoid having to use 100 to represent infinity, we also graph gallons-per-mile, the reciprocal of miles-per-gallon, as shown in Figure C.1.2. In this graph the minimum occurs at 55 mph. And the straight line part of the graph on the speed axis (horizontal) records zero gallons per mile.

Life Insurance

The graphs in Figure C.1.3 compare the cost of a million-dollar life insurance policy for a non-smoker and for a smoker, for men at various ages. (By defini-

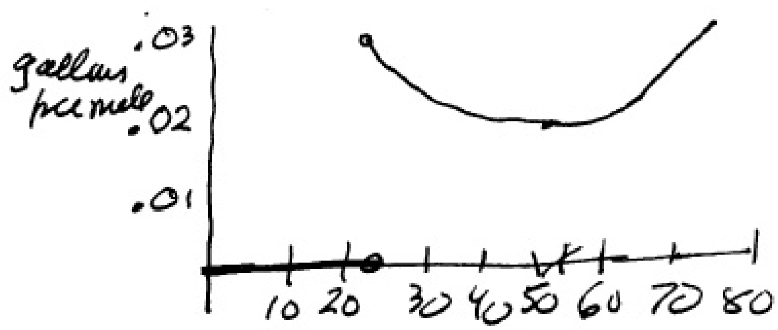


Figure C.1.2:

tion, a "non-smoker" has not used any tobacco product in the previous three years.) A glance at the graphs shows that at a given age the smoker pays almost three times what a non-smoker pays. One can also see, for instance, that a 20-year-old smoker pays more than a 40-year-old non-smoker.

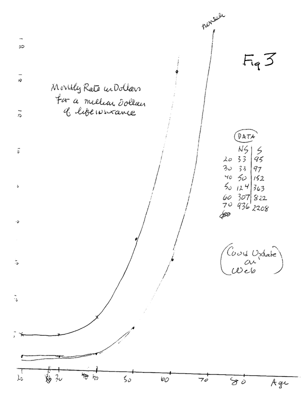


Figure C.1.3: Source: American General Life Insurance Company advertisement

Traffic and Accidents

Figure C.1.4 appears in S.K. Stein's, *Risk Factors of Sober and Drunk Drivers by Time of Day*, Alcohol, Drugs, and Driving 5 (1989), pp. 215–227. The vertical scale is described in the paper.

Glancing at the graph labeled "traffic" we see that there are peaks at the morning and afternoon rush hours, with minimum traffic around 3A.M.. However, the number of accidents is fairly high at that hour. "Risk" is measured by the quotient, "accidents divided by traffic." This reaches a peak at 1a.m.. The high risk cannot be explained by the darkness at that hour, for the risk rapidly decreases the rest of the night. It turns out that the risk has the same shape as the graph that records the number of drunk drivers.

It is a sobering thought that at any time of day a drunk's risk of being involved in an accident is on the order of one hundred times that of an alcohol-free driver at any time of day.

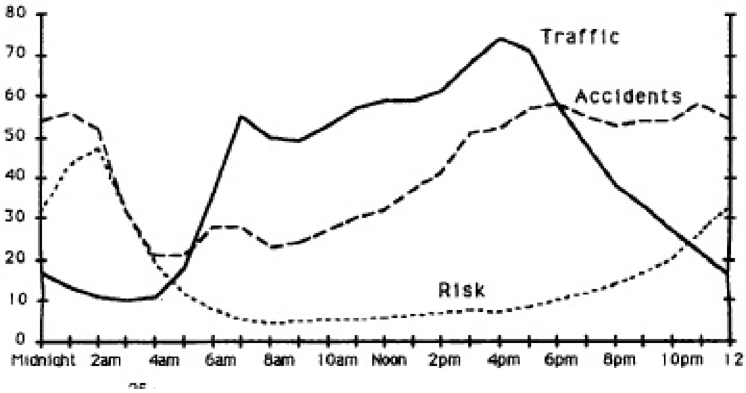


Figure C.1.4:

Petroleum

The three graphs in Figure C.1.5 show the rate of crude oil production in the United States, the rate at which it was imported, and their sum, the rate of consumption. They are expressed in millions of barrels per day, as a function of time. A barrel contains 42 gallons. (For a few years after the discovery of oil in Pennsylvania in 1859 oil was transported in barrels.)

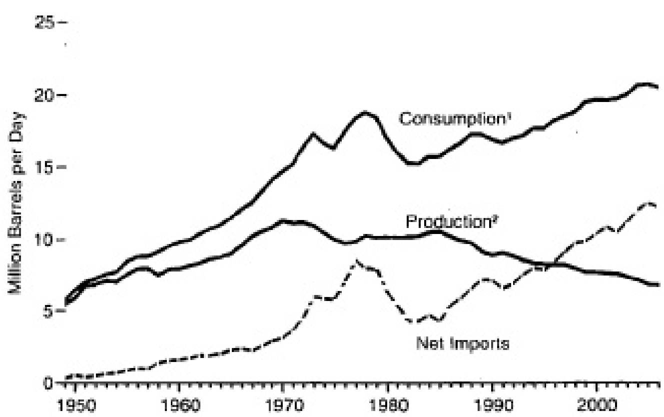


Figure C.1.5: Source: Energy Information Administration (Annual Energy Review, 2006)

The graphs convey a good deal of history and a warning. In 1950 the United States produced almost enough petroleum to meet its needs, but by 2006 it had to import most of the petroleum consumed. Moreover, domestic production peaked in 1970.

The imbalance between production and consumption raises serious questions, especially as exporting countries need more oil to fuel their own growing economies, and developing nations, such as India and China, place rapidly increasing demands on world production. Also, since the total amount of petroleum in the earth is finite, it will run out, and the Age of Oil will end. Geologists, having gone over the globe with a "fine-tooth comb," believe they have already found all the major oil deposits. No wonder that the development of alternative sources of energy has become a high priority.

Calculus is Everywhere # 2

Where Does All That Money Come From?

As of 2007 there were over 7 trillion dollars. Some were in the form of currency, some as deposits in banks, some in money market mutual funds, and so on. Where did they all come from? How is money created?

Banks create some of the supply, and this is how they do it.

When someone makes a deposit at a bank, the bank lends most of it. However, it cannot lend all of it, for it must keep a reserve to meet the needs of depositors who may withdraw money from their account. The government stipulates what this reserve must be, usually between 10 and 20 percent of the deposit Let's use the figure of 20 percent.

If a person deposits \$1,000, the bank can lend \$800. Assume that the borrower deposits that amount in another bank; that second bank can lend 80 percent of the \$800, or \$640. The recipient of the \$640 can then deposit it at a bank, which must retain 20 percent, but is free to lend 80 percent, which is \$512. At this point there are now

$$1000 + 800 + 640 + 512 \text{ dollars in circulation.} \tag{C.2.1}$$

Each summand is 0.8 times the preceding summand. The sum (C.14.42) can be written as

$$1000\left(1 + 0.8 + 0.8^2 + 0.8^3\right). \tag{C.2.2}$$

The process goes on indefinitely, through a fifth person, a sixth, and so on. A good approximation of the impact of that initial deposit of \$1000 after n stages is 1000 times the sum

$$1 + 0.8 + 0.8^2 + 0.8^3 + \cdots + 0.8^n. \tag{C.2.3}$$

A picture shows what happens to such sums as n increases.

The sum (C.14.44), being the sum of a geometric progression, equals $(1 - 0.8^{n+1})/(1 - 0.8)$. As n increases, this approaches $1/0.2 = 5$. In short, the original \$1000 could create an amount approaching \$5000. Economists say that in this case the multiplier is 5, the total impact is five times the initial deposit. There are now magically \$4000 more dollars than at the start. This can happen because a bank can lend money it does not have. The sequence of "deposits and lends" all involved having faith in the future. If that faith is destroyed, then there may be a run on the bank as depositors rush to take their money out. If such a disaster can be avoided, then banking is a delightful business, for bankers can lend money they don't have.

SHERMAN: Update after the financial crisis in 2008-2009?

The concept of the multiplier also appears in measuring economic activity. Assume that the government spends a million dollars on a new road. That amount goes to various firms and individuals who build the road. In turn, those firms and individuals spend a certain fraction. This process of "earn and spend" continues to trickle through the economy. The total impact may be much more than the initial amount that the government spent. Again, the ratio between the total impact and the intial expenditure is called the **multiplier**.

The mathematics behind the multiplier is the theory of the geometric series, summing the successive powers of a fixed number.

EXERCISES

1.[R] If the amount a bank must keep on reserve is cut in half, what effect does this have on the multiplier?

Calculus is Everywhere # 3
Bank Interest and the Annual Percentage Yield

The Truth in Savings Act, passed in 1991, requires a bank to post the Annual Percentage Yield (APY) on deposits. That yield depends on how often the bank computes the interest earned, perhaps as often as daily or as seldom as once a year. Imagine that you open an account on January 1 by depositing \$1000. The bank pays interest monthly at the rate of 5 percent a year. How much will there be in your account at the end of the year? For simplicity, assume all the months have the same length. To begin, we find out how much there is in the account at the end of the first month. The account then has the initial amount, \$1000, plus the interest earned during January. Because there are 12 months, the interest rate in each month is 5 percent divided by 12, which is 0.05/12 percent per month. So the interest earned in January is \$1000 times 0.05/12. At the end of January the account then has

$$\$1000 + \$1000(0.05/12) = \$1000(1 + 0.05/12).$$

The initial deposit is "magnified" by the factor $(1 + 0.05/12)$.

The amount in the account at the end of February is found the same way, but the initial amount is $\$1000(1 + 0.05/12)$ instead of \$1000. Again the amount is magnified by the factor $1 + 0.05/12$, to become

$$\$1000(1 + 0.05/12)^2.$$

The amount at the end of March is

$$\$1000(1 + 0.05/12)^3.$$

At the end of the year the account has grown to

$$\$1000(1 + 0.05/12)^{12},$$

which is about \$1051.16.

The deposit earned \$51.16. If instead the bank computed the interest only once, at the end of the year, so-called "simple interest," the deposit would earn only 5 percent of \$1000, which is \$50. The depositor benefits when the interest is computed more than once a year, so-called "compound interest." A competing bank may offer to compute the interest every day. In that case, the account would grow to

$$\$1000(1 + 0.05/365)^{365},$$

n	$(1 + 1/n)^n$	$(1 + 1/n)^n$
1	$(1 + 1/1)^1$	2.00000
2	$(1 + 1/2)^2$	2.25000
3	$(1 + 1/3)^3$	2.37037
10	$(1 + 1/10)^{10}$	2.59374
100	$(1 + 1/100)^{100}$	2.70481
1000	$(1 + 1/1000)^{1000}$	2.71692

Table C.3.1:

which is about \$1051.27, eleven cents more than the first bank offers. More generally, if the initial deposit is A, the annual interest rate is r, and interest is computed n times a year, the amount at the end of the year is

$$A(1 + r/n)^n. \tag{C.3.4}$$

In the examples, A is \$1000, r is 0.05, and n is 12 and then 365. Of special interest is the case when A is 1 and r is a generous 100 percent, that is, $r = 1$. Then (C.2.1) becomes

$$(1 + 1/n)^n. \tag{C.3.5}$$

How does (C.2.2) behave as n increase? Table C.2.1 shows a few values of (C.2.2), to five decimal places. The base, $1+1/n$, approaches 1 as n increases, suggesting that (C.2.2) may approach a number near 1. However, the exponent gets large, so we are multiplying lots of numbers, all of them larger than 1. It turns out that as n increases $(1 + 1/n)^n$ approaches the number e defined in Section **??**. One can write

$$\lim_{x\to 0^+} (1 + x)^{1/x} = e.$$

Note that the exponent, $1/x$, is the reciprocal of the "small number" x.

With that fact at our disposal, we can figure out what happens when an account opens with \$1000, the annual interest rate is 5 percent, and the interest is compounded more and more often. In that case we would be interested in

$$1000 \lim_{n\to\infty} \left(1 + \frac{0.05}{n}\right)^n.$$

Unfortunately, the exponent n is not the reciprocal of the small number $0.05/n$. But a little algebra can overcome that nuisance, for

$$\left(1 + \frac{0.05}{n}\right)^n = \left(\left(1 + \frac{0.05}{n}\right)^{\frac{n}{0.05}}\right)^{0.05}. \tag{C.3.6}$$

The expression in parentheses has the form "(1 + small number) raised to the reciprocal of that small number." Therefore, as n increases, (C.2.3)) approaches $e^{0.05}$, which is about 1.05127. No matter how often interest is compounded, the \$1000 would never grow beyond \$1051.27.

The definition of e given in Section **??** has no obvious connection to the fact that $\lim_{x \to 0^+}(1+x)^{1/x}$ equals the number e. It seems "obvious," by thinking in terms of banks, that as n increases, so does $(1+1/n)^n$. Without thinking about banks, try showing that it does increase. (This limit will be evaluated in Section **??**.)

EXERCISES

1.[R] A dollar is deposited at the beginning of the year in an account that pays an interest rate r of 100% a year. Let $f(t)$, for $0 \le t \le 1$, be the amount in the account at time t. Graph the function if the bank pays

(a) only simple interest, computed only at $t = 1$.

(b) compound interest, twice a year computed at $t = 1/2$ and 1.

(c) compound interest, three times a year computed at $t = 1/3$, $2/3$, and 1.

(d) compound interest, four times a year computed at $t = 1/4$, $1/2$, $3/4$, and 1.

(e) Are the functions in (a), (b), (c), and (d) continuous?

(f) One could expect the account that is compounded more often than another would always have more in it. Is that the case?

Calculus is Everywhere # 4
Solar Cookers

A satellite dish is parabolic in shape. It is formed by rotating a parabola about its axis The reason is that all radio waves parallel to the axis of the parabola, after bouncing off the parabola, pass through a common point. This point is called the **focus** of the parabola. (See Figure C.3.6.) Similarly, the reflector behind a flashlight bulb is parabolic.

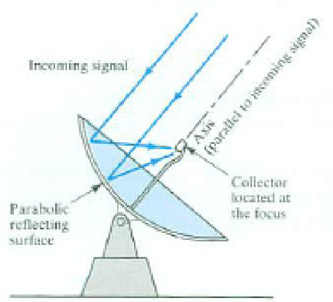


Figure C.4.6:

An ellipse also has a reflection property. Light, or sound, or heat radiating off one focus, after bouncing off the ellipse, goes through the other focus. This is applied, for instance, in the construction of computer chips where it is necessary to bake a photomask onto the surface of a silicon wafer. The heat is focused at the mask by placing a heat source at one focus of an ellipse and positioning the wafer at the other focus, as in Figure C.3.7.

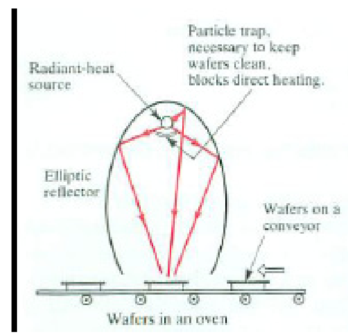


Figure C.4.7:

The reflection property is used in wind tunnel tests of aircraft noise. The test is run in an elliptical chamber, with the aircraft model at one focus and a microphone at the other.

Whispering rooms, such as the rotunda in the Capitol in Washington, D.C., are based on the same principle. A person talking quietly at one focus can be heard easily at the other focus and not at other points between the foci. (The whisper would be unintelligible except for the additional property that all the paths of the sound from one focus to the other have the same length.)

An ellipsoidal reflector cup is used for crushing kidney stones. (An ellipsoid is formed by rotating an ellipse about the line through its foci.) An electrode is placed at one focus and an ellipsoid positioned so that the stone is at the other focus. Shock waves generated at the electrode bounce off the ellipsoid, concentrate on the other focus, and pulverize the stones without damaging other parts of the body. The patient recovers in three to four days instead of the two to three weeks required after surgery. This advance also reduced the mortality rate from kidney stones from 1 in 50 to 1 in 10,000.

The reflecting property of the ellipse also is used in the study of air pollution. One way to detect air pollution is by light scattering. A laser is aimed through one focus of a shiny ellipsoid. When a particle passes through this focus, the light is reflected to the other focus where a light detector is located. The number of particles detected is used to determine the amount of pollution in the air.

The Angle Between Two Lines

To establish the reflection properies just mentioned we will use the principle that the angle of reflection equals the angle of incidence, as in Figure C.3.8, and work with the angle between two lines, given their slopes.

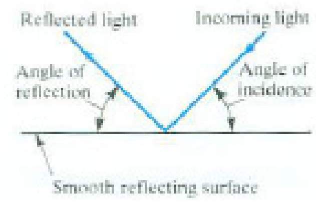
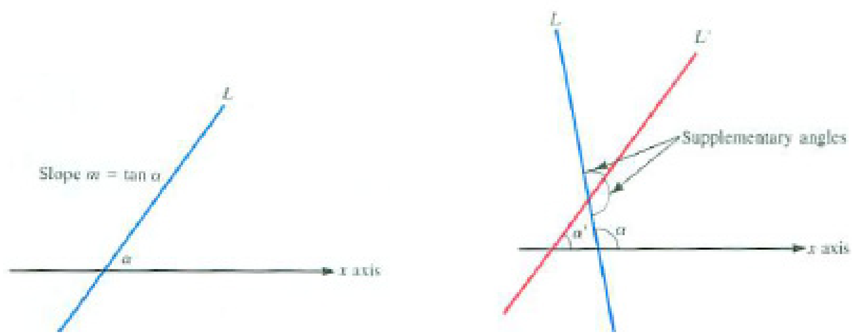


Figure C.4.8:

Consider a line L in the xy-plane. It forms an **angle of inclination** α, $0 \leq \alpha < \pi$, with the positive x-axis. The slope of L is $\tan(\alpha)$. (See Figure C.3.9(a).) If $\alpha = \pi/2$, the slope is not defined.



(a) (b)

Figure C.4.9:

Consider two lines L and L' with angles of inclination α and α' and slopes m and m', respectively, as in Figure C.3.9(b). There are two (supplementary) angles between the two lines. The following definition serves distinguishes one of these two angles as *the* angle between L and L'.

DEFINITION *(Angle between two lines.)* Let L and L' be two lines in the xy-plane, named so that L has the larger angle of inclination, $\alpha > \alpha'$. The angle θ between L and L' is defined to be

$$\theta = \alpha - \alpha'.$$

If L and L' are parallel, define θ to be 0.

Note that θ is the counterclockwise angle from L' to L and that $0 \leq \theta < \pi$. The tangent of θ is easily expressed in terms of the slopes m of L and m' of L'. We have

$$\begin{aligned}
\tan(\theta) &= \tan(\alpha - \alpha') && \text{definition of } \theta \\
&= \frac{\tan(\alpha)-\tan(\alpha')}{1+\tan(\alpha)\tan(\alpha')} && \text{by the identity for } \tan(A-B) \\
&= \frac{m-m'}{1+mm'}.
\end{aligned}$$

Thus

$$\tan(\theta) = \frac{m - m'}{1 + mm'}. \tag{C.4.7}$$

The Reflection Property of a Parabola

Consider the parabola $y = x^2$. (The geometric description of this parabola is the set of all points whose distance from the point $(0, \frac{1}{4})$ equals its distance from the line $y = \frac{-1}{4}$, but this information is not needed here.)

In Figure C.3.10 we wish to show that angles A and B at the typical point (a, a^2) on the parabola are equal. We will do this by showing that $\tan(A) = \tan(B)$.

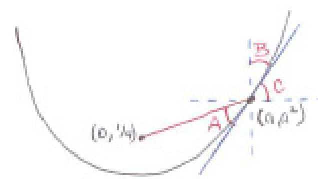


Figure C.4.10:

First of all, $\tan(C) = 2a$, the slope of the parabola at (a, a^2). Since A is the complement of C, $\tan(A) = 1/(2a)$.

The slope of the line through the focus $(0, \frac{1}{4})$ and a point on the parabola (a, a^2) is

$$\frac{a^2 - \frac{1}{4}}{a - 0} = \frac{4a^2 - 1}{4a}.$$

Therefore,

$$\tan(B) = \frac{2a - \frac{4a^2-1}{4a}}{1 + 2a\left(\frac{4a^2-1}{4a}\right)}.$$

Exercise 1 asks you to supply the algebraic steps to complete the proof that $\tan(B) = \tan(A)$.

The Reflection Property of an Ellipse

An ellipse consists of every point such that the sum of the distances from the point to two fixed points is constant. Let the two fixed points, called the **foci** of the ellipse, be a distance $2c$ apart, and the fixed sum of the distances be $2a$, where $a > c$. If the foci are at $(c, 0)$ and $(-c, 0)$ and $b^2 = a^2 - c^2$, the equation of the ellipse is

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

where $b^2 = a^2 - c^2$. (See Figure C.3.11.)

As in the case of the parabola, one shows $\tan(A) = \tan(B)$.

One reason to do Exercise 2 is to appreciate more fully the power of vector calculus, developed later in Chapter **??**, for with that tool you can establish the reflection property of either the parabola or the ellipse in one line.

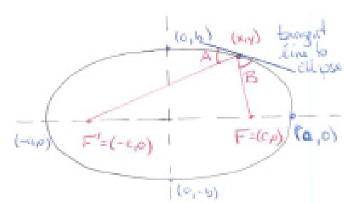


Figure C.4.11: Diocles, *On Burning Mirrors*, edited by G. J. Toomer, Springer, New York, 1976.

Diocles, in his book *On Burning Mirrors*, written around 190 B.C., studied spherical and parabolic reflectors, both of which had been considered by earlier scientists. Some had thought that a spherical reflector focuses incoming light at a single point. This is false, and Diocles showed that a spherical reflector subtending an angle of 60° reflects light that is parallel to its axis of symmetry to points on this axis that occupy about one-thirteenth of the radius. He proposed an experiment, "Perhaps you would like to make two examples of a burning-mirror, one spherical, one parabolic, so that you can measure the burning power of each." Though the reflection property of a parabola was already known, *On Burning Mirrors* contains the first known proof of this property.

Exercise 3 shows that a spherical oven is fairly effective. After all, a potato or hamburger is not a point.

EXERCISES

1.[R] Do the algebra to complete the proof that $\tan(A) = \tan(B)$.

2.[R] This exercise establishes the reflection property of an ellipse. Refer to Figure C.3.11 for a description of the notation.

(a) Find the slope of the tangent line at (x, y).

(b) Find the slope of the line through $F = (c, 0)$ and (x, y).

(c) Find $\tan(B)$.

(d) Find the slope of the line through $F' = (c', 0)$ and (x, y).

(e) Find $\tan(A)$.

(f) Check that $\tan(A) = \tan(B)$.

3.[M] Use trigonometry to show that a spherical mirror of radius r and subtending an angle of 60° causes light parallel to its axis of symmetry to reflect and meet the axis in an interval of length $\left(\frac{1}{\sqrt{3}} - \frac{1}{2}\right) r \approx r/12.9$.

Calculus is Everywhere # 5
Calculus Reassures a Bicyclist

Both authors enjoy bicycling for pleasure and running errands in our flat towns. One of the authors (SS) often bicycles to campus through a parking lot. On each side of his route is a row of parked cars. At any moment a car can back into his path. Wanting to avoid a collision, he wonders where he should ride. The farther he rides from a row, the safer he is. However, the farther he rides from one row, the closer he is to the other row. Where should he ride?

Instinct tells him to ride midway between the two rows, an equal distance from both. But he has second thoughts. Maybe it's better to ride, say, one-third of the way from one row to the other, which is the same as two-thirds of the way from the other row. That would mean he has two safest routes, depending on which row he is nearer. Wanting a definite answer, he resorted to calculus.

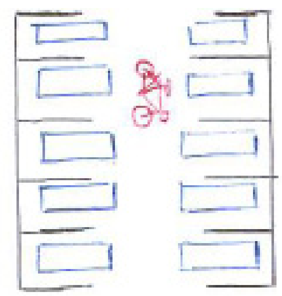


Figure C.5.12: ARTIST:picture of two rows of parked cars, with bicycle

He introduced a function, $f(x)$, which is the probability that he gets through safely when his distance from one row is x, considering only cars in that row. Then he calls the distance between the two rows be d. When he was at a distance x from one row, he was at a distance $d - x$ from the other row. The probability that he did not collide with a car backing out from either row is then the product, $f(x)f(d - x)$. His intuition says that this is maximized when $x = d/2$, putting him midway between the two rows.

What did he know about f? First of all, the farther he rode from one line of cars, the safer he is. So f is an increasing function; thus f' is positive. Moreover, when he was very far from the cars, the probability of riding safely through the lot approached 1. So he assumed $\lim_{x\to\infty} f(x) = 1$ (which it turned out he did not need).

The derivative of f' measured the rate at which he gained safety as he increased his distance from the cars. When x is small, and he rode near the cars, $f'(x)$ was large: he gained a great deal of safety by increasing x. However, when he was far from the cars, he gained very little. That means that f' was a decreasing function. In other words f'' is negative.

Does that information about f imply that midway is the safest route?

In other words, does the maximum of $f(x)f(d - x)$ occur when $x = d/2$? Symbolically, is

$$f(d/2)f(d/2) \geq f(x)f(d - x)?$$

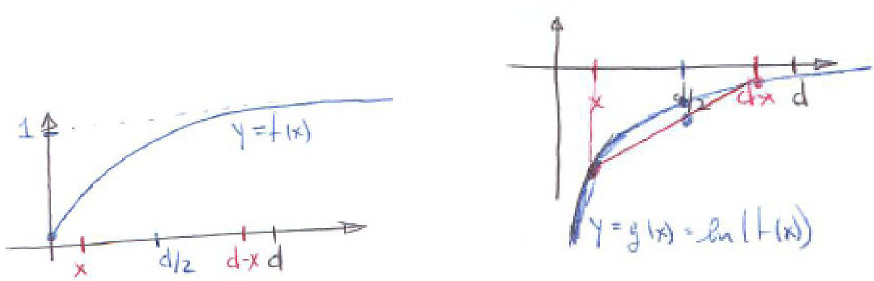
To begin, he took the logarithm of that expression, in order to replace a product by something easier, a sum. He wanted to see if

$$2\ln(f(d/2)) \geq \ln(f(x)) + \ln(f(d - x)).$$

Letting $g(x)$ denote the composite function $\ln(f(x))$, he faced the inequality,

$$\begin{aligned} 2g(d/2) &\geq g(x) + g(d-x), \\ \text{or} \qquad g(d/2) &\geq \frac{1}{2}(g(x) + g(d-x)). \end{aligned}$$

This inequality asserts that the point $(d/2, g(d/2))$ on the graph of g is at least



(a) (b)

Figure C.5.13:

as high as the midpoint of the chord joining $(x, g(x))$ to $(d-x, g(d-x))$. This would be the case if the second derivative of g were negative, and the graph of g were concave down. He had to compute g'' and hope it is negative. First of all, $g'(x)$ is $f'(x)/f(x)$. Then $g''(x)$ is

$$\frac{f(x)f''(x) - (f'(x))^2}{f(x)^2}.$$

The denominator is positive. Because $f(x)$ is positive and concave down, the numerator is negative. So the quotient is negative. That means that the safest path is midway between the two rows. The bicyclist continues to follow that route, but, after these calculations, with more confidence that it is indeed the safest way.

Calculus is Everywhere # 6

Graphs in Economics

Elementary economics texts are full of graphs. They provide visual images of a variety of concepts, such as production, revenue, cost, supply, and demand. Here we show how economists use graphs to help analyze production as a function of the amount of labor, that is, the number of workers.

Let $P(L)$ be the amount of some product, such as cell phones, produced by a firm employing L workers. Since both workers and wireless network cards come in integer amounts, the graph of $P(L)$ is just a bunch of dots. In practice, these dots suggest a curve, and the economists use that curve in their analysis. So $P(L)$ is viewed as a differentiable function defined for some interval of the form $[0, b]$.

If there are no workers, there is no production, so $P(0) = 0$. When the first few workers are added, production may increase rapidly, but as more are hired, production may still increase, but not as rapidly. Figure C.5.14 is a typical **production curve**. It seems to have an inflection point when the gain from adding more workers begins to decline. The inflection point of $P(L)$ occurs at L_2 in Figure C.5.15.

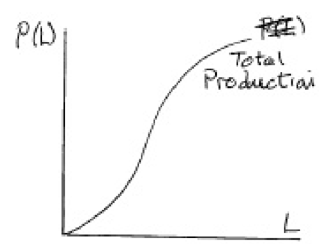


Figure C.6.14:

When the firm employs L workers and adds one more, production increases by $P(L+1) - P(L)$, the marginal production. Economists manage to relate this to the derivative by a simple trick:

$$P(L+1) - P(L) = \frac{P(L+1) - P(L)}{(L+1) - L} \qquad \text{(C.6.8)}$$

The right-hand side of (C.5.5) is "change in output" divided by "change in input," which is, by the definition of the derivative, an approximation to the derivative, $P'(L)$. For this reason economists define the **marginal production** as $P'(L)$, and think of it as the extra product produced by the "L plus first" worker. We denote the marginal product as $m(L)$, that is, $m(L) = P'(L)$.

The **average production** per worker when there are L workers is defined as the quotient $P(L)/L$, which we denote $a(L)$. We have three functions: $P(L)$, $m(L) = P'(L)$, and $a(L) = P(L)/L$.

Now the fun begins.

At what point on the graph of the production function is the average production a maximum?

Since $a(L) = P(L)/L$, it is the slope of the line from the origin to the point $(L, P(L))$ on the graph. Therefore we are looking for the point on the graph where the slope is a maximum. One way to find that point is to rotate a straightedge around the origin, clockwise, starting at the vertical axis until it meets the graph, as in Figure C.5.15. Call the point of tangency $(L_1, P(L_1))$. For L less than L_1 or greater than L_1, average productivity is less than $a(L_1)$.

Note that at L_1 the average product is the same as the marginal product, for the slope of the tangent at $(L_1, P(L_1))$ is both the quotient $P(L_1)/L_1$ and the derivative $P'(L_1)$. We can use calculus to obtain the same conclusion:

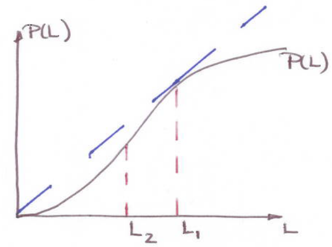


Figure C.6.15:

Since $a(L)$ has a maximum when the input is L_1, its derivative is 0 then. The derivative of $a(L)$ is

$$\frac{d}{dL}\left(\frac{P(L)}{L}\right) = \frac{LP'(L) - P(L)}{L^2}. \tag{C.6.9}$$

At L_1 the quotient in (C.5.6) is 0. Therefore, its numerator is 0, from which it follows that $P'(L_1) = P(L_1)/L_1$. (You might take a few minutes to see why this equation should hold, without using graphs or calculus.)

In any case, the graphs of $m(L)$ and $a(L)$ cross when L is L_1. For smaller values of L, the graph of $m(L)$ is above that of $a(L)$, and for larger values it is below, as shown in Figure C.5.16.

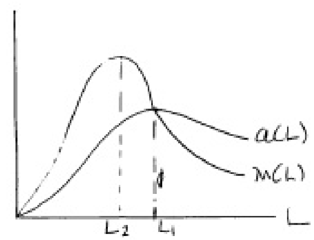


Figure C.6.16:

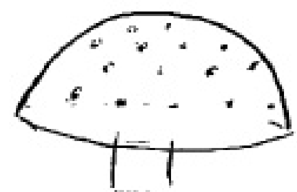
What does the maximum point on the marginal product graph tell about the production graph?

Assume that $m(L)$ has a maximum at L_2. For smaller L than L_2 the derivative of $m(L)$ is positive. For L larger than L_2 the derivative of $m(L)$ is negative. Since $m(L)$ is defined as $P'(L)$, the second derivative of $P(L)$ switches from positive to negative at L_2, showing that the production curve has an inflection point at $(L_2, P(L_2))$.

Economists use similar techniques to deal with a variety of concepts, such as marginal and average cost or marginal and average revenue, viewed as functions of labor or of capital.

Calculus is Everywhere # 7
The Uniform Sprinkler

One day one of the authors (S.S.) realized that the sprinkler did not water his lawn evenly. Placing empty cans throughout the lawn, he discovered that some places received as much as nine times as much water as other places. That meant some parts of the lawn were getting too much water or not enough water.



The sprinkler, which had no moving parts, consisted of a hemisphere, with holes distributed uniformly on its surface, as in Figure C.6.17. Even though the holes were uniformly spaced, the water was not supplied uniformly to the lawn. Why not?

A little calculus answered that question and advised how the holes should be placed to have an equitable distribution. For convenience, it was assumed that the radius of the spherical head was 1, that the speed of the water as it left the head was the same at any hole, and air resistance was disregarded.

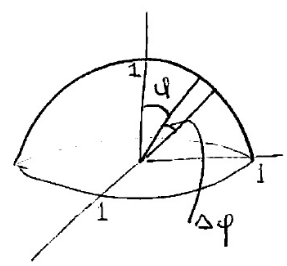


Figure C.7.18:

Consider the water contributed to the lawn by the uniformly spaced holes in a narrow band of width $d\phi$ near the angle ϕ, as shown in Figure C.6.18. To be sure the jet was not blocked by the grass, the angle ϕ is assumed to be no more than $\pi/4$.

Water from this band wets the ring shown in Figure C.6.19.

The area of the band on the sprinkler is roughly $2\pi\sin(\phi)\ d\phi$. As shown in Section **??**, see Exercises **??** and **??**, water from this band lands at a distance from the sprinkler of about

$$x = kv^2\sin(2\phi).$$

Here k is a constant and v is the speed of the water as it leaves the sprinkler. The width of the corresponding ring on the lawn is roughly

$$dx = 2kv^2\cos(2\phi)d\phi.$$

Since its radius is approximately $kv^2\sin(2\phi)$, its area is approximately

$$2\pi\left(kv^2\sin(2\phi)\right)\left(2kv^2\cos(2\phi)\ d\phi\right),$$

which is proportional to $\sin(2\phi)\cos(2\phi)$, hence to $\sin(4\phi)$.

Thus the water supplied by the band was proportional to $\sin(\phi)$ but the area watered by that band was proportional to $\sin(4\phi)$. The ratio

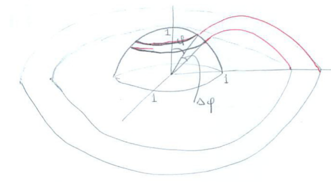


Figure C.7.19:

$$\frac{\sin(4\phi)}{\sin(\phi)} = \frac{\text{Area watered on lawn}}{\text{Area of supply on sprinkler}}$$

is the key to understanding both why the distribution was not uniform and to finding out how the holes should be placed to water the lawn uniformly.

By l'Hôpital's rule, this fraction approaches 4 as ϕ approaches zero:

$$\lim_{\phi\to 0}\frac{\sin(4\phi)}{\sin(\phi)} = 4. \qquad \text{(C.7.10)}$$

This means that for angles ϕ near 0 that ratio is near 4. When ϕ is $\pi/4$, that ratio is $\frac{\sin(\pi)}{\sin(\pi/4)} = 0$, and water was supplied much more heavily far from the sprinkler than near it. To compensate for this bias the number of holes in the band should be proportional to $\sin(4\phi)/\sin(\phi)$. Then the amount of water is proportional to the area watered, and watering is therefore uniform.

Professor Anthony Wexler of the Mechanical Engineering Department of UC-Davis calculated where to drill the holes and made a prototype, which produced a beautiful fountain and a much more even supply of water. Moreover, if some of the holes were removed, it would water a rectangular lawn.

We offered the idea to the firm that made the biased sprinkler. After keeping the prototype for half a year, it turned it down because "it would compete with the product we have."

Perhaps, when water becomes more expensive our uniform sprinkler may eventually water many a lawn.

EXERCISES

1.[R] Show that the limit (C.6.7) is 4

(a) using only trigonometric identities.

(b) using l'Hôpital's rule.

2.[R] Show that $\sin(4x)/\sin(x)$ is a decreasing function for x in the interval $[0, \pi/4]$. HINT: Use trigonometric identities and no calculus. (However, you may be amused if you also do this by calculus.)

3.[R] An oscillating sprinkler goes back and forth at a fixed angular speed.

(a) Does it water a lawn uniformly?

(b) If not, how would you modify it to provide more uniform coverage?

Calculus is Everywhere # 8
Peak Oil Production

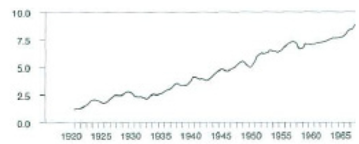


Figure C.8.20:

The United States in 1956 produced most of the oil it consumed, and the rate of production was increasing. Even so, M. King Hubbert, a geologist at Shell Oil, predicted that production would peak near 1970 and then gradually decline. His prediction did not convince geologists, who were reassured by the rising curve in Figure C.7.20.

Hubbert was right and the moment of maximum production is known today as Hubbert's Peak.

We present below Hubbert's reasoning in his own words, drawn from "Nuclear Energy and the Fossil Fuels," available at `http://www.hubbertpeak.com/hubbert/1956/1956.pdf`. In it he uses an integral over the entire positive x-axis, a concept we will define in Section **??**. However, since a finite resource is exhausted in a finite time, his integral is an ordinary definite integral, whose upper bound is not known.

First he stated two principles when analyzing curves that describe the rate of exploitation of a finite resource:

1. For any production curve of a finite resource of fixed amount, two points on the curve are known at the outset, namely that at $t = 0$ and again at $t = \infty$. The production rate will be zero when the reference time is zero, and the rate will again be zero when the resource is exhausted; that is to say, in the production of any resource of fixed magnitude, the production rate must begin at zero, and then after passing through one or several maxima, it must decline again to zero.

2. The second consideration arises from the fundamental theorem of integral calculus; namely, if there exists a single-valued function $y = f(x)$, then

$$\int_0^{x_1} y\, dx = A, \tag{C.8.11}$$

where A is the area between the curve $y = f(x)$ and the x-axis from the origin out to the distance x_1.

In the case of the production curve plotted against time on an arithmetical scale, we have as the ordinate

$$P = \frac{dQ}{dt}, \tag{C.8.12}$$

where dQ is the quantity of the resource produced in time dt. Likewise, from equation (C.7.8) the area under the curve up to any time t is given by

$$A = \int_0^t P\ dt = \int_0^t \left(\frac{dQ}{dt}\right)\ dt = Q, \tag{C.8.13}$$

where Q is the cumulative production up to the time t. Likewise, the ultimate production will be given by

$$Q_{max} = \int_0^\infty P\ dt, \tag{C.8.14}$$

and will be represented on the graph of production-versus-time as the total area beneath the curve.

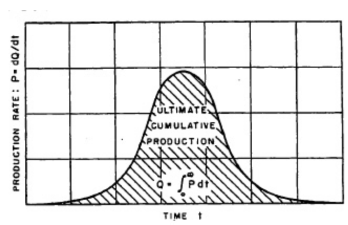


Figure C.8.21:

These basic relationships are indicated in Figure C.7.21. The only a priori information concerning the magnitude of the ultimate cumulative production of which we may be certain is that it will be less than, or at most equal to, the quantity of the resource initially present. Consequently, if we knew the production curves, all of which would exhibit the common property of beginning and ending at zero, and encompassing an area equal to or less than the initial quantity.

That the production of exhaustible resources does behave this way can be seen by examining the production curves of some of the older producing areas.

He then examines those curves for Ohio and Illinois. They resembled the curves below, which describe more recent data on production in Alaska, the United States, the North Sea, and Mexico.

Hubbert did not use a particular formula. Instead he employed the key idea in calculus, expressed in terms of production of oil, "The definite integral of the rate of production equals the total production."

He looked at the data up to 1956 and extrapolated the curve by eye, and by logic. This is his reasoning:

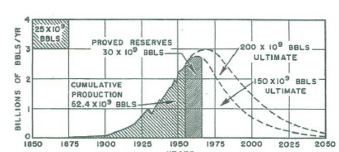


Figure C.8.23: Ultimate United States crude-oil production based on assumed initial reserves of 150 and 200 billion barrels.

> Figure C.7.23 shows "a graph of the production up to the present, and two extrapolations into the future. The unit rectangle in this case represents 25 billion barrels so that if the ultimate potential production is 150 billion barrels, then the graph can encompass but six rectangles before returning to zero. Since the cumulative production is already a little more than 50 billion barrels, then only four more rectangles are available for future production. Also, since the production rate is still increasing, the ultimate production peak must be greater than the present rate of production and must occur sometime in the future. At the same

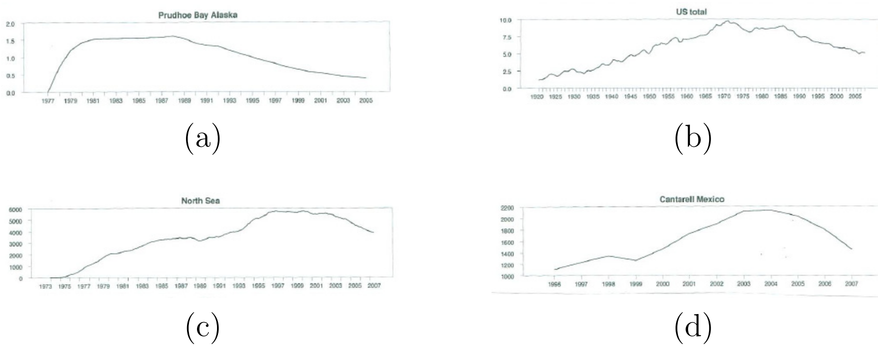


Figure C.8.22: Annual production of oil in millions of barrels per day for (a) Annual oil production for Prudhoe Bay in Alaska, 1977–2005 [Alaska Department of Revenue], (b) moving average of preceding 12 months of monthly oil production for the United States, 1920–2008 [EIA, "Crude Oil Production"], (c) moving average of preceding 12 months of sum of U.K. and Norway crude oil production, 1973–2007 [EIA, Table 11.1b], and (d) annual production from Cantarell complex in Mexico, 1996–2007 [Pemex 2007 Statistical Yearbook and Green Car Congress (`http://www.greencarcongress.com/2008/01/mexicos-cantare.html`).

> time it is possible to delay the peak for more than a few years and still allow time for the unavoidable prolonged period of decline due to the slowing rates of extraction from depleting reservoirs.
>
> With due regard for these considerations, it is almost impossible to draw the production curve based upon an assumed ultimate production of 150 billion barrels in any manner differing significantly from that shown in Figure C.7.23, according to which the curve must culminate in about 1965 and then must decline at a rate comparable to its earlier rate of growth.
>
> If we suppose the figure of 150 billion barrels to be 50 billion barrels too low — an amount equal to eight East Texas oil fields — then the ultimate potential reserve would be 200 billion barrels. The second of the two extrapolations shown in Figure C.7.23 is based upon this assumption; but it is interesting to note that even then the date of culmination is retarded only until about 1970."

Geologists are now trying to predict when world production of oil will peak. (Hubbert predicted the peak to occur in the year 2000.) In 2009 oil was being extracted at the rate of 85 million barrels per day. Some say the peak occurred as early as 2005, but others believe it may not occur until after 2020.

To see some of the latest estimates, do a web search for "Hubbert peak oil estimate".

What is just as alarming is that the world is burning oil faster than we are discovering new deposits.

In the CIE on Hubbert's Peak in Chapter **??** (see page 34) we present a later work of Hubbert, in which he uses a specific formula to analyze oil use and depletion.

Calculus is Everywhere # 9

Escape Velocity

In Example **??** in Section **??** we saw that the total work required to lift a 1-pound payload from the surface of the earth to the moon is 3,933 mile-pounds. Since the radius of the earth is about 4,000 miles, the work required to launch a payload on an endless journey is given by the improper integral

$$\int_{4,000}^{\infty} \left(\frac{4,000}{r}\right)^2 dr = 4,000 \text{ mile-pounds.}$$

Because the integral is convergent, only a finite amount of energy is needed to send a payload on an endless journey — as the Voyager spacecraft has demonstrated. It takes only a little more energy than is required to lift the payload to the moon.

That the work required for the endless journey is finite raises the question "With what initial velocity must we launch the payload so that it never falls back, but continues to rise forever away from the earth?" If the initial velocity is too small, the payload will rise for a while, then fall back, as anyone who has thrown a ball straight up knows quite well.

The energy we supply the payload is kinetic energy. The force of gravity slows the payload and reduces its kinetic energy. We do not want the kinetic energy to shrink to zero. It it were ever zero, then the velocity of the payload would be zero. At that point the payload would start to fall back to earth.

As we will show, the kinetic energy of the payload is reduced by *exactly* the amount of work done on the payload by gravity. If v_{esc} is the minimal velocity needed for the payload to "escape" and not fall back, then

$$\frac{1}{2}mv_{\text{esc}}^2 = 4,000 \text{ mile-pounds,} \tag{C.9.15}$$

where m is the mass of the payload. Equation (C.8.12) can be solved for v_{esc}, the **escape velocity**.

In order to solve (C.8.12) for v_{esc}, we must calculate the mass of a payload that weighs 1 pound at the surface of the earth. The weight of 1 pound is the gravitational force of the earth pulling on the payload. Newton's equation

$$\text{Force} = \text{Mass} \times \text{Acceleration}, \tag{C.9.16}$$

known as his "second law of motion," provides the relationship among force, mass, and the acceleration of that mass that is needed.

The acceleration of an object at the surface of the earth is 32 feet per second per second, or 0.0061 miles per second per second. Then (C.8.13), for the 1-pound payload, becomes

$$1 = m(0.0061). \qquad \text{(C.9.17)}$$

Combining (C.8.12) and (C.8.14) gives

$$\begin{aligned} \frac{1}{2}\frac{1}{0.0061}\left(v_{\text{esc}}\right)^2 &= 4,000 \\ \text{or} \qquad \left(v_{\text{esc}}\right)^2 = (8,000)(0.0061) &= 48.8. \end{aligned}$$

Hence $v_{\text{esc}} \approx 7$ miles per second, which is about 25,000 miles per hour, a speed first attained by human beings when Apollo 8 traveled to the moon in December 1968. All that remains is to justify the claim that the change in kinetic energy equals the work done by the force.

Let $v(r)$ be the velocity of the payload when it is r miles from the center of the earth. Let $F(r)$ be the force on the payload when it is r miles from the center of the earth. Since the force is in the opposite direction from the motion, we will define $F(r)$ to be negative.

Let a and b be numbers, $4,000 \leq a < b$. (See Figure C.8.24.) We wish to show that

$$\underbrace{\frac{1}{2}m(v(b))^2 - \frac{1}{2}m(v(a))^2}_{\text{change in kinetic energy}} = \underbrace{\int_a^b F(r)dr}_{\text{work done by gravity}} . \qquad \text{(C.9.18)}$$

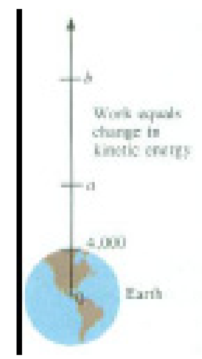


Figure C.9.24:

In this equation m is the payload mass. Note that both sides of (C.8.15) are negative.

Equation (C.8.15) resembles the Fundamental Theorem of Calculus. If we could show that $\frac{1}{2}m(v(r))^2$ is an antiderivative of $F(r)$, then (C.8.15) would follow immediately. Let us find the derivative of $\frac{1}{2}m(v(r))^2$ with respect to r and show that it equals $F(r)$:

$$\begin{aligned} \frac{d}{dr}\left(\frac{1}{2}m(v(r))^2\right) &= mv(r)\frac{dv}{dr} = mv(r)\frac{dv/dt}{dr/dt} && \text{(Chain Rule; } t \text{ is time)} \\ &= mv(r)\frac{d^2r/dt^2}{v(r)} = m\frac{d^2r}{dt^2} && (v(r) = \tfrac{dr}{dt}) \\ &= \text{mass} \times \text{acceleration} \\ &= F(r) && \text{(Newton's 2}^{\text{nd}} \text{ Law of Motion.} \end{aligned}$$

Hence (C.8.15) is valid and we have justified our calculation of escape velocity.

Incidentally, the escape velocity is $\sqrt{2}$ times the velocity required for a satellite to orbit the earth (and not fall into the atmosphere and burn up).

EXERCISES 1.[R] The earth is not a perfect sphere. The"mean radius" of the earth is about 3,959 miles. A more accurate value for the force of gravity is 32.174 feet per second per second. Repeat the derivation of the escape velocity using these values. References: `http://en.wikipedia.org/wiki/Earth_radius` and `http://en.wikipedia.org/wiki/Standard_gravity`.

2.[R] Repeat the derivation of the escape velocity using CGS units. That is, assume the radius of the earth is 6,371 kilometers and the force of gravity is 9.80665 meters per second per second.

3.[R] Determine the escape velocity from the moon. NOTE: What information do you need to complete this calculation?

4.[R] Determine the escape velocity from the sun.

Calculus is Everywhere # 10
Average Speed and Class Size

There are two ways to define your average speed when jogging or driving a car. You could jot down your speed at regular intervals of time, say, every second. Then you would just average those speeds. That average is called an average *with respect to time*. Or, you could jot down your velocity at regular intervals of distance, say, every hundred feet. The average of those velocities is called an average *with respect to distance*.

How do you think they would compare? If you kept a constant speed, c, the averages would both be c. Are they always equal, even if your speed varies? Would one of the averages always tend to be larger? Try to answer the question before we analyze it mathematically, with the aid of the **Cauchy-Schwartz inequality**.

pronounced: "ko-shee' shwartz"

There are several versions of the Cauchy-Schwartz inequality. The version we need here concerns two continuous functions, f and g, defined on an interval $[a, b]$. If $\int_a^b f(x)^2\ dx$ and $\int_a^b g(x)^2\ dx$ are small, then the absolute value of $\int_a^b f(x)g(x)\ dx$ ought to be small too. It is, as the following Cauchy-Schwartz inequality implies:

$$\left(\int_a^b f(x)g(x)\ dx\right)^2 \leq \int_a^b f(x)^2\ dx \int_a^b g(x)^2\ dx. \qquad \text{(C.10.19)}$$

After showing some of its applications, we will use the quadratic formula to show that it is true.

First we use the inequality (C.9.16) to answer the question, "Which average of speed is larger, the one with respect to time or the one with respect to distance?"

Let the speed at time t be $v(t)$ and let $s(t)$ be the distance traveled up to time t. During the time interval from time a to time b the average of velocity with respect to time is

$$\frac{\int_a^b v(t)\ dt}{b-a} = \frac{s(b)-s(a)}{b-a}.$$

On the other hand, the average of velocity with respect to distance is defined as

$$\frac{\int_{s(a)}^{s(b)} v(s)\ ds}{s(b)-s(a)}, \qquad \text{(C.10.20)}$$

where $v(s)$ denotes the velocity when the distance covered is s. Changing the independent variable in the numerator of (C.9.17) from s to t by the relation $ds = v(t)\ dt$, we obtain

$$\frac{\int_{s(a)}^{s(b)} v(s)\ ds}{s(b) - s(a)} = \frac{\int_a^b v(t)v(t)\ dt}{s(b) - s(a)}.$$

Noting that $s(b) - s(a) = \int_a^b v(t)\ dt$ and $b - a = \int_a^b 1\ dt$, we will show that the average with respect to time is less than or equal to the average with respect to distance, that is,

$$\frac{\int_a^b v(t)\ dt}{\int_a^b 1\ dt} \leq \frac{\int_a^b v(t)^2\ dt}{\int_a^b v(t)\ dt}.$$

Or, equivalently,

$$\left(\int_a^b v(t)\ dt \right)^2 \leq \int_a^b 1\ dt \int_a^b v(t)^2\ dt. \qquad \text{(C.10.21)}$$

But, (C.9.18) is a special case of (C.9.16), with $f(t) = 1$ and $g(t) = v(t)$.

This implies that the average with respect to time is always less than or equal to the average with respect to distance. Exercise 1 shows a bit more: if the speed is not constant, then the average with respect to time is less than the average with respect to distance.

The way to show that inequality (C.9.16) holds is indirect but short. Introduce a new function, $h(t)$, defined by

$$h(t) = \int_a^b (f(x) - tg(x))^2\ dx = \int_a^b f(x)^2\ dx - 2t \int_a^b f(x)g(x)\ dx + t^2 \int_a^b g(x)^2\ dx. \qquad \text{(C.10.22)}$$

Because the first integrand in (C.9.19) is never negative, $h(t) \geq 0$. Now, $h(t) = pt^2 + qt + r$, where

$$p = \int_a^b g(x)^2\ dx, \qquad q = -2 \int_a^b f(x)g(x)\ dx, \quad \text{and} \quad r = \int_a^b f(x)^2\ dx.$$

The parabola $y = h(t)$ never drops below the t-axis, and touches the t-axis at at most one point. Otherwise, if it touches the t-axis at two points, it would dip below that axis, forcing $h(t)$ to take on some negative values.

Because the equation $h(t) = 0$ has at most one solution, the discriminant $q^2 - 4pr$ must not be positive. Thus, $q^2 - 4pr \leq 0$, from which the Cauchy-Schwartz inequality follows.

EXERCISES

1.[M] Show that the only case when equality holds in (C.9.16) is when $g(x)$ is a constant times $f(x)$.

2.[M] The "discrete" form of the Cauchy-Schwartz inequality asserts that if a_1, a_2, a_3, …, a_n and b_1, b_2, b_3, …, b_n are real numbers, then

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 \leq \sum_{i=1}^{n} a_i^2 \sum_{i=1}^{n} a_i^2.$$

(a) Prove this inequality.

(b) When does equality hold?

3.[M] Use the inequality in Exercise 2 to show that the average class size at a university as viewed by the registrar is usually smaller than the average class size as viewed by the students.

It is also the case that the average time between buses as viewed by the dispatcher is usually shorter than the average time between buses as viewed by passengers arriving randomly at a bus stop.
Reference: S. K. Stein, *An Inequality Between Two Average Speeds*, Transportation Research 22B (1988), pp. 469–471.
4.[C] A region R is bounded by the x-axis, the lines $x = 2$ and $x = 5$, and the curve $y = f(x)$, where f is a positive function. The area of R is A. When revolved around the x-axis it produces a solid of volume V.

(a) How large can V be?

(b) How small can V be?

HINT: In one of these two cases the Cauchy-Schwartz inequality on 25 may help.

5.[C] If the region R in the preceding exercise is revolved around the y-axis, what can be said about the maximum and minimum values for the volume of the resulting solid? Explain.

Calculus is Everywhere # 11
An Improper Integral in Economics

Both business and government frequently face the question, "How much money do I need today to have one dollar t years in the future?"

Implicit in this question are such considerations as the present value of a business being dependent on its future profit and the cost of a toll road being weighed against its future revenue. We determine the present value of a business which depends on the future rate of profit.

To begin the analysis, assume that the annual interest rate r remains constant and that 1 dollar deposited today is worth e^{rt} dollars t years from now. This assumption corresponds to continuously compounded interest or to natural growth. Thus A dollars today will be worth Ae^{rt} dollars t years from now. What is the present value of the promise of 1 dollar t years from now? In other words, what amount A invested today will be worth 1 dollar t years from now? To find out, solve the equation $Ae^{rt} = 1$ for A. The solution is

t need not be an integer

The present value of \$1 t years from now is \$ e^{-rt}

$$A = e^{-rt}. \tag{C.11.23}$$

Now consider the present value of the future profit of a business (or future revenue of a toll road). Assume that the profit flow t years from now is at the rate $f(t)$. This rate may vary within the year; consider f to be a continuous function of time. The profit in the small interval of time dt, from time t to time $t + dt$, would be approximately $f(t)dt$. The total future profit, $F(T)$, from now, when $t = 0$, to some time T in the future is therefore

$$F(T) = \int_0^T f(t)dt. \tag{C.11.24}$$

But the **present value** of the future profit is *not* given by (C.10.21). It is necessary to consider the present value of the profit earned in a typical short interval of time from t to $t + dt$. According to (C.10.20), its present value is approximately

$$e^{-rt} f(t)dt.$$

Hence the present value of future profit from $t = 0$ to $t = T$ is given by

$$\int_0^T e^{-rt} f(t)dt. \tag{C.11.25}$$

The present value of all future profit is, therefore, the improper integral $\int_0^\infty e^{-rt} f(t)dt$.

To see what influence the interest rate r has, denote by $P(r)$ the present value of all future revenue when the interest rate is r; that is,

$$P(r) = \int_0^\infty e^{-rt} f(t)dt. \tag{C.11.26}$$

If the interest rate r is raised, then according to (C.10.23) the present value of a business declines. An investor choosing between investing in a business or placing the money in a bank account may find the bank account more attractive when r is raised.

A proponent of a project, such as a toll road, will argue that the interest rate r will be low in the future. An opponent will predict that it will be high. Of course, neither knows what the inscrutable future will do to the interest rate. Even so, the prediction is important in a cost-benefit analysis.

Equation (C.10.23) assigns to a profit function f (which is a function of time t) a present-value function P, which is a function of r, the interest rate. In the theory of differential equations, P is called the **Laplace transform** of f. This transform can replace a differential equation by a simpler equation that looks quite different.

The Laplace transform was first encountered in Exercises ?? to ?? in Section ?? and reappeared in Exercises ?? to ?? in Section ??.

EXERCISES

In Exercises 1 to 8 $f(t)$ is defined on $[0, \infty)$ and is continuous. Assume that for $r > 0$, $\int_0^\infty e^{-rt} f(t)dt$ converges and that $e^{-rt} f(t) \to 0$ as $t \to \infty$. Let $P(r) = \int_0^\infty e^{-rt} f(t)dt$. Find $P(r)$, the Laplace transform of $f(t)$, in Exercises 1 to 5.

1.[R] $f(t) = t$

2.[R] $f(t) = e^t$, assume $r > 1$

3.[R] $f(t) = t^2$

4.[R] $f(t) = \sin(t)$

5.[R] $f(t) = \cos(t)$

6.[M] Let P be the Laplace transform of f, and let Q be the Laplace transform of f'. Show that $Q(r) = -f(0) + rP(r)$.

7.[M] Let P be the Laplace transform of f, a a positive constant, and $g(t) = f(at)$. Let Q be the Laplace transform of g. Show that $Q(t) = \frac{1}{a} P\left(\frac{r}{a}\right)$.

8.[R] Which is worth more today, \$100, 8 years from now or \$80, five years from now?

(a) Assume $r = 4\%$.

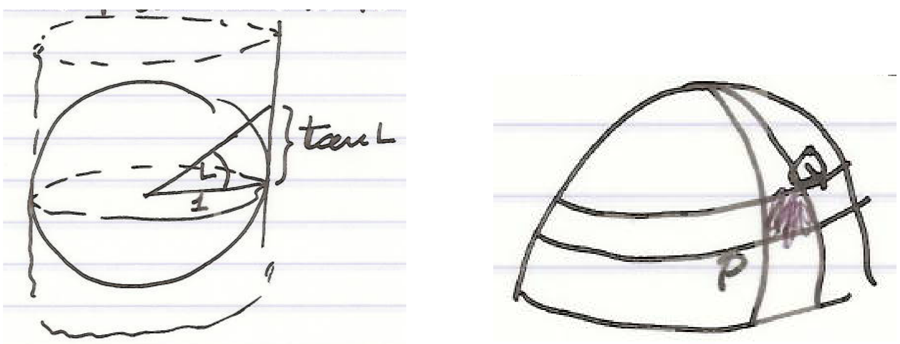
(b) Assume $r = 8\%$.

(c) For which interest rate are the two equal?

Calculus is Everywhere # 12
The Mercator Map

A web search for "map projection" leads to detailed information about this and other projections. The US Geological Society has some particularly good resources.

One way to make a map of a sphere is to wrap a paper cylinder around the sphere and project points on the sphere onto the cylinder by rays from the center of the sphere. This **central cylindrical projection** is illustrated in Figure C.11.25(a).



(a) (b)

Figure C.12.25:

Points at latitude L project onto points at height $\tan(L)$ from the equatorial plane.

A **meridean** is a great circle passing through the north and south poles. It corresponds to a fixed longitude. A short segment on a meridian at latitude L of length dL projects onto the cylinder in a segment of length approximately $d(\tan(L)) = \sec(L)^2\, dL$. This tells us that the map magnifies short vertical segments at latitude L by the factor $\sec^2(L)$.

Points on the sphere at latitude L form a circle of radius $\cos(L)$. The image of this circle on the cylinder is a circle of radius 1. That means the projection magnifies horizontal distances at latitude L by the factor $\sec(L)$.

Consider the effect on a small "almost rectangular" patch bordered by two meridians and two latitude lines. The patch is shaded in Figure C.11.25(b). The vertical edges are magnified by $\sec^2(L)$, but the horizontal edges by only $\sec(L)$. The image on the cylinder will not resemble the patch, for it is stretched more vertically than horizontally.

The path a ship sailing from P to Q makes a certain angle with the latitude line through P. The map just described distorts that angle.

The ship's caption would prefer a map without such a distortion, one that preserves direction. That way, to chart a voyage from point A to point B on the sphere of the Earth at a fixed compass heading, he would simply draw a straight line from A to B on the map to determine the compass setting.

Gerhardus Mercator, in 1569, designed a map that does not distort small patches hence preserves direction. He figured that since the horizontal magnification factor is $\sec(L)$, the vertical magnification should also be $\sec(L)$, not $\sec^2(L)$.

This condition can be stated in the language of calculus. Let y be the height on the map that represents latitude L_0. Then Δy should be approximately $\sec(L)\Delta L$. Taking the limit of $\Delta y/\Delta L$ and ΔL approaches 0, we see that $dy/dL = \sec(L)$. Thus

$$y = \int_0^{L_0} \sec(L)\ dL. \qquad \text{(C.12.27)}$$

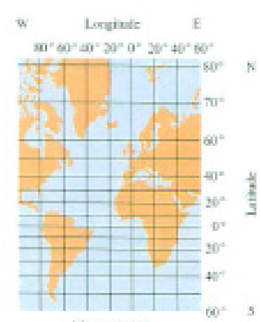
Mercator, working a century before the invention of calculus, did not have the concept of the integral or the Fundamental Theorem of Calculus. Instead, he had to break the interval $[0, L_0]$ into several short sections of length ΔL, compute $(\sec(L))\Delta L$ for each one, and sum these numbers to estimate y in (C.11.24).

We, coming after Newton and Leibniz, can write

$$y = \int_0^{L_0} \sec(L)\, dL = \ln|\sec(L) + \tan(L)|\ |_0^{L_0} = \ln(\sec(L_0)+\tan(L_0)) \qquad \text{for } 0 \le L_0 \le \pi/2.$$

In 1645, Henry Bond conjectured that, on the basis of numerical evidence, $\int_0^{\alpha} \sec(\theta)\ d\theta = \ln(\tan(\alpha/2+\pi/4))$ but offered no proof. In 1666, Nicolaus Mercator (no relation to Gerhardus) offered the royalties on one of his inventions to the mathematician who could prove Bond's conjecture was right. Within two years James Gregory provided the missing proof.

Figure 11 shows a Mercator map. Such a map, though it preserves angles, greatly distorts areas: Greenland looks bigger than South America even though it is only one eighth its size. The first map we described distorts areas even more than does a Mercator map.



Mercator map

EXERCISES

1.[R] Draw a clear diagram showing why segments at latitude L are magnified vertically by the factor $\sec(L)$.

2.[R] Explain why a short arc of length dL in Figure C.11.25(a) projects onto a short interval of length approximately $\sec^2(L)\ dL$.

3.[R] On a Mercator map, what is the ratio between the distance between the lines representing latitudes 60° and 50° to the distance between the lines representing latitudes 40° amd 30°?

4.[M] What magnifying effect does a Mercator map have on areas?

Calculus is Everywhere # 13
Hubbert's Peak

SHERMAN: This CIE needs an ending. Did you have something in mind?

In the CIE for Chapter **??**, Hubbert combined calculus concepts with counting squares. Later he developed specific functions and used more techniques of calculus in "Oil and Gas Supply Modeling", NBS Special Publication 631, U.S. Department of Commerce, National Bureau of Standards, May, 1982. (NOTE: NBS is now the National Institute of Standards and Technology (NIST).)

In his approach, Q_∞ denotes the total amount of oil reserves a the time oil is first extracted and t, time. The derivative dQ/dt is the rate at which oil is extracted. $Q(t)$ denotes the amount extracted up to time t. Hubbert assumes $Q(0) = 0$ and $(dQ/dt)(0) = 0$. He wants to obtain a formula for $Q(t)$.

"The curve of dQ/dt versus Q between 0 and Q_∞ can be represented by the Maclaurin series

$$\frac{dQ}{dt} = c_0 + c_1 Q + c_2 Q^2 + c_3 Q^3 + \cdots .$$

Since, when $Q = 0$, $dQ/dt = 0$, it follows that $c_0 = 0$.

"Since the curve must return to 0 when $Q = Q_\infty$, the minimum number of terms that permit this, and the simplest form of the equation, becomes the second degree equation

$$\frac{dQ}{dt} = c_1 Q + c_2 Q^2.$$

By letting $a = c_1$ and $b = -c_2$, this can be rewritten as

$$\frac{dQ}{dt} = aQ - bQ^2.$$

"Since when $Q = Q_\infty$, $dQ/dt = 0$,

$$aQ_\infty - bQ_\infty^2 = 0$$

or

$$b = \frac{a}{Q_\infty}$$

and

$$\frac{dQ}{dt} = a\left(Q - \frac{Q^2}{Q_\infty}\right). \qquad \text{(C.13.28)}$$

"This is the equation of a parabola .... The maximum value occurs when the slope is 0, or when

$$a - \frac{2a}{Q_\infty}Q = 0,$$

or

$$Q = \frac{Q_\infty}{2}.$$

"It is to be emphasized that the curve of dQ/dt versus Q does not have to be a parabola, but that a parabola is the simplest mathematical form that this curve can assume. We may regard the parabolic form as a sort of idealization for all such actual data curves."

He then points out that

$$\frac{dQ/dt}{Q} = a - \frac{a}{Q_\infty}.$$

"This is the equation of a straight line. The plotting of this straight line gives the values for its constraints Q_∞ and a."

Because the rate of production, dQ/dt, and the total amount produced up to time t, namely, $Q(t)$ and observable, the line can be drawn and its intercepts read off the graph. (The two intercepts are $(0, a)$ and $(Q_\infty, 0)$.)

Hubbert then compares this with actual data, which it approximates fairly well.

Equation (C.12.25) can be written as

$$\frac{dQ}{dt} = \frac{a}{Q_\infty} Q \left(Q_\infty - Q\right),$$

which says, "The rate of production is proportional both to the amount already produced and to the reserves $Q_\infty - Q$." This is related to the logistic equation describing bounded growth. (See Exercises **??** to **??** in Section **??**.)

This approach, which is more formal than the one in CIE 7 at the end of Chapter **??**, concludes that as Q approaches Q_∞, the rate of production will decline, approaching 0. This means the Age of Oil will end.

Calculus is Everywhere # 14

$E = mc^2$

This could also appear as a boxed item in Chapter ??.

The equation $E = mc^2$ relates the energy associated with an object to its mass and the speed of light. Where does it come from?

Newton's second law of motion reads: "Force is the rate at which the momentum of an object changes." The momentum of an object of mass m and velocity v is the product mv. Denoting the force by F, we have

$$F = \frac{d}{dt}(mv).$$

If the mass is constant, this reduces to the familiar "force equals mass times acceleration." But what if the mass m is not constant? What if the mass of an object changes as its velocity changes?

According to Einstein's Special Theory of Relativity, announced in 1905, the mass does change, in a manner described by the equation:

$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{C.14.29}$$

Here m_0 is the mass at rest, v is the velocity, and c is the velocity of light. If v is not zero, m is larger than m_0. When v is small (compared to the velocity of light) then m is only slightly larger than m_0. However, as v approaches the velocity of light, the mass becomes arbitrarily large.

For a satellite circling the Earth at 17,000 miles per hour, v/c is less than 1/2500.

Consider moving an object, initially at rest, in a straight line. If the velocity at time t is $v(t)$, then the displacement is $x(t) = \int_0^t v(s)\, ds$. Assuming the object is initially at rest $v(0) = 0$, the work done by a varying force F in moving the object during the time interval $[0, T]$ is

$$\begin{aligned}
W &= \int_0^T F(t)v(t)\, dt = \int_0^T (mv)'v\, dt \\
&= (mv)v\big|_0^T - \int_0^T mv(v')\, dt && \text{integration by parts} \\
&= m(v(T))^2 - \int_0^T \frac{m_0 vv'}{\sqrt{1-\frac{v^2}{c^2}}}\, dt \\
&= m(v(T))^2 - \left(-c^2 m_0 \sqrt{1 - \frac{v^2}{c^2}}\right)\Bigg|_0^T && \text{FTC} \\
&= m(v(T))^2 - \left(-c^2 m_0 \sqrt{1 - \frac{(v(T))^2}{c^2}} + c^2 m_0 \sqrt{1 - \frac{0^2}{c^2}}\right) && \text{since } v(0) = 0 \\
&= m(v(T))^2 + c^2 m_0 \sqrt{1 - \frac{(v(T))^2}{c^2}} - m_0 c^2 \\
&= m(v(T))^2 + mc^2 \left(1 - \frac{(v(T))^2}{c^2}\right) - m_0 c^2 && \text{using (C.15.45)} \\
&= m(v(T))^2 + mc^2 - m(v(T))^2 - m_0 c^2 \\
&= mc^2 - m_0 c^2.
\end{aligned}$$

We can interpret this as saying that the "total energy associated with the object" increases from m_0c^2 to mc^2. The energy of the object at rest is then m_0c^2, called its **rest energy**.

That is the mathematics behind the equation $E = mc^2$. It suggests that mass may be turned into energy, as Einstein predicted. For instance, in a nuclear reactor some of the mass of the uranium is indeed turned into energy in the fission process. Also, the mass of the sun decreases as it emits radiant energy.

What about the equation that states kinetic energy is half the product of the mass and the square of the velocity? That is what (C.15.46) resembles when v is small (compared to c). In this case the first two terms of the binomial series for $(1-x^2)^{-1/2}$, with $x = v^2/c^2$, give

$$\begin{aligned} mc^2 - m_0c^2 = m_0\left(1-\frac{v^2}{c^2}\right)^{-1/2} c^2 - mc^2 &\approx m_0c^2\left(1+\frac{1}{2}\frac{v^2}{c^2}\right) - m_0c^2 \\ &= m_0c^2 + \frac{1}{2}\frac{m_0c^2v^2}{c^2} - m_0c^2 \\ &= \frac{m_0v^2}{2}. \end{aligned}$$

So the increase in energy is well approximated by the familiar kinetic energy, $\frac{1}{2}m_0v^2$.

Calculus is Everywhere # 15

Sparse Traffic

Customers arriving at a checkout counter, cars traveling on a one-way road, raindrops falling on a street and cosmic rays entering the atmosphere all illustrate one mathematical idea — the theory of sparse traffic involving independent events. We will develop the mathematics, which is the basis of the study of waiting time – whether customers at the checkout counter or telephone calls at a switchboard.

First we sketch briefly a bit of probability theory.

Some Probability Theory

The probability that an event occurs is measured by a number p, which can be anywhere from 0 up to 1; $p = 1$ implies the event will certainly occur with negligible exceptions and $p = 0$ that it will not occur with negligible exceptions. The probability that a penny turns up heads is $p = 1/2$ and that a die turns up 2 is $p = 1/6$. (The phrase "certainly occurs with negligible exceptions" means, roughly, that the times the event does not occur are so rare that we may disregard them. Similarly, the phrase "certainly will not occur with negligible exceptions" means, roughly, that the times the event does not occur are so rare that we may disregard them.)

The probability that two events that are independent of each other both occur is the product of their probabilities. For instance, the probability of getting heads when tossing a penny and a 2 when tossing the die is $p = \left(\frac{1}{2}\right)\left(\frac{1}{6}\right) = \frac{1}{12}$.

The probability that exactly one of several mutually exclusive events occurs is the sum of their probabilities. For instance, the probability of getting a 2 or a 3 with a die is $\frac{1}{6} + \frac{1}{6} = \frac{1}{3}$.

With that thumbnail introduction, we will analyze sparse traffic on a one-way road. We will assume that the cars enter the traffic independently of each other and travel at the same speed. Finally, to simplify matters, we assume each car is a point.

The Model

To construct our model we introduce the functions $P_0, P_1, P_2, \ldots, P_n, \ldots$ where $P_n(x)$ shall be the probability that any interval of length x contains exactly n

cars (independently of the location of the interval). Thus $P_0(x)$ is the probability that an interval of length x is empty. We shall assume that

$$P_0(x) + P_1(x) + \cdots + P_n(x) + \cdots = 1 \qquad \text{for any } x.$$

We also shall assume that $P_0(0) = 1$ ("the probability is 1 that a given point contains no cars").

For our model we make the following two major assumptions:

(a) The probability that exactly one car is in any fixed short section of the road is approximately proportional to the length of the section. That is, there is some positive number k such that

$$\lim_{\Delta x \to 0} \frac{P_1(\Delta x)}{\Delta x} = k.$$

(b) The probability that there is more than one car in any fixed short section of the road is neglible, even when compared to the length of the section. That is,

$$\lim_{\Delta x \to 0} \frac{P_2(\Delta x) + P_3(\Delta x) + P_4(\Delta x) + \cdots}{\Delta x} = 0. \tag{C.15.30}$$

We shall now put assumptions (a) and (b) into more useful forms. If we let

$$\epsilon = \frac{P_1(\Delta x)}{\Delta x} - k \tag{C.15.31}$$

where ϵ depends on Δx, assumption (a) tells us that $\lim_{\Delta x} \epsilon = 0$. Thus, solving (C.13.27) for $P_1(\Delta x)$, we see that assumption (a) can be phrased as

$$P_1(\Delta x) = k\Delta x + \epsilon \Delta x \tag{C.15.32}$$

where $\epsilon \longrightarrow 0$ as $\Delta x \longrightarrow 0$.

Since $P_0(\Delta x) + P_1(\Delta x) + \cdots + P_n(\Delta x) + \cdots = 1$, assumption (b) may be expressed as

$$\lim_{\Delta x \to 0} \frac{1 - P_0(\Delta x) - P_1(\Delta x)}{\Delta x} = 0. \tag{C.15.33}$$

In light of assumption (a), equation (C.13.29) is equivalent to

$$\lim_{\Delta x \to 0} \frac{1 - P_0(\Delta x)}{\Delta x} = k. \tag{C.15.34}$$

In the manner in which we obtained (C.13.28), we may deduce that

$$1 - P_0(\Delta x) = k\Delta x + \delta \Delta x,$$

where $\delta \to 0$ as $\Delta x \to 0$. Thus

$$P_0(\Delta x) = 1 - k\Delta x - \delta\Delta x, \qquad \text{(C.15.35)}$$

where $\delta \to 0$ as $\Delta x \to 0$. *On the basis of (a) and (b), expressed in* (C.13.28) *and* (C.13.31)*, we shall obtain an explicit formula for each* P_n*.*

Let us determine P_0 first. Observe that a section of length $x+\Delta x$ is vacant if its left-hand part of length x is vacant and its right-hand part of length Δx is also vacant. Since the cars move independently of each other, the probability that the whole interval of length $x + \Delta x$ being empty is the product of the probabilities that the two smaller intervals of lengths x and Δx are both empty. (See Figure C.13.26.) Thus we have

Figure C.15.26: No cars in a section of length $x+\Delta x$.

$$P_0(x + \Delta x) = P_0(x)P_0(\Delta x). \qquad \text{(C.15.36)}$$

Recalling (C.13.31), we write (C.13.32) as

$$P_0(x + \Delta x) = P_0(x)(1 - k\Delta x - \delta\Delta x)$$

which a little algebra transforms to

$$\frac{P_0(x + \Delta x) - P_0(x)}{\Delta x} = -(k + \delta)P_0(x). \qquad \text{(C.15.37)}$$

Taking limits on both sides of (C.13.33) as $\Delta x \to 0$, we obtain

$$P_0'(x) = -kP_0(x). \qquad \text{(C.15.38)}$$

(Recall that $\delta \to 0$ as $\Delta x \to 0$.) From (C.13.34) it follows that there is a constant A such that $P_0(x) = Ae^{-kx}$. Since $1 = P_0(0) = Ae^{-k0} = A$, we conclude that $A = 1$, hence

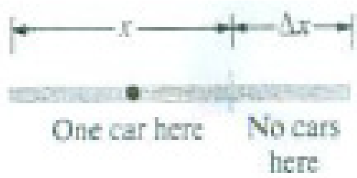
$$P_0(t) = e^{-kx}.$$

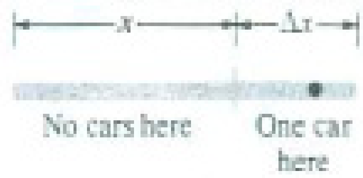
This explicit formula for P_0 is reasonable; e^{-kx} is a decreasing function of x, so that the larger an interval, the less likely that it is empty.

Now let us determine P_1. To do so, we examine $P_1(x + \Delta x)$ and relate it to $P_0(x)$, $P_0(\Delta x)$, $P_1(x)$, and $P_1(\Delta x)$, with the goal of finding an equation involving the derivative of P_1.

Again, imagine an interval of length $x + \Delta x$ cut into two intervals, the left-hand subinterval of length x and the right-hand subinterval of length Δx. Then there is precisely one car in the whole interval if *either* there is exactly one car in the left-hand interval and none in the right-hand subinterval *or* there is none in the left-hand subinterval and exactly one in the right-hand subinterval. (See Figure C.13.27.) Thus we have

$$P_1(x + \Delta x) = P_1(x)P_0(\Delta x) + P_0(x)P_1(\Delta x). \qquad \text{(C.15.39)}$$





(a) (b)

Figure C.15.27: The two ways to have exactly one care in an interval of length $x + \Delta x$.

In view of (C.13.28) and (C.13.31), we may write (C.13.35) as

$$P_1(x + \Delta x) = P_1(x)(1 - k\Delta x - \delta\Delta x) + P_0(x)(k\Delta x + \epsilon\Delta x)$$

which a little algebra changes to

$$\frac{P_1(x + \Delta x) - P_1(x)}{\Delta x} = -(k + \delta)P_1(x) + (k + \epsilon)P_0(x). \qquad \text{(C.15.40)}$$

Letting $\Delta x \to 0$ in (C.13.36) and remembering that $\delta \to 0$ and $\epsilon \to 0$ as $\Delta x \to 0$, we obtain $P_1'(x) = -kP_1(x) + kP_0(x)$; recalling that $P_0(x) = e^{-kx}$, we deduce that

$$P_1'(x) = -kP_1(x) + ke^{-kx}. \qquad \text{(C.15.41)}$$

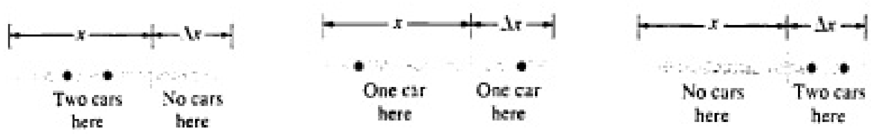
From (C.13.37) we shall obtain an explicit formula for $P_1(x)$. Since $P_0(x)$ involves e^{-kx} and so does (C.13.37). it is reasonable to guess that $P_1(x)$ involves e^{-kx}. Therefore let us express $P_1(x)$ as $g(x)e^{-kx}$ and determine the form of $g(x)$. (Since we have the identity $P_1(x) = \left(P_1(x)e^{kx}\right)e^{-kx}$, we know that $g(x)$ exists.)

According to (C.13.37) we have $\left(g(x)e^{-kx}\right)' = -kg(x)e^{-kx} + ke^{-kx}$; hence

$$g(x)(ke^{-kx}) + g'(x)e^{-kx} = -kg(x)e^{-kx} + ke^{-kx}$$

from which it follows that $g'(x) = k$. Hence $g(x) = kx + c_1$, where c_1 is some constant: $P_1(x) = (kx + c_1)e^{-kx}$. Since $P_1(0) = 0$, we have $P_1(0) = (k \cdot 0 + c_1)e^{-k\cdot 0} = c_1$ and hence $c_1 = 0$. Thus we have shown that

$$P_1(x) = kxe^{-kx} \qquad \text{(C.15.42)}$$



(a) (b) (c)

Figure C.15.28: The three ways to have exactly two cars in an interval of length $x + \Delta x$.

and P_1 is completely determined.

To obtain P_2 we argue as we did in obtaining P_1. Instead of (C.13.35) we have

$$P_2(x + \Delta x) = P_2(x)P_0(\Delta x) + P_1(x)P_1(\Delta x) + P_0(x)P_2(\Delta x) \qquad \text{(C.15.43)}$$

an equation that records the three ways in which two cars in a section of length $x + \Delta x$ can be situated in a section of length x and a section of length Δx. (See Figure C.13.28.)

Similar reasoning shows that

See Exercise 8.

$$P_2(x) = \frac{k^2x^2}{2}. \qquad \text{(C.15.44)}$$

Then, applying the same reasoning inductively leads to

See Exercises 9 and 10.

$$P_n(x) = \frac{(kx)^n}{n!}e^{-kx}. \qquad \text{(C.15.45)}$$

We have obtained in (C.13.41) the formulas on which the rest of our analysis will be based. Note that these formujlas refer to a road section of any length, though the assumptions (a) and (b) refer only to short sections. What has enabled us to go from the "microscopic" to the "macroscopic" is the additional assumption that the traffic in any one section is independent of the traffic in any other section. The formulas (C.13.41) are known as the **Poisson formulas**.

The Meaning of k

The constant k was defined in terms of arbitrarily short intervals, at the "microscopic level". How would we compute k in terms of observable data, at the "macroscopic level"? It turns out that k records the traffic density: the average number of events during an interval of length x is kx.

The average number of events in a section of length x is defined as $\sum_{n=0}^{\infty} nP_n(x)$. This weights each possible number of events (n) with it's likelihood of occurring ($P_n(x)$). This average is

$$\begin{aligned}\sum_{n=0}^{\infty} nP_n(x) &= \sum_{n=1}^{\infty} nP_n(x) = \sum_{n=1}^{\infty} n\frac{(kx)^n e^{-kx}}{n!} \\ &= kxe^{-kx}\sum_{n=1}^{\infty}\frac{(kx)^{n-1}}{(n-1)!} \\ &= kxe^{-kx}\sum_{n=0}^{\infty}\frac{(kn)^n}{n!} = kxe^{-kx}e^{kx} = kx.\end{aligned}$$

Thus the expected number of cars in a section is proportional to the length of the section. This shows that the k appearing in assumption (a) is the measure of traffic density, the number of cars per unit length of road.

To estimate k, in the case of traffic for instance, divide the number of cars in a long section of the road by the length of that section.

EXAMPLE 1 (Traffic at a checkout counter.) Customers arrive at a checkout counter at the rate of 15 per hour. What is the probability that exactly five customers will arrive in any given 20-minute period?
SOLUTION We may assume that the probability of exactly one customer coming in a short interval of time is roughly proportional to the duration of that interval. Also, there is only a negligible probability that more than one customer may arrive in a brief interval of time. Therefore conditions (a) and (b) hold, if we replace "length of section" by "length of time". Without further ado, we conclude that the probability of exactly n customers arriving in a period of x minutes is given by (C.13.41). Moreover, the "customer density" is one per 4 minutes; hence $k = 1/4$, and thus the probability that exactly five customers arrive during a 20-minute period, $P_5(20)$, is

$$\left(\frac{1}{4}\cdot 20\right)^5 \frac{e^{-(1/4)\cdot 20}}{5!} = \frac{5^5 e^{-5}}{120} \approx 0.17547.$$

◇

Modeling of the type within this section is of use in predicting the length of waiting lines (or times) or the waiting time to cross. This is part of the theory of queues. See, for instance, Exercises 2 and 3. (See also Exercise **??** in the Summary Exercises in Chapter **??**.)

EXERCISES

1.[R]

(a) Why would you expect that $P_0(a+b) = P_0(a) \cdot P_0(b)$ for any a and b?

(b) Verify that $P_0(x) = e^{-kx}$ satisfies the equation in (a).

2.[R] A cloud chamber registers an average of four cosmic rays per second.

(a) What is the probability that no cosmic rays are registered for 6 seconds?

(b) What is the probability that exactly two are registered in the next 4 seconds?

3.[R] Telephone calls during the busy hour arrive at a rate of three calls per minute. What is the probability that none arrives in a period of (a) 30 seconds, (b) 1 minute, (c) 3 minutes?

4.[R] In a large continually operating factory there are, on the average, two accidents per hour. Let $P_n(x)$ denote the probability that there are exactly n accidents in an interval of time of length x hours.

(a) Why is it reasonable to assume that there is a constant k such that $P_0(x)$, $P_1(x)$, ... satisfy 1 and 2 on page 37?

(b) Assuming that these conditions are satisfied, show that $P_n(x) = (kx)^n e^{-kx}/n!$.

(c) Why must $k = 2$?

(d) Compute $P_0(1)$, $P_1(1)$, $P_2(1)$, $P_3(1)$, and $P_4(1)$.

5.[R] A typesetter makes an average of one mistake per page. Let $P_n(x)$ be the probability that a section of x pages (x need not be an integer) has exactly n errors.

(a) Why would you expect $P_n(x) = x^n e^{-x}/n!$?

(b) Approximately how many pages would be error-free in a 300-page book?

6.[R] In a light rainfall you notice that on one square foot of pavement there are an average of 3 raindrops. Let $P_n(x)$ be the probability that there are n raindrops on an area of x square feet.

(a) Check that assumptions 1 and 2 are likely to hold.

(b) Find the probability that an area of 3 square feet has exactly two raindrops.

(c) What is the most likely number of raindrops to find on an area of one square foot?

7.[R] Write x^2 in the form $g(x)e^{-kx}$.

8.[R] Show that $P_2(x) = \frac{k^2x^2}{2}e^{-kx}$.

9.[R] Show that $P_3(x) = \frac{(kx)^3}{3!}e^{-kx}$.

10.[M] Show that $P_n(x) = \frac{(kx)^n}{n!}e^{-kx}$.

11.[R]

(a) Why would you expect $P_3(a+b) = P_0(a)P_3(b) + P_1(a)P_2(b) + P_2(a)P_1(b) + P_3(a)P_0(b)$?

(b) Do functions defined in (C.13.41) satisfy the equation in (a)?

12.[R]

(a) Why would you expect $\lim_{n\to\infty} P_n(x) = 0$?

(b) Show that the functions defined in (C.13.41) have the limit in (a).

13.[R]

(a) Why would you expect $\lim_{x\to 0} P_1(x) = 1$ and, for all $n \geq 1$, $\lim_{x\to 0} P_n(x) = 0$?

(b) Show that the functions defined in (C.13.41) satisfy the limit in (a).

14.[R] We obtained $P_0(x) = e^{-kx}$ and $P_1(x) = kxe^{-kx}$. Verify that $\lim_{\Delta x\to 0} P_1(\Delta x)/\Delta x = k$, and $\lim_{\Delta x\to 0} P_0(\Delta x)/\Delta x = 1-k$. Hence show that $\lim_{\Delta x\to 0}(P_2(\Delta x)+P_3(\Delta x)+\cdots+)/\Delta x = 0$, and that assumptions 1 and 2 on page 37 are indeed satisfied.

15.[R]

(a) Obtain assumption 1 from equation (C.13.28).

(b) Obtain equation (C.13.28) from assumption 2.

(c) Obtain assumption 2 from equation (C.13.31).

16.[M] What length of road is most likely to contain exactly one car? That is, what x maximizes $P_1(x)$?

17.[M] What length of road is most likely to contain three cars?

18.[M] For any $x \geq 0$, $\sum_{n=0}^{\infty} P_n(x)$ should equal 1 because it is certain that some number of cars is in a given section of length x (maybe 0 cars). Check that $\sum_{n=0}^{\infty} P_n(x) = 1$. NOTE: This provides a probabilistic argument that $e^u = \sum_{n=0}^{\infty} u^n/n!$ for $n \geq 0$.

19.[M] Planes arrive randomly at an airport at the rate of one per 2 minutes. What is the probability that more than three planes arrive in a 1-minute interval?

Calculus is Everywhere # 16

Space Flight: The Gravitational Slingshot

For vector-algebra chapter

In a "slingshot" or "gravitational assist" a spacecraft picks up speed as it passes near a planet and exploits the planet's gravity. For instance, New Horizons, launched on January 19, 2006, enjoys a gravitational assist as it passed by Jupiter, February 27, 2007 on its long journey to Pluto. With the aid of that slingshot the speed of the spacecraft increased from 47,000 to 50,000 miles per hour (mph). As a result, it will arrive near Pluto in 2015, instead of 2018.

Before we see how this technique works, let's look at a simple situation on earth that illustrates the idea. Later we will replace the truck with a planet's gravitational field.

A playful lad throws a perfectly elastic tiny ball at 30 mph directly at a truck approaching him at 70 miles per hour, as shown in Figure C.17.34.

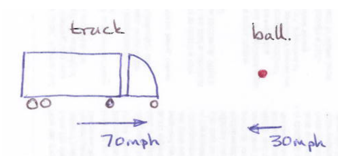


Figure C.16.29:

The truck driver sees the ball coming toward her at $70 + 30 = 100$ mph. The balls hits the windshield and, because the ball is perfectly elastic, the driver sees it bounce off at 100 mph in the opposite direction.

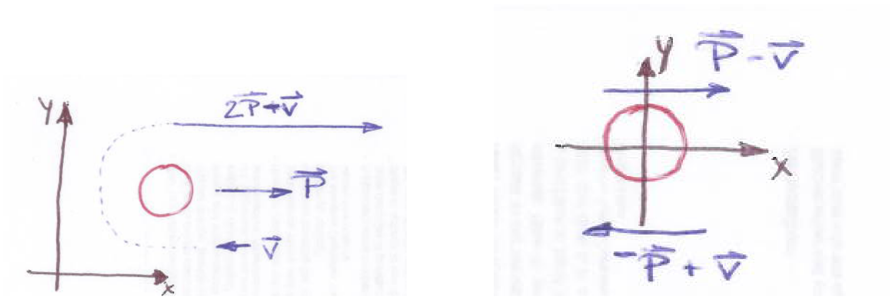
However, because the truck is moving in the same direction as the ball, the ball is moving through the air at $100 + 70 = 170$ mph as it returns to the boy. The ball has gained 140 mph, twice the speed of the truck.

Now, instead of picturing a truck, think of a planet whose velocity relative to the solar system is represented by the vector $\mathbf{P}$. A spacecraft, moving in the opposite direction with the velocity $\mathbf{v}$ relative to the solar system comes close to the planet.

An observer on the planet sees the spacecraft approaching with velocity $-vP + \mathbf{v}$. The spacecraft swings around the planet as gravity controls its orbit and sends it off in the opposite direction. Whatever speed it gained as it arrived, it loses as it exits. Its velocity vector when it exits is $-(-vP + \mathbf{v}) = \mathbf{P} - \mathbf{v}$, as viewed by the observer on the planet. Since the planet is moving through the solar system with velocity vector $\mathbf{P}$, the spacecraft is now moving through the solar system with velocity $\mathbf{P} + (\mathbf{P} - \mathbf{v}) = 2\mathbf{P} - \mathbf{v}$. See Figure C.17.35.

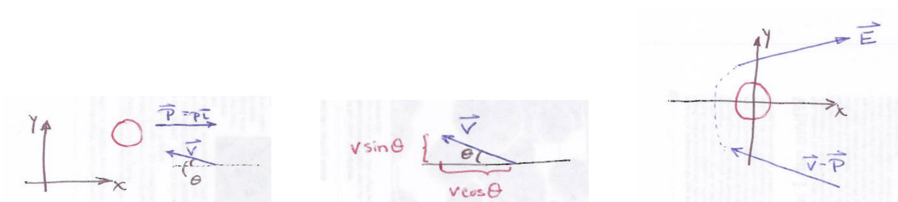
If $\mathbf{P} = 70\mathbf{i}$ and $\mathbf{v} = -30\mathbf{i}$, we have the vector $2(70\mathbf{i}) - (-30\mathbf{i}) = 170\mathbf{i}$, the case of the ball and truck.

But the direction of the spacecraft as it arrives may not be exactly opposite the direction of the planet. To treat the more general case, assume that $\mathbf{P} = p\mathbf{i}$, where p is positive and $\mathbf{v}$ makes an angle θ, $0 \leq \theta \leq \pi/2$, with $-\mathbf{i}$, as shown in Figure C.17.36(a). Let $v = |\mathbf{v}|$ be the speed of the spacecraft relative to the solar system. We will assume that the spacecraft's speed (relative to the planet) as it exits is the same as its speed relative to the planet on its arrival. (Figure C.17.36(b)) shows the arrival and exit vectors. Note that $\mathbf{E}$ and $\mathbf{v} - \mathbf{P}$



(a) (b)

Figure C.16.30: (a) The velocity vector relative to the solar system. (b) The velocity vector relative to the planet.



(a) (b) (c)

Figure C.16.31:

have the same y-components, but the x-component of $\mathbf{E}$ is the negative of the x-component of $\mathbf{v} - \mathbf{P}$.

Figure C.17.36(c) shows the arrival vector relative to the solar system. So, $\mathbf{v} = -w\cos(\theta)\mathbf{i} + v\sin(\theta)\mathbf{j}$.

Relative to the planet we have

$$\begin{array}{lrcl} \text{Arrival Vector:} & \mathbf{v} - \mathbf{P} & = & -p\mathbf{i} + (-v\cos(\theta)\mathbf{i} + v\sin(\theta)\mathbf{j}) \\ \text{Exit Vector:} & \mathbf{E} & = & p\mathbf{i} + v\cos(theta)\mathbf{i} + v\sin(\theta)\mathbf{j} \end{array}$$

The exit vector relative to the solar system, $\mathbf{E}$, is therefore

$$\mathbf{E} = (2p + v\cos(\theta))\mathbf{i} + v\sin(theta)\mathbf{j}.$$

The magnitude of $\mathbf{E}$ is

$$\sqrt{(2p + v\cos(\theta))^2 + (v\sin(theta))^2} = \sqrt{v^2 + 2pv\cos(\theta) + 4p^2}.$$

When $\theta = 0$, we have the case of the truck and ball or the planet and spacecraft in Figure C.17.35. Then $\cos(\theta) = 1$ and $|\mathbf{E}| = \sqrt{v^2 + 2pv + 4p^2} = v + 2p$, in agreement with our earlier observations.

The scientists controlling a slingshot carry out much more extensive calculations, which take into consideration the masses of the spacecraft and the planet, and involve an integration while the spacecraft is near the planet. Incidentally, the diameter of Jupiter is 86,000 miles.

"Near" in the case of the slingshot around Jupiter means 1.4 million miles. If the spacecraft gets too close, the atmosphere slows down or destroys the craft.

The gravity assist was proposed by Michael Minovitch in 1963 when he was still a graduate student at UCLA. Before then it was felt that to send a spacecraft to the outer solar system and beyond would require launch vehicles with nuclear reactors to achieve the necessary thrust.

Calculus is Everywhere # 17
How to Find Planets around Stars

Astronomers have discovered that other stars than the sun have planets circling them. How do they do this, given that the planets are too small to be seen? It turns out that they combine some vector calculus with observations of the star. Let us see what they do.

Imagine a star S and a planet P in orbit around S. To describe the situation, we are tempted to choose a coordinate system attached to the star. In that case the star would appear motionless, hence having no acceleration. However, the planet exerts a gravitational force F on the star and the equation force = mass × acceleration would be violated. After introducing the appropriate mathematical tools, we will choose a proper coordinate system.

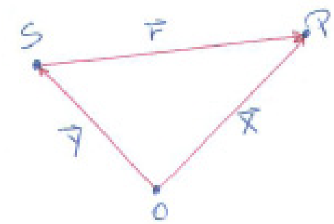


Figure C.17.32:

Let $\mathbf{X}$ be the position vector of the planet P and $\mathbf{Y}$ be the position vector of the star S, relative to our inertial system. Let M be the mass of the sun and m the mass of planet P. Let $\mathbf{r} = \mathbf{X} - \mathbf{Y}$ be the vector from the star to the planet, as shown in Figure C.18.37.

The gravitational pull of the star on the planet is proportional to the product between them:

$$\mathbf{F} = \frac{-GmM\mathbf{r}}{r^3}.$$

Here G is a universal constant, that depends on the units used to measure mass, length, time, and force. Equating the force with mass times acceleration, we have

$$\begin{aligned} M\mathbf{X}'' &= \tfrac{-GmM\mathbf{r}}{r^3}. \\ \text{Thus} \qquad \mathbf{X}'' &= \tfrac{-Gm\mathbf{r}}{r^3}. \end{aligned}$$

Similarly, by calculating the force that the planet exerts on the star, we have

$$\mathbf{Y}'' = \frac{Gm\mathbf{r}}{r^3}.$$

The center of gravity of the system consisting of the planet and the star, which we will denote C (see Figure C.18.38), is given by

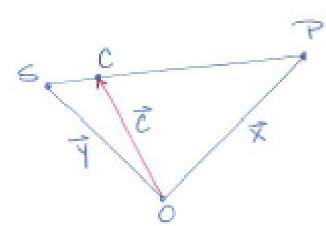


Figure C.17.33:

$$\mathbf{C} = \frac{M\mathbf{Y} + m\mathbf{X}}{M + m}.$$

The center of gravity is much closer to the start than to the planet. In the case of our sun and Earth, the center of gravity is a mere 300 miles from the center of the sun.

The acceleration of the center of gravity is

$$\mathbf{C}'' = \frac{M\mathbf{Y}'' + m\mathbf{X}''}{M+m} = \frac{1}{M+m}\left(M\left(\frac{Gm\mathbf{r}}{r^3}\right) + m\left(\frac{-Gm\mathbf{r}}{r^3}\right)\right) = \mathbf{0}.$$

Because the center of gravity has $\mathbf{0}$-acceleration, it is moving at a constant velocity relative to the coordinate system we started with. Therefore a coordinate system rigidly attached to the center of gravity may also serve as an inertial system in which the laws of physics still hold.

We now describe the position of the star and planet to this new coordinate system. Star S has the vector $\mathbf{x}$ from $\mathbf{C}$ to it and planet P has the vector $\mathbf{y}$ from $\mathbf{C}$ to it, as shown in Figure C.18.39. Note that $\mathbf{r} = \mathbf{x} - \mathbf{y}$.

Figure C.17.34:

To obtain a relation between $\mathbf{x}$ and $\mathbf{y}$, we first express each in terms of $\mathbf{r}$. We have

$$\mathbf{y} = \mathbf{Y} - \vec{OC} = \mathbf{Y} - \frac{M\mathbf{Y} - m\mathbf{X}}{M+m} = \frac{m}{M+m}\mathbf{Y} + \frac{m}{M+m}\mathbf{X}.$$

Letting $k = m/M$, a very small quantity, we have

$$\mathbf{y} = \frac{k}{1+k}(\mathbf{Y} - \mathbf{X}) = \frac{-k}{1+k}\mathbf{r}. \qquad \text{(C.17.46)}$$

Since $\mathbf{r} = \mathbf{x} - \mathbf{y}$, it follows that $\mathbf{x} = \mathbf{r} + \mathbf{y}$, hence

$$\mathbf{x} = \mathbf{r} + \left(\frac{-k}{1+k}\right)\mathbf{r} = \frac{1}{1+k}\mathbf{r}. \qquad \text{(C.17.47)}$$

Combining (C.18.56) and (C.18.57) shows that

$$\mathbf{y} = -k\mathbf{x}. \qquad \text{(C.17.48)}$$

Equation (C.18.58) tells us a good deal about the relation between the orbits of the star and planet in terms of the second inertial system:

SHERMAN: First use of "second inertial system;" what is the first?

1. The star and planet remain on opposite sides of C on a straight line through C.
2. The star is always much closer to C than the planet is.
3. The orbit of the star is similar in shape to the orbit of the planet, but smaller and reflected through C.

4. If the orbit of the star is periodic so is the orbit of the planet, and both have the same period.

Equation (C.18.58) is the key to the discover of planets around stars. The astronomers look for a star that "wobbles" a bit. That wobble is the sign that the star is in orbit around the center of gravity of it and some planet. Moveover, the time it takes for the planet to orbit the star is simply the time it takes for the star to oscillate back and forth once.

The reference cited below shows that the star and the planet sweep out elliptical orbits in the second coordinate system (the one relative to C).

Astronomers have found over two hundred stars with planets, some with several planets. A registry of these **exoplanets** is maintained at `http://exoplanets.org/`.

Reference: Robert Osserman, *Kepler's Laws, Newton's Laws, and the Search for New Planets*, Am. Math. Monthly **108** (2001), pp. 813–820.

EXERCISES 1.[R] The mass of the sun is about 330,000 times that of Earth.

The closest Earth gets to the sun is about 91,341,000 miles, and the farthest from it is about 94,448,000 miles. What is the closest the center of the sun gets to the center of gravity of the sun-Earth system? What is the farthest it gets from it? Hint: It lies within the sun itself.

2.[M] Find the condition that must be satisfied if the center of gravity of a sun-planet system will lie outside the sun.

SHERMAN: See `http://en.wikipedia.org/wiki/Center_of_mass`, particularly the animations at the end of the section on "Barycenter in astrophysics and astronomy".

Calculus is Everywhere # 18
Newton's Law Implies Kepler's Three Laws

After hundreds of pages of computation based on observations by the astronomer Tycho Brahe (1546—1601) in the last 30 years of the sixteenth century, plus lengthy detours and lucky guesses, Kepler (1571–1630) arrived at these three laws of planetary motion:

Kepler's Three Laws

1. Every planet travels around the sun in an elliptical orbit such that the sun is situated at one focus (discovered in 1605, published in 1609).
2. The velocity of a planet varies in such a way that the line joining the planet to the sun sweeps out equal areas in equal times (discovered 1602, published 1609).
3. The square of the time required by a planet for one revolution around the sun is proportional to the cube of its mean distance from the sun (discovered 1618, published 1619).

The work of Kepler shattered the crystal spheres which for 2,000 years had carried the planets. Before him astronomers admitted only circular motion and motion compounded of circular motions. Copernicus (1473–1543), for instance, used five circles to describe the motion of Mars.

The ellipse got a cold reception.

The ellipse was not welcomed. In 1605 Kepler complained to a skeptical astronomer:

> You have disparaged my oval orbit . . . . If you are enraged because I cannot take away oval flight how much more you should be enraged by the motions assigned by the ancients, which I did take away . . . . You disdain my oval, a single cart of dung, while you endure the whole stable. (If indeed my oval is a cart of dung.)

But the astronomical tables that Kepler based on his theories, and published in 1627, proved to be more accurate than any other, and the ellipse gradually gained acceptance.

The three laws stood as mysteries alongside a related question: If there are no crystal spheres, what propels the planets? Bullialdus (1605–1694), a French astronomer and mathematician, suggested in 1645:

The inverse square law was conjectured.

> The force with which the sun seizes or pulls the planets, a physical force which serves as hands for it, is sent out in straight lines into all the world's space . . . ; since it is physical it is decreased in greater space; . . . the ratio of this distance is the same as that for light, namely as the reciprocal of the square of the distance.

In 1666, Hooke (1635–1703), more of an experimental scientist than a mathematician, wondered:

> why the planets should move about the sun . . . being not included in any solid orbs . . . nor tied to it . . . by any visible strings . . . . I cannot imagine any other likely cause besides these two: The first may be from an unequal density of the medium . . . ; if we suppose that part of the medium, which is farthest from the centre, or sun, to be more dense outward, than that which is more near, it will follow, that the direct motion will be always deflected inwards, by the easier yielding of the inwards . . . .
>
> But the second cause of inflecting a direct motion into a curve may be from an attractive property of the body placed in the center; whereby it continually endeavours to attract or draw it to itself. For if such a principle be supposed all the phenomena of the planets seem possible to be explained by the common principle of mechanic motions. . . . By this hypothesis, the phenomena of the comets as well as of the planets may be solved.

In 1675, Hooke, in an announcement to the Royal Society, went further:

> All celestial bodies have an attraction towards their own centres, whereby they attract not only their own parts but also other celestial bodies that are within the sphere of their activity . . . . All bodies that are put into direct simple motion will so continue to move forward in a single line till they are, by some other effectual powers, deflected and bent into a motion describing a circle, ellipse, or some other more compound curve . . . . These attractive powers are much more powerful in operating by how much the nearer the body wrought upon is to their own centers . . . . It is a notion which if fully prosecuted as it ought to be, will mightily assist the astronomer to reduce all the celestial motions to a certain rule . . . .

Hooke pressed Newton to work on the problem.

Trying to interest Newton in the question, Hooke wrote on November 24, 1679: "I shall take it as a great favor if . . . you will let me know your thoughts of that of compounding the celestial motion of planets of a direct motion by the tangent and an attractive motion toward the central body." But four days later Newton replied:

> My affection to philosophy [science] being worn out, so that I am almost as little concerned about it as one tradesman used to be about another man's trade or a countryman about learning. I must acknowledge myself averse from spending that time in writing about it which I think I can spend otherwise more to my own content and the good of others . . . .

In a letter to Newton on January 17, 1680, Hooke returned to the problem of planetary motion:

> It now remains to know the properties of a curved line (not circular . . .) made by a central attractive power which makes the velocities of descent from the tangent line or equal straight motion at all distances in a duplicate proportion to the distances reciprocally taken. I doubt not that by your excellent method you will easily find out what that curve must be, and its properties, and suggest a physical reason for this proportion.

Hooke succeeded in drawing Newton back to science, as Newton himself admitted in his *Principia*, published in 1687: "I am beholden to him only for the diversion he gave me from the other studies to think on these things and for his dogmaticalness in writing as if he had found the motion in the ellipse, which inclined me to try it."

It seems that Newton then obtained a proof — perhaps containing a mistake (the history is not clear) — that if the motion is elliptical, the force varies as the inverse square. In 1684, at the request of the astronomer Halley, Newton provided a correct proof. With Halley's encouragement, Newton spent the next year and a half writing the *Principia*.

Halley, of Halley's comet, paid for publication of the *Principia*.

In the *Principia*, which develops the science of mechanics and applies it to celestial motions, Newton begins with two laws:

1. Every body continues in its state of rest, or of uniform motion in a straight line, unless it is compelled to change this state by forces impressed upon it.

2. The change of momentum is proportional to the motive force impressed, and is made in the direction of the straight line in which that force is impressed.

To state these in the language of vectors, let $\mathbf{v}$ be the velocity of the body, $\mathbf{F}$ the impressed force, and m the mass of the body. The first law asserts that $\mathbf{v}$ is constant if $\mathbf{F}$ is $\mathbf{0}$. **Momentum** is defined as $m\mathbf{v}$; the second law asserts that

$$\mathbf{F} = \frac{d}{dt}(m\mathbf{v}).$$

If m is constant, this reduces to

$$\mathbf{F} = m\mathbf{a},$$

where $\mathbf{a}$ is the acceleration vector.

Newton assumed a universal **law of gravitation**. Any particle P exerts an attractive force on any other particle Q, and the direction of the force is from Q toward P. Then *assuming* that the orbit of a planet moving about the sun (both treated as points) is an ellipse, he *deduced* that this force is inversely proportional to the square of the distance between the particles P and Q.

Nowhere in the *Principia* does he deduce from the inverse-square law of gravity that the planets' orbits are ellipses. (However, there are general theorems in *Principia* on the basis of which this deduction could have been made.) In the *Principia* he showed that Kepler's second law (concerning areas) was equivalent to the assumption that the force acting on a planet is directed toward the sun. Finally, he deduced Kepler's third law.

Newton's universal law of gravitation asserts that any particle, of mass M, exerts a force on any other particle, of mass m, and that the magnitude of this force is proportional to the product of the two masses, mM, inversely proportional to the square of the distance between them, and is directed toward the particle with the larger mass. (Here, we assume $M > m$.)

We will assume that the sun is fixed at O.

Assume that the sun has mass M and is located at point O and that the planet has mass m and is located at point P. (See Figure C.19.40.) Let $\mathbf{r} = \vec{OP}$ and $r = \|\mathbf{r}\|$. Then the sun exerts a force $\mathbf{F}$ on the planet given by the formula

$$\mathbf{F} = -\frac{GmM}{r^3}\mathbf{r}, \quad \text{(C.18.49)}$$

where G is a universal constant. It is convenient to introduce the unit vector $\mathbf{u} = \mathbf{r}/r$, which points in the direction of $\mathbf{r}$. Then (C.19.59) reads

$$\mathbf{F} = -\frac{GmM}{r^2}\mathbf{u}.$$

Now, $\mathbf{F} = m\mathbf{a}$, where $\mathbf{a}$ is the acceleration vector of the planet. Thus

$$m\mathbf{a} = -\frac{GmM}{r^2}\mathbf{u},$$

from which it follows that

$$\mathbf{a} = -\frac{q\mathbf{u}}{r^2}, \quad \text{(C.18.50)}$$

where $q = GM$ is independent of the planet.

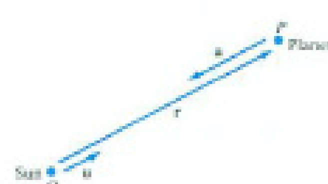


Figure C.18.35:

The vectors $\mathbf{u}$, $\mathbf{r}$, and $\mathbf{a}$ are indicated in Figure C.19.40.

The following exercises show how to obtain Kepler's three laws from the single law of Newton, $\mathbf{a} = -q\mathbf{u}/r^2$.

EXERCISES

Exercises 1 to 3 obtain Kepler's "area" law.

1.[R] Let $\mathbf{r}(t)$ be the position vector of a given planet at time t. Let $\Delta\mathbf{r} = \mathbf{r}(t+\Delta t) - \mathbf{r}(t)$. Show that for small Δt,

$$\frac{1}{2}\left\|\mathbf{r}\times\Delta\mathbf{r}\right\|$$

approximates the area swept out by the position vector during the small interval of time Δt. Hint: Draw a picture.

Figure C.18.36:

2.[R] From Exercise 1 deduce that $\frac{1}{2}\left\|\mathbf{r}\times\frac{d\mathbf{r}}{dt}\right\|$ is the rate at which the position vector $\mathbf{r}$ sweeps out area. (See Figure C.19.41.)

Let $\mathbf{v} = d\mathbf{r}/dt$. The vector $\mathbf{r}\times\mathbf{v}$ will play a central role in the argument leading to Kepler's area law. (See also Exercise **??** in Section **??**.)

3.[R] With the aid of (C.19.60), show that the vector $\mathbf{r}\times\mathbf{v}$ is constant, independent of time.

Since $\mathbf{r}\times\mathbf{v}$ is constant, $\frac{1}{2}\left\|\mathbf{r}\times\mathbf{v}\right\|$ is constant. In view of Exercise 2, it follows that the radius vector of a given planet sweeps out area at a constant rate. **To put it another way, the radius vector sweeps out equal areas in equal times. This is Kepler's second law.**

Introduce an xyz-coordinate system such that the unit vector $\mathbf{k}$, which points in the direction of the positive z axis, has the same direction as the constant vector $\mathbf{r}\times\mathbf{v}$. Thus there is a positive constant h such that

$$\mathbf{r}\times\mathbf{v} = h\mathbf{k}. \tag{C.18.51}$$

Exercises 4 to 13 obtain Kepler's "ellipse" law.

4.[R] Show that h in (C.19.61) is twice the rate at which the position vector of the planet sweeps out area.

5.[R] Show that the planet remains in the plane perpendicular to $\mathbf{k}$ that passes through the sun.

By Exercise 5, the orbit of the planet is planar. We may assume that the orbit lies in the xy plane; for convenience, locate the origin of the xy coordinates at the sun. Also introduce polar coordinates in this plane, with the pole at the sun and the polar axis along the positive x axis, as in Figure C.19.42.

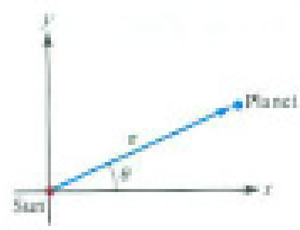


Figure C.18.37:

6.[R]

(a) Show that during the time interval $[t_0, t]$ the position vector of the planet sweeps out the area

$$\frac{1}{2}\int_{t_0}^{t} r^2 \frac{d\theta}{dt} dt.$$

(b) From (a) deduce that the radius vector sweeps out area at the rate $\frac{1}{2}r^2\frac{d\theta}{dt}$.

Henceforth use the dot notation for differentiation with respect to time. Thus $\dot{\mathbf{r}} = \mathbf{v}$, $\dot{\mathbf{v}} = \mathbf{a}$, and $\dot{\theta} = \frac{d\theta}{dt}$.

7.[R] Show that $\mathbf{r} \times \mathbf{v} = r^2\dot{\theta}\mathbf{k}$.

8.[R] Show that $\dot{\mathbf{u}} = \frac{d\mathbf{u}}{d\theta}\dot{\theta}$ and is perpendicular to $\mathbf{u}$. Recall that $\mathbf{u}$ is defined as $\mathbf{r}/\|\mathbf{r}\|$.

9.[R] Recalling that $\mathbf{r} = r\mathbf{u}$, show that $h\mathbf{k} = r^2(\mathbf{u} \times \dot{\mathbf{u}})$.

10.[R] Using (C.19.60) and Exercise 9, show that $\mathbf{a} \times h\mathbf{k} = q\dot{\mathbf{u}}$. HINT: What is the vector identity for $\mathbf{A} \times (\mathbf{B} \times \mathbf{C})$?

11.[R] Deduce from Exercise 10 that $\mathbf{v} \times h\mathbf{k}$ and $q\mathbf{u}$ differ by a constant vector.

By Exercise 11, there is a constant vector $\mathbf{C}$ such that

$$\mathbf{v} \times h\mathbf{k} = q\mathbf{u} + \mathbf{C}. \qquad \text{(C.18.52)}$$

Then the angle between $\mathbf{r}$ and $\mathbf{C}$ is the angle θ of polar coordinates.
The next exercise requires the vector identity $(\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C} = \mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})$, which is valid for any three vectors $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$.
12.[R]

(a) Show that $(\mathbf{r} \times \mathbf{v}) \cdot h\mathbf{k} = h^2$.

(b) Show that $\mathbf{r} \cdot (\mathbf{v} \times h\mathbf{k}) = rq + \mathbf{r} \cdot \mathbf{C}$.

(c) Combining (a) and (b), deduce that $h^2 = rq + rc\cos(\theta)$, where $c = \|\mathbf{C}\|$

It follows from Exercise 12 that the polar equation for the orbit of the planet is given by

$$r(\theta) = \frac{h^2}{q + c\cos(\theta)}. \qquad \text{(C.18.53)}$$

13.[R] By expressing (C.19.63) in rectangular coordinates, show that it describes a conic section.

Since the orbit of a planet is bounded and is also a conic section, it must be an ellipse. This establishes Kepler's first law.
Kepler's third law asserts that the square of the time required for a planet to complete one orbit is proportional to the cube of its mean distance from the sun.
First the term **mean distance** must be defined. For Kepler this meant the average of the shortest distance and the longest distance from the planet to the sun in its orbit. Let us compute this average for the ellipse of semimajor axis a and semiminor axes b, shown in Figure C.19.43. The sun is at the focus F, which is also the pole of the polar coordinate system we are using. The line through the two foci contains the polar axis.

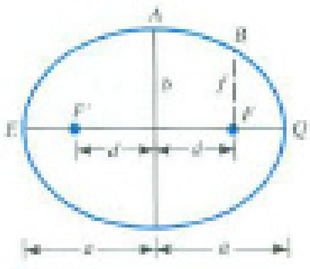


Figure C.18.38:

Recall that an ellipse is the set of points P such that the sum of the distances from P to the two foci F and F' is constant, $2a$. The shortest distance from the planet to the sun is $\overline{FQ} = a - d$ and the longest distance is $\overline{EF} = a + d$. Thus Kepler's mean distance is

$$\frac{(a-d)+(a+d)}{2} = a.$$

Now let T be the time required by the given planet to complete one orbit. Kepler's third law asserts that T^2 is proportional to a^3. Exercises 14 to 18 establish this law by showing that T^2/a^3 is the same for all planets.

14.[R] Using the fact that the area of the ellipse in Figure C.19.43 is πab, show that $Th/2 = \pi ab$, hence that

$$T = \frac{2\pi ab}{h}. \qquad \text{(C.18.54)}$$

The rest of the argument depends only on (C.19.63) and (C.19.64) and the "fixed sum of two distances" property of an ellipse.

15.[R] Using (C.19.63), show that f in Figure C.19.43 equals h^2/q.

16.[R] Show that $b^2 = af$, as follows:

(a) From the fact that $\bar{F'}A + \bar{F}A = 2a$, deduce that $a^2 = b^2 + d^2$.

(b) From the fact that $\bar{F'}B + \bar{F}B = 2a$, deduce that $d^2 = a^2 - af$.

(c) From (a) and (b), deduce that $b^2 = af$.

17.[R] From Exercises 15 and 16, deduce that $b^2 = ah^2/q$.

18.[R] Combining (C.19.64) and Exercise 17, show that

$$\frac{T^2}{a^3} = \frac{4\pi^2}{q}.$$

Since $4\pi^2/q$ is a constant, the same for all points, Kepler's third law is established.

For Further Reading

1. M. Caspar, *Kepler*, Hellman, New York, 1959.
2. J. Kepler, "Copernican astronomy," *Great Books of the Western World*, Encyclopedia Britannica, Chicago, 1952.
3. A. Koestler, *The Watershed*, Doubleday, New York, 1960.
4. I. Newton, *Principia*, University of California, Berkeley, 1947.
5. V. Frederick Rickey, "Isaac Newton: Man, Myth, and Mathematician," *College Mathematics Journal*, **18** (1987), 362—388.
6. J. L. Russell, "Kepler's Laws of Planetary Motion 1609–1666," *British Journal for the History of Science*, **2** (1964), 1—24.
7. S. Stein, "Exactly how did Newton deal with his planets?", *Math. Intell.*, **18** (1996), 6—11.

Calculus is Everywhere # 19

The Suspension Bridge and the Hanging Cable

In a suspension bridge the roadway hangs from a cable, as shown in Figure C.20.44. We will use calculus to find the shape of the cable. To begin, we assume that the weight of any section of the roadway is proportional to its length. That is, there is a constant k such that x feet of the roadway weighs kx pounds. We will assume that the cable itself is weightless. That is justified for it weighs little in comparison to the roadway.

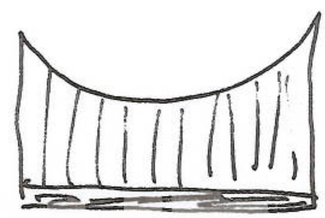
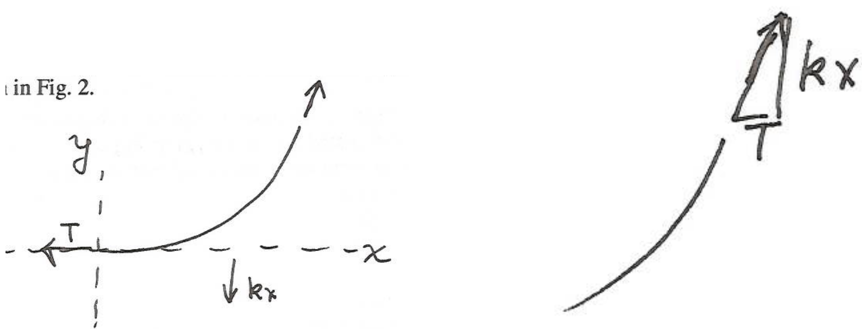


Figure C.19.39:

We introduce an xy-coordinate system with origin at the lowest point of the cable, and consider a typical section of the cable, which goes from $(0, 0)$ to (x, y), as shown in Figure C.20.45(a). Three forces act on this section. The



(a) (b)

Figure C.19.40:

force at $(0, 0)$ is horizontal and pulls the cable to the left. Call its magnitude T. Gravity pulls the cable down with the force kx (the weight of the roadway beneath the cable). At the top of the section, at (x, y) the cable above it pulls the cable to the right and upward, along the tangent line to the cable.

The section does not move, neither up nor down, neither to the left nor to the right. That means the horizontal part of the force at (x, y) must have magnitude T and the vertical part of the force must have magnitude kx, as shown in Figure C.20.45(b). (Think of one person pulling horizontally at (x, y) and another pulling vertically to duplicate the effect of the part of the cable above (x, y) that is pulling on the section.)

Since the force at the point (x, y) is directed along the tangent line there, we have

$$\frac{dy}{dx} = \frac{kx}{T}. \tag{C.19.55}$$

Therefore,

$$y = \frac{kx^2}{T} + C.$$

for some constant C. Since $(0,0)$ is on the curve, $C = 0$, and the cable has the equation

$$y = \frac{kx^2}{tT}.$$

The cable forms a parabola.

But what if, instead, we have the cable but no roadway? That is the case with a laundry line or a telephone wire or a hanging chain. In this case the downward force is due to the weight of the cable. If s feet of cable weighs ks pounds, reasoning almost identical to the case of the suspension bridge leads to the equation

$$\frac{dy}{dx} = \frac{ks}{T}. \tag{C.19.56}$$

Since

$$s = \int_0^x \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx$$

we have the equation

$$\frac{dy}{dx} = \frac{k\int_0^x \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx}{T}. \tag{C.19.57}$$

We get rid of the integral by differentiating both sides of (C.20.67), and using part of the fundamental theorem of calculus, obtaining

$$\frac{d^2y}{dx^2} = \frac{k}{T}\sqrt{1 + \left(\frac{dy}{dx}\right)^2}. \tag{C.19.58}$$

This equation is solved in a differential equations course, where it is shown that

$$y = \frac{k}{T}\left(e^{\frac{kx}{T}} + e^{\frac{-kx}{T}}\right) - 2\frac{k}{T}. \tag{C.19.59}$$

This curve is called a **catenary**, after the Latin "catena," meaning "chain." (Hence the word "concatenation," referring to a chain of events.) It may look like a parabola, but it isn't. The 630-foot tall Gateway Arch in St. Louis, completed October 28, 1965, is the most famous catenary.

EXERCISES

1.[M] Check that the solution to

$$\frac{d^2y}{dx^2} = \frac{k}{T}\sqrt{1 + \left(\frac{dy}{dx}\right)^2}$$

that passes through $(0, 0)$ is

$$y = \frac{k}{T}\left(e^{\frac{kx}{T}} + e^{\frac{-kx}{T}}\right) - 2\frac{k}{T}. \tag{C.19.60}$$

Calculus is Everywhere # 20

The Path of the Rear Wheel of a Scooter

When the front wheel of a scooter follows a certain path, what is the path of the rear wheel? This question could be phrased in terms of a bicycle or car, but the scooter is more convenient for carrying out real-life experiments.

In **??** we considered the special case when the front wheel moves in a straight line, as may occur when parking a car. Now, using vectors, we will look at the case when the front wheel sweeps out a circular path.

The Basic Equation

Figure C.16.29 shows the geometry at any instant. Let s denote the arc length of the path swept out by the rear wheel as measured from its starting point. Let a be the length of the wheel base, that is, the distance between the front and rear axels. The vector $\mathbf{r}(s)$ records the position of the rear wheel and $\mathbf{f}(s)$ records the position of the front wheel. Because the rear wheel is parallel to $\mathbf{f}(s) - \mathbf{r}(s)$, the vector $\mathbf{r}'(s)$ points directly toward the front wheel or directly away from it. Note that $\mathbf{r}'(s)$ is a unit vector.

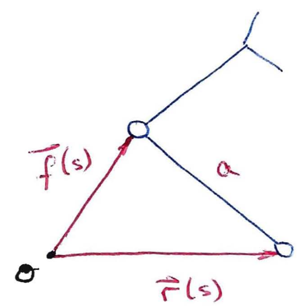


Figure C.20.41:

Thus

$$\mathbf{f}(s) = \mathbf{r}(s) + a\mathbf{r}'(s) \tag{C.20.61}$$

or

$$\mathbf{f}(s) = \mathbf{r}(s) - a\mathbf{r}'(s). \tag{C.20.62}$$

In short, we will write $\mathbf{f}(s) = \mathbf{r}(s) \pm a\mathbf{r}'(s)$.

Assume that the front wheel moving, say, counterclockwise traces out a circular path with center O and radius c. Because

$$\mathbf{f}(s) \cdot \mathbf{f}(s) = c^2,$$

we have

$$(\mathbf{r}(s) \pm a\mathbf{r}'(s)) \cdot (\mathbf{r}(s) \pm a\mathbf{r}'(s)) = c^2.$$

By distributivity of the dot product,

$$\mathbf{r}(s) \cdot \mathbf{r}(s) + a^2\mathbf{r}'(s) \cdot \mathbf{r}'(s) \pm 2a\mathbf{r}(s) \cdot \mathbf{r}'(s) = c^2. \tag{C.20.63}$$

Letting $r(s) = \|\mathbf{r}(s)\|$, we may rewrite (C.16.48) as

$$(r(s))^2 + a^2 \pm 2a\mathbf{r}(s) \cdot \mathbf{r}'(s) = c^2. \tag{C.20.64}$$

Differentiating $\mathbf{r}(s) \cdot \mathbf{r}(s) = r(s)^2$ to obtain, $\mathbf{r}(s) \cdot \mathbf{r}'(s) = r(s)r'(s)$, which changes (C.16.48) to an equation involving the scalar function $r(s)$. For simplicity, we write $r(s)$ as r and $r'(s)$ as r', obtaining

$$r^2 + a^2 \pm 2arr' = c^2. \tag{C.20.65}$$

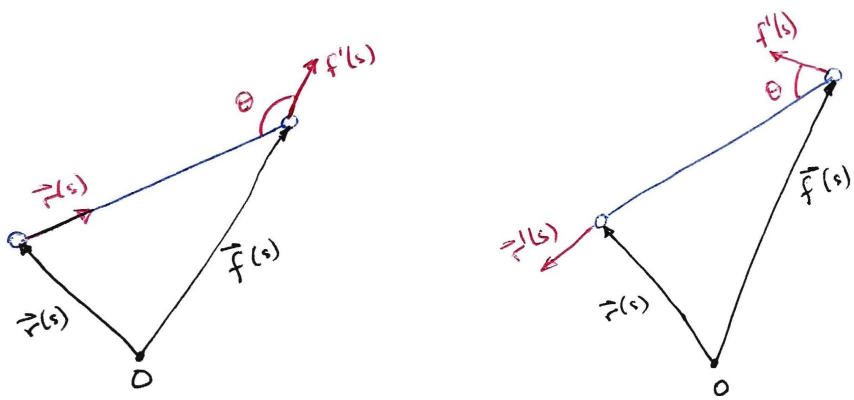
This is the basic equation we will use to analyze the path of the rear wheel of a scooter.

The Direction of $\mathbf{r}'$

Before going further we examine when $\mathbf{r}'$ points towards the front wheel and when it points away from the front wheel.

The movement of the back wheel is determined by the projection of $\mathbf{f}'$ on the line of the scooter. That projection is the same as $\mathbf{r}'$.

Thus, when the angle θ between the front wheel and the line of the scooter is obtuse, as in Figure C.16.30(a), $\mathbf{r}'$ points towards the front wheel. When θ is acute, the scooter backs up and $\mathbf{r}'$ points away from the front wheel, as shown in Figure C.16.30(b).



(a) (b)

Figure C.20.42: The direction of $\mathbf{r}'$ depends on the angle θ between the front wheel and the line of the scooter. (a) θ is obtuse, (b) θ is acute.

When the direction of $\mathbf{r}'$ abruptly shifts from pointing towards the front wheel to pointing away from the front wheel, the path of the rear wheel also abruptly changes, as shown in Figure C.16.31.

The path of the rear wheel is continuous but the unit tangent vector $\mathbf{r}'$ is not defined at the point where its direction suddenly shifts. The path is said to contain a "cusp" and the point at which $\mathbf{r}'(s)$ shifts direction by the angle π is the "vertex" of the cusp.

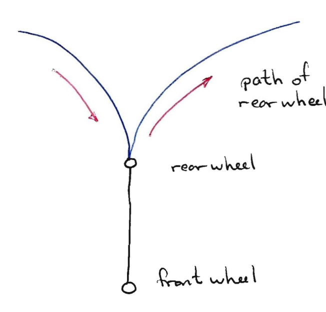


Figure C.20.43:

The Path of the Rear Wheel for a Short Scooter

Assume that the wheel base, a, is less than the radius of the circle, c, that, initially, θ is obtuse, and that r^2 is less than $c^2 - a^2$. Thus, $c^2 - a^2 - r^2$ is positive. (Exercise 1 shows the special significance of $c^2 - a^2$.)

We rewrite the equation $c^2 = a^2 + r^2 + 2rr'a$ in the form

$$\frac{-2rr'}{c^2 - a^2 - r^2} = \frac{-1}{a}. \tag{C.20.66}$$

Integration of both sides of (C.16.51) with respect to arc length s shows that there is a constant k such that

$$\ln(c^2 - a^2 - r^2) = \frac{-s}{a} + k,$$

hence

$$c^2 - a^2 - r^2 = e^k e^{-s/a}. \tag{C.20.67}$$

Equation (C.16.52) tells us that r^2 increases but remains less than $c^2 - a^2$, and approaches $c^2 - a^2$ as s increases. Thus the rear wheel traces a spiral path that gets arbitrarily close to the circle of radius $\sqrt{c^2 - a^2}$ and center O, as in Figure C.16.32.

The Path of the Rear Wheel for a Long Scooter

Assume that the wheel base is longer than the radius of the circle on which the front wheel moves, that is, $a > c$. Assume also that initially the scooter is moving forward, so we again have the equation

$$c^2 = a^2 + r^2 + 2rr'a. \tag{C.20.68}$$

The initial position is indicated in Figure C.16.33(a).

Now $c^2 - a^2 - r^2$ is negative, and we have

$$\frac{2rr'}{a^2 + r^2 - c^2} = \frac{-1}{a},$$

where the denominator on the left-hand side is positive. Thus there is a constant k such that

$$a^2 + r^2 - c^2 = e^k e^{-s/a}. \tag{C.20.69}$$

If s gets arbitrarily large, (C.16.54) implies that r^2 approaches $c^2 - a^2$. *But, $c^2 - a^2$ is negative, so this cannot happen.* Our assumption that (C.16.53) holds for all s must be wrong. Instead, there must be a cusp and the governing equation switches to

$$c^2 = a^2 + r^2 - 2arr'.$$

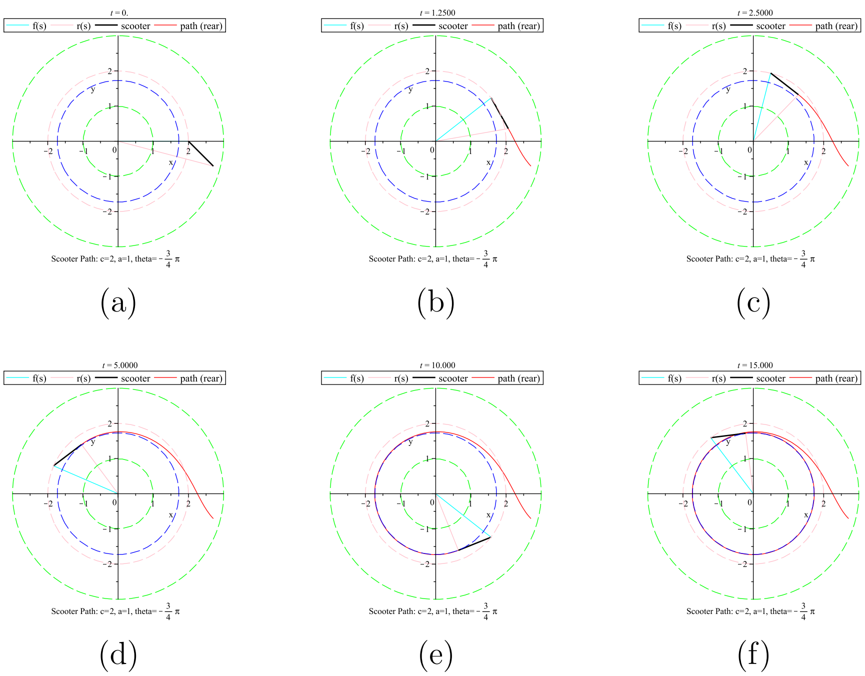


Figure C.20.44: The path of the rear wheel of a scooter with length $a = 1$, whose front wheel moves counter-clockwise around the circle with radius $c = 2$ from the point $(2, 0)$ with the line of the scooter at an angle $\theta = -3\pi/4$ with the front wheel. The snapshots are taken when (a) $s = 0$, (b) $s = 1.25$, (c) $s = 2.50$, (d) $s = 5.0$, (e) $s = 10.0$, and (f) $s = 15.0$. Because this is a short scooter ($a < c$), the rear wheel approaches the circle with radius $r = \sqrt{c^2 - a^2} = \sqrt{3}$. (Recall that s is the arclength of the rear wheel's path.)

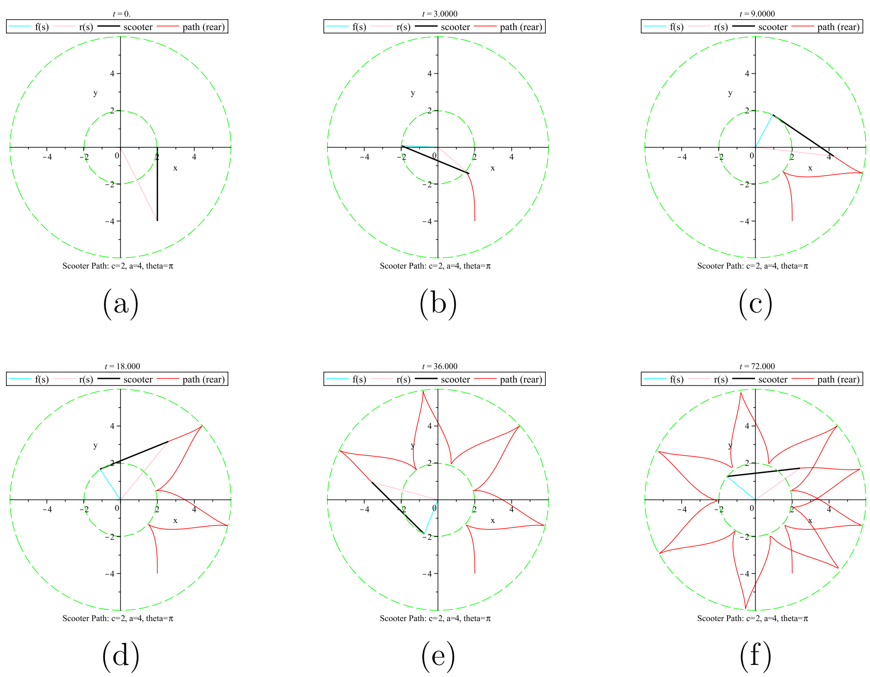


Figure C.20.45: The path of the rear wheel of a scooter with length $a = 4$, whose front wheel moves counter-clockwise around the circle with radius $c = 2$ from the point $(2, 0)$ with the line of the scooter at an angle $\theta = pi$ with the front wheel. The snapshots are taken when (a) $s = 0$, (b) $s = 3$, (c) $s = 9$, (d) $s = 18$, (e) $s = 36$, and (f) $s = 72$. Because this is a long scooter ($a > c$), the rear wheel travels along path that has cusps whenever $r = c+a$ and $f = |c-a|$. (Recall that s is the arclength of the rear wheel's path.)

This leads to the equation

$$a^2 + r^2 - c^2 = e^k e^{s/a}. \tag{C.20.70}$$

Equation (C.16.55) implies that as s increases r becomes arbitrarily large. However, r can never exceed $c + a$. So, another cusp must form.

It can be shown that the cusps occur when $r = a - c$ (assuming $a > c$) and $r = a + c$. At the vertex of a cusp, $\mathbf{r}'$ is not defined; it changes direction by π.

Figure C.16.33(b) shows the shape of the path of the rear wheel for a long scooter, $a > c$. (For $a > 2c$, that path remains outside the circle.)

EXERCISES

1.[R]

(a) Assume a and c are positive numbers with $c > a$ and that the front wheel moves on a circle of radius c. Show that when the front wheel moves along a circle of radius c the rear wheel could remain on a concentric circle of radius $b = \sqrt{(}c^2 - a^2)$.

(b) Draw the triangle whose sides are a, b, and c and explain why the result in (a) is plausible.

2.[M] We assumed in the case of the short scooter that initially $r^2 < c^2 - a^2$. Examine the case in which initially $r^2 > c^2 - a^2$. Again, assume that initially the scooter is not backing up.

3.[M] We assumed in the case of the short scooter that initially $r^2 < c^2 - a^2$ and that the scooter is not backing up. Investigate what happens when we assume that initially $r^2 < c^2 - a^2$ and the scooter is backing up.

(a) Draw such an initial position.

(b) Predict what will happen.

(c) Carry out the mathematics.

4.[R] It is a belief among many bicyclists that the rear tire wears out more slowly than the front tire. Decide whether their belief is justified. (Assume both tires support the same weight.)

5.[M] Show that if the path of the front wheel is a circle and a cusp forms in the path of the rear wheel, the scooter at that moment lies on a line through the center of the circle.

6.[M] In the case of the long scooter, $a > c$, do cusps always form, whatever the initial value of r and θ?

7.[C] Extend the analysis of the scooter to the case when $a = c$.

8.[C] Assume that the path of the front wheel is a straight line. For convenience, choose that line as the x-axis. Write $\mathbf{r}(s)$ as $x(s)\mathbf{i} + y(s)\mathbf{j}$.

(a) Show that $y(s) + y'(s)a = 0$.

(b) Deduce that there is a constant k such that $y(s) = ke^{-s/a}$. Thus the distance from the rear wheel to the x-axis "decays" exponentially.

Calculus is Everywhere # 21
The Wave in a Rope

We will develop what may be the most famous partial differential equation. In the CIE of the next chapter we will solve that equation and, then, use it in the final chapter to show how it helped Maxwell discover that light is an electrical-magnetic phenomenon.

As Morris Kline writes in *Mathematical Thought from Ancient to Modern Times*, "The first real success with partial differential equations came in renewed attacks on the vibrating string problem, typified by the violin string. The approximation that the vibrations are small was imposed by d'Alembert (1717-1783) in his papers of 1746."

Imagine shaking the end of a rope up and down gently, as in Figure C.21.46.

Figure C.21.46:

That motion starts a wave moving along the rope. The individual molecules in the rope move up and down, while the wave travels to the right. In the case of a sound wave, the wave travels at 700 miles per hour, but the air just vibrates back and forth. (When someone says "good morning" to us, we are not struck with a hurricane blast of wind.)

To develop the mathematics of the wave in a weightless rope, we begin with some simplifying assumptions. First, each molecule moves only up and down. Second, the distance each one moves is very small and the slope of the curve assumed by the rope remains close to zero. (Think of a violin string.)

At time t the vertical position of the molecule whose x-coordinate is x is $y = y(x,t)$, for it depends on both x and t. Consider a very short section of the rope at time t, shown as PQ in Figure C.21.47.

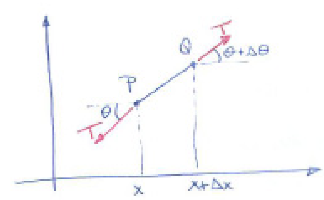


Figure C.21.47:

We assume that the tension T is the same throughout the rope. Apply Newton's Second Law, "force equals mass times acceleration," to the mass in PQ.

If the linear density of the rope is λ, the mass of the segment is λ times the length of the segment. Because we are assuming small displacements, we will approximate that length by Δx. The upward force exerted by the rope on the segment is $T\sin(\theta+\Delta\theta)$ and the downward force is $T\sin(\theta)$. The net vertical force is $T\sin(\theta+\Delta\theta) - T\sin(\theta)$. Thus

$$\underbrace{Tsin(\theta+\Delta\theta) - T\sin(\theta)}_{\text{net vertical force}} = \underbrace{\lambda\Delta x}_{\text{mass}} \underbrace{\frac{\partial^2 y}{\partial t^2}}_{\text{acceleration}} . \qquad \text{(C.21.71)}$$

(Because y is a function of x and t, we have a partial derivative, not an ordinary derivative.)

Next we express $\sin(\theta)$ and $\sin(\theta + \Delta\theta)$ in terms of the partial derivative $\partial y/\partial x$.

First of all, because θ is near 0, $\cos(\theta)$ is near 1. Thus $\sin(\theta)$ is approximately $\sin(\theta)/\cos(\theta) = \tan(\theta)$, the slope of the rope at time t above (or below) x, which is $\partial y/\partial x$ at x and t. Similarly, $\sin(\theta + \Delta\theta)$ is approximately $\partial y/\partial x$ at $x + \Delta x$ and t. So (C.21.71) is approximated by

$$T\frac{\partial y}{\partial x}(x + \Delta x, t) - T\frac{\partial y}{\partial x}(x, t) = \lambda \Delta x \frac{\partial^2 y}{\partial t^2}(x, t). \tag{C.21.72}$$

Dividing both sides of (C.21.72) by Δx gives

$$\frac{T\left(\frac{\partial y}{\partial x}(x + \Delta x, t) - \frac{\partial y}{\partial x}(x, t)\right)}{\Delta x} = \lambda \frac{\partial^2 y}{\partial t^2}(x, t). \tag{C.21.73}$$

Letting Δx in (C.21.73) approach 0, we obtain

$$T\frac{\partial^2 y}{\partial x^2}(x, t) = \lambda \frac{\partial^2 y}{\partial t^2}(x, t). \tag{C.21.74}$$

Since both T and λ are positive, we can write (C.21.74) in the form

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 y}{\partial t^2}. \tag{C.21.75}$$

This is the famous **wave equation**. It relates the acceleration of the molecule to the geometry of the curve; the latter is expressed by $\partial^2 y/\partial x^2$. Since we are assuming that the slope of the rope remains near 0, $\frac{\partial^2 y}{\partial x^2}$ is approximately

$$\frac{\frac{\partial^2 y}{\partial x^2}}{\left(\sqrt{1 + \left(\frac{\partial y}{\partial x}\right)^2}\right)^3}$$

which is the curvature at a given location and time. At the curvier part of the rope, the acceleration is greater.

As the CIE in the next chapter shows, the constant c turns out to be the velocity of the wave.

EXERCISES

1.[M] Figure C.21.48 shows a vibrating string whose ends are fixed at A and B. Assume that each part of the string moves parallel to the y-axis (a reasonable approximation of the vibrations are small.) Let $y = f(x, t)$ be the height of the string at the point with abscissa x at time t, as shown in the figure. In this case, the partial derivatives are denoted $\partial y/\partial x$ and $\partial y/\partial t$.

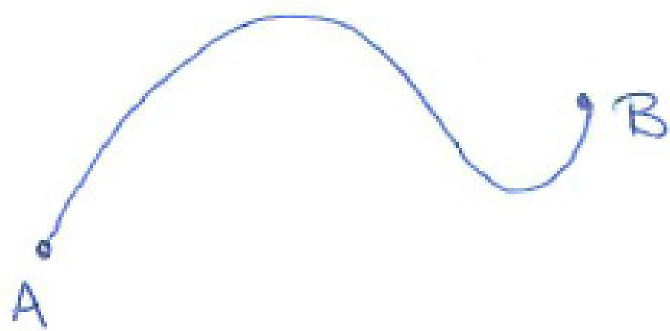


Figure C.21.48:

(a) What is the meaning of y_x?

(b) What is the meaning of y_t?

Calculus is Everywhere # 22
Solving the Wave Equation

In the *The Wave in a Rope* Calculus is Everywhere in the previous chapter we encountered the partial differential equation

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 y}{\partial t^2}. \qquad \text{(C.22.76)}$$

Now we will solve this equation to find y as a function of x and t. First, we solve some simpler equations, which will help us solve (C.22.76).

EXAMPLE 2 Let $u(x, y)$ satisfy the equation $\partial u/\partial x = 0$. Find the form of $u(x, y)$.
SOLUTION Since $\partial u/\partial x$ is 0, $u(x, y)$, for a fixed value of y, is constant. Thus, $u(x, y)$ depends only on y, and can be written in the form $h(y)$ for some function h of a single variable.

On the other hand, any function $u(x, y)$ that can be written in the form $h(y)$ has the property that $\partial u/\partial x = 0$ is any function that can be written as a function of y alone. ⋄

EXAMPLE 3 Let $u(x, y)$ satisfy

$$\frac{\partial^2 u}{\partial x \partial y} = 0. \qquad \text{(C.22.77)}$$

Find the form of $u(x, y)$.
SOLUTION We know that

$$\frac{\partial\left(\frac{\partial u}{\partial y}\right)}{\partial x} = \frac{\partial^2 u}{\partial x \partial y} = 0.$$

By Example 2,

$$\frac{\partial u}{\partial y} = h(y) \qquad \text{for some function } h(y).$$

By the Fundamental Theorem of Calculus, for any number b,

$$u(x, b) - u(x, 0) = \int_0^b \frac{\partial u}{\partial y} dy = \int_0^b h(y) dy.$$

Let H be an antiderivative of h. Then

$$u(x,b) - u(x,0) = H(b) - H(0).$$

Replacing b by y shows that

$$u(x,y) = u(x,0) + H(y) - H(0).$$

That tells us that $u(x,y)$ can be expressed as the sum of a function of x and a function of y,

$$u(x,y) = f(x) + g(y). \tag{C.22.78}$$

◇

We will solve the wave equation (C.22.76) by using a suitable change of variables that transforms that equation into the one solved in Example 3.

The new variables are

$$p = x + ct \qquad \text{and} \qquad q = x - ct.$$

One could solve these equations and express x and t as functions of p and q. We will apply the chain rule, where y is a function of p and q and p and q are functions of x and t, as indicated in Figure C.22.49. Thus $y(x,t) = u(p,q)$.

$x = \frac{1}{2}(p+q)$ and $t = \frac{1}{2c}(p-q)$.

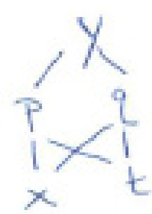


Figure C.22.49:

Keeping in mind that

$$\frac{\partial p}{\partial x} = 1, \qquad \frac{\partial p}{\partial t} = c, \qquad \frac{\partial q}{\partial x} = 1, \qquad \text{and } \frac{\partial q}{\partial t} = -c,$$

we have

$$\frac{\partial y}{\partial x} = \frac{\partial u}{\partial p}\frac{\partial p}{\partial x} + \frac{\partial u}{\partial q}\frac{\partial q}{\partial x} = \frac{\partial u}{\partial p} + \frac{\partial u}{\partial q}.$$

Then

$$\begin{aligned}
\frac{\partial^2 y}{\partial x^2} &= \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial p} + \frac{\partial u}{\partial q}\right) \\
&= \frac{\partial}{\partial p}\left(\frac{\partial u}{\partial p} + \frac{\partial u}{\partial q}\right)\frac{\partial p}{\partial x} + \frac{\partial}{\partial q}\left(\frac{\partial u}{\partial p} + \frac{\partial u}{\partial q}\right)\frac{\partial q}{\partial x} \\
&= \left(\frac{\partial^2 u}{\partial p^2} + \frac{\partial^2 u}{\partial p \partial q}\right)\cdot 1 + \left(\frac{\partial^2 u}{\partial q \partial p} + \frac{\partial^2 u}{\partial^2 q}\right)\cdot 1.
\end{aligned}$$

Thus

$$\frac{\partial^2 y}{\partial x^2} = \frac{\partial^2 u}{\partial p^2} + 2\frac{\partial^2 u}{\partial p \partial q} + \frac{\partial^2 u}{\partial q^2}. \tag{C.22.79}$$

A similar calculation shows that

$$\frac{\partial^2 y}{\partial t^2} = c^2\left(\frac{\partial^2 u}{\partial p^2} - 2\frac{\partial^2 u}{\partial p \partial q} + \frac{\partial^2 u}{\partial q^2}\right). \tag{C.22.80}$$

Substituting (C.22.79) and (C.22.80) in (C.22.76) leads to

$$\frac{\partial^2 u}{\partial p^2} + 2\frac{\partial^2 u}{\partial p \partial q} + \frac{\partial^2 u}{\partial q^2}. = \frac{1}{c^2}\left(c^2\right)\left(\frac{\partial^2 u}{\partial p^2} - 2\frac{\partial^2 u}{\partial p \partial q} + \frac{\partial^2 u}{\partial q^2}\right),$$

which reduces to

$$4\frac{\partial^2 u}{\partial p \partial q} = 0.$$

By Example 3, there are function $f(p)$ and $g(q)$ such that

$$y(x,t) = u(p,q) = f(p) + g(q).$$

or

$$y(x,t) = f(x+ct) + g(x-ct). \tag{C.22.81}$$

The expression (C.22.81) is the most general solution of the wave equation (C.22.76).

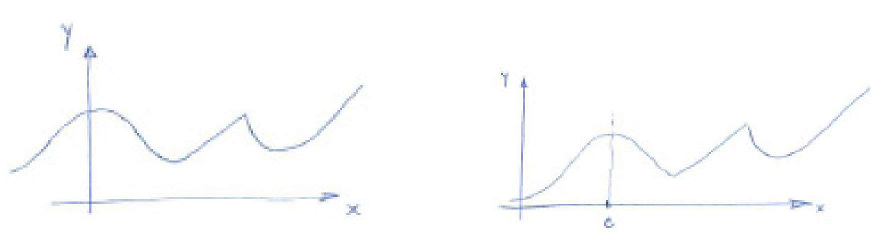
What does a solution (C.22.81) look like? What does the constant c tell us? To answer these questions, consider just

$$y(x,t) = g(x-ct). \tag{C.22.82}$$

Here t represents time. For each value of t, $y(x,t) = g(x-ct)$ is simply a function of x and we can graph it in the xy plane. For $t=0$, (C.22.82) becomes

$$y(x,0) = g(x).$$

That is just the graph of $y = g(x)$, whatever g is, as shown in Figure C.22.50(a).



(a) (b)

Figure C.22.50: (a) $t=0$, (b) $t=1$.

Now consider $y(x,t)$ when $t=1$, which we may think of as "one unit of time later." Then

$$y = y(x,1) = g(x - c \cdot 1) = g(x-c).$$

The value of $y(x, 1)$ is the same as the value of g at $x - c$, c units to the left of x. So the graph at $t = 1$ is the graph of f in Figure C.22.50(a) shifted to the *right* c units, as in Figure C.22.50(b).

As t increases, the initial "wave" shown in Figure C.22.50(a) moves further to the right at the constant speed, c. Thus c tells us the velocity of the moving wave. That fact will play a role in Maxwell's prediction that electro-magnetic waves travel at the speed of light, as we will see in the Calculus is Everywhere at the end of Chapter **??**.

EXERCISES

1.[R] Which functions $u(x, y)$ have both $\partial u/\partial x$ and $\partial u/\partial y$ equal to 0 for all x and y?

2.[R] Let $u(x, y)$ satisfy the equation $\partial^2 u/\partial x^2 = 0$. Find the form of $u(x, y)$.

3.[R] Show that any function of the form (C.22.78) satisfies equation (C.22.77).

4.[R] Verify that any function of the form (C.22.81) satisfies the wave equation.

5.[M] We interpreted $y(x, t) = g(x - ct)$ as the description of a wave moving with speed c to the right. Interpret the equation $y(x, t) = f(x + ct)$.

6.[M] Let k be a positive constant.

(a) What are the solutions to the equation

$$\frac{\partial^2 y}{\partial x^2} = k\frac{\partial^2 y}{\partial t^2}?$$

(b) What is the speed of the "waves"?

Calculus is Everywhere # 23
How Maxwell Did It

In a letter to his cousin, Charles Cay, dated January 5, 1965, Maxwell wrote:

> I have also a paper afloat containing an electromagnetic theory of light, which, till I am convinced to the contrary, I hold to be great guns. [Everitt, F., *James Clerk Maxwell: a force for physics*, Physics World, Dec 2006, `http://physicsworld.com/cws/article/print/26527`]

It indeed was "great guns," for out of his theory has come countless inventions, such as television, cell phones, and remote garage door openers. In a dazzling feat of imagination, Maxwell predicted that electrical phenomena create waves, that light is one such phenomenon, and that the waves travel at the speed of light, in a vacuum.

In this section we will see how those predictions came out of the four equations (I'), (II'), (III'), and (IV') in Section **??**.

First, we take a closer look at the dimensions of the constants ε_0 and μ_0 that appear in (IV'),

$$\frac{1}{\mu_0\varepsilon_0}\nabla\times\mathbf{B} = \frac{\mathcal{J}}{\varepsilon_0}.$$

The constant ε_0 makes its appearance in the equation

$$\text{Force} = F = \frac{1}{4\pi\varepsilon_0}\frac{qq_0}{r^2}. \qquad \text{(C.23.83)}$$

Since the force F is "mass times acceleration" its dimensions are

$$\text{mass}\cdot\frac{\text{length}}{\text{time}^2},$$

or, in symbols

$$m\frac{L}{T^2}.$$

The number 4π is a pure number, without any physical dimension.

The quantity qq_0 has the dimensions of "charge squared," q^2, and R^2 has dimensions L^2, where L denotes length.

Solving (C.23.83) for ε_0, we find the dimensions of ε_0. Since

$$\varepsilon_0 = \frac{q^2}{4\pi F r^2},$$

its dimensions are

$$\left(\frac{T^2}{mL}\right)\left(\frac{q^2}{L^2}\right) = \frac{T^2q^2}{mL^3}.$$

To figure out the dimensions of μ_0, we will use its appearance in calculating the force between two wires of length L each carrying a current I in the same direction and separated by a distance R. (Each generates a magnetic field that draws the other towards it.) The equation that describes that force is

$$\mu_0 = \frac{2\pi RF}{I^2L}.$$

Since R has the dimensions of length L and F has dimensions mL/T^2, the numerator has dimensions mL^2/T^2. The current I is "charge q per second," so I^2 has dimensions q^2/T^2. The dimension of the denominator is, therefore,

$$\frac{q^2L}{T^2}.$$

Hence μ_0 has the dimension

$$\frac{mL^2}{T^2} \cdot \frac{T^2}{q^2L} = \frac{mL}{q^2}.$$

The dimension of the product $\mu_0\varepsilon_0$ is therefore

$$\frac{mL}{q^2} \cdot \frac{T^2q^2}{mL^3} = \frac{T^2}{L^2}.$$

The dimension of $1/\mu_0\varepsilon_0$, the same as the square of speed. In short, $1/\sqrt{\mu_0\varepsilon_0}$ has the dimension of speed, "length divided by time."

Now we are ready to do the calculations leading to the prediction of waves traveling at the speed of light. We will use the equations (I'), (II'), (III'), and (IV'), as stated on page **??**, where the fields **B** and **E** vary with time. However, we assume there is no current, so $\mathcal{J} = /$. We also assume that there is no charge q.

Recall the equation (IV')

$$\nabla \times \mathbf{B} = \mu_0\varepsilon_0\frac{\partial \mathbf{E}}{\partial t}.$$

Differentiating this equation with respect to time t we obtain

$$\frac{\partial}{\partial t}(\nabla \times \mathbf{B}) = \mu_0\varepsilon_0\frac{\partial^2 \mathbf{E}}{\partial t^2}. \tag{C.23.84}$$

As is easy to check, the operator $\frac{\partial}{\partial t}$ can be moved past the $\nabla\times$ to operate directly on **B**. Thus (C.23.84) becomes

$$\nabla \times \frac{\partial \mathbf{B}}{\partial t} = \mu_0\varepsilon_0\frac{\partial^2 \mathbf{E}}{\partial t^2}. \tag{C.23.85}$$

Recall the equation (II')

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}.$$

Taking the curl of both sides of this equation leads to

$$\nabla\left(-\nabla \times \mathbf{E}\right) = \nabla \times \frac{\partial \mathbf{B}}{\partial t}. \qquad \text{(C.23.86)}$$

Combining (C.23.85) and (C.23.86) gives us an equation that involves **E** alone:

$$\nabla \times \left(-\nabla \times \mathbf{E}\right) = \mu_0\varepsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}. \qquad \text{(C.23.87)}$$

An identity concerning "the curl of the curl," which tells us that

$$\nabla \times \left(\nabla \times \mathbf{E}\right) = \nabla\left(\nabla \cdot \mathbf{E}\right) - \left(\nabla \cdot \nabla\right)\mathbf{E}. \qquad \text{(C.23.88)}$$

But $\nabla \cdot \mathbf{E} = 0$ is one of the four assumptions, namely (I), on the electromagnetic fields. By (C.23.87) and (C.23.88), we arrive at

$$\begin{aligned} \left(\nabla \cdot \nabla\right)\mathbf{E} &= \mu_0\varepsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} \\ \text{or} \qquad \frac{\partial^2 \mathbf{E}}{\partial t^2} - \frac{1}{\mu_0\varepsilon_0}\nabla^2 \mathbf{E} &= \mathbf{0}. \end{aligned} \qquad \text{(C.23.89)}$$

The expression ∇^2 in (C.23.89) is short for

$$\begin{aligned} \nabla \cdot \nabla &= \left(\frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j} + \frac{\partial}{\partial z}\mathbf{k}\right) \cdot \left(\frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j} + \frac{\partial}{\partial z}\mathbf{k}\right) \\ &= \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \end{aligned} \qquad \text{(C.23.90)}$$

In $(\nabla \cdot \nabla)\mathbf{E}$ we apply (C.23.90) to each of the three components of **E**. Thus $\nabla^2\mathbf{E}$ is a vector. So is $\partial^2\mathbf{E}/\partial t^2$ and (C.23.90) makes sense.

For the sake of simplicity, consider the case in which **E** has only an x-component, which depends only on x and t, $\mathbf{E}(x, y, z, t) = E(x, t)\mathbf{i}$, where E is a scalar function. Then (C.23.90) becomes

$$\frac{\partial^2}{\partial t^2}E(x,t)\mathbf{i} - \frac{1}{\mu_0\varepsilon_0}\left(\frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial y^2} + \frac{\partial^2 E}{\partial z^2}\right)\mathbf{i} = \mathbf{0},$$

from which it follows

$$\frac{\partial^2}{\partial t^2}E(x,t) - \frac{1}{\mu_0\varepsilon_0}\frac{\partial^2 E}{\partial x^2} = 0. \qquad \text{(C.23.91)}$$

Multiply (C.23.91) by $-\mu_0\varepsilon_0$ to obtain

$$\frac{\partial^2 E}{\partial x^2} - \mu_0\varepsilon_0\frac{\partial^2 E}{\partial t^2} = 0.$$

This looks like the wave equation (see (**??**) on page **??**). The solutions are waves traveling with speed $1/\sqrt{\mu_0\varepsilon_0}$.

Maxwell then compares $\sqrt{\mu_0\varepsilon_0}$ with the velocity of light:

> In the following table, the principal results of direct observation of the velocity of light, are compared with the principal results of the comparison of electrical units $(1/\sqrt{\mu_0 v_0})$.

Velocity of light (meters per second)		Ratio of electrical units	
Fizeau	314,000,000	Weber	310,740,000
Sun's Parallax	308,000,000	Maxwell	288,000,000
Foucault	298,360,000	Thomson	282,000,000

Table C.23.2:

> It is magnificent that the velocity of light and the ratio of the units are quantities of the same order of magnitude. Neither of them can be said to be determined as yet with such a degree of accuracy as to enable us to assert that the one is greater or less than the other. It is to be hoped that, by further experiment, the relation between the magnitude of the two quantities may be more accurately determined.
>
> In the meantime our theory, which asserts that these two quantities are equal, and assigns a physical reason for this equality, is certainly not contradicted by the comparison of these results such as they are. [reference?]

On this basis Maxwell concluded that light is an "electromagnetic disturbance" and predicted the existence of other electromagnetic waves. In 1887, eight years after Maxwell's death, Heinrich Hertz produced the predicted waves, whose frequency placed them outside what the eye can see.

By 1890 experiments had confirmed Maxwell's conjecture. First of all, experiments gave the velocity of light as 299,766,000 meters per second and $\sqrt{1/\mu_0\varepsilon_0}$ as 299,550,000 meters per second.

Newton, in his *Principia* of 1687 related gravity on earth with gravity in the heavens. Benjamin Franklin, with his kite experiments showed that lightning was simply an electric phenomenon. From then through the early nineteenth century, Faraday, ???, . . . showed that electricity and magnetism were inseparable. Then Maxwell joined them both to light. Einstein, in 1905(?), also by a mathematical argument, hypothesized that mass and energy were related, by his equation $E = mc^2$.

Calculus is Everywhere # 24
Heating and Cooling

Engineers who design a car radiator or a home air conditioner are interested in the distribution of temperature of a fin attached to a tube. We present one of the mathematical tools they use. Incidentally, the example shows how Green's Theorem is applied in practice.

A plane region $\mathcal{A}$ with boundary curve C is occupied by a sheet of metal. By various heating and cooling devices, the temperature along the border is held constant, independent of time. Assume that the temperature in $\mathcal{A}$ eventually stabilizes. This steady-state temperature at point P in $\mathcal{A}$ is denoted $T(P)$. What does that imply about the function $T(x, y)$?

First of all, heat tends to flow "from high to low temperatures," that is, in the direction of $-\nabla T$. According to Fourier's law, flow is proportional to the conductivity of the material k (a positive constant) and the magnitude of the gradient $\|\nabla T\|$. Thus

$$\oint_C (-k\nabla T) \cdot \mathbf{n} ds$$

measures the rate of heat loss across C.

Since the temperature in the metal is at a steady state, the heat in the region bounded by C remains constant. Thus

$$\oint_C (-k\nabla T) \cdot \mathbf{n} ds = 0.$$

Now, Green's theorem then tells us that

$$\int_{\mathcal{A}} \nabla \cdot (-k\nabla T) dA = 0$$

for any region $\mathcal{A}$ in the metal plate. Since $\nabla \cdot \nabla T$ is the Laplacian of T and k is not 0, we conclude that

$$\int_{\mathcal{A}} \left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} \right) dA = 0. \tag{C.24.92}$$

By the "zero integrals" theorem, the integrand must be 0 throughout $\mathcal{A}$,

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} = 0.$$

This is an important step, since it reduces the study of the temperature distribution to solving a partial differential equation.

The expression
$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2},$$
which is $\nabla \cdot \nabla T$, the divergence of the gradient of T, is called the **Laplacian** of T. If T is a function of x, y, and z, then its Laplacian has one more summand, $\partial^2 T/\partial z^2$. However, the vector notation remains the same, $\nabla \cdot \nabla T$. Even more compactly, it is often reduced to $\nabla^2 T$. Note that in spite of the vector notation, the Laplacian of a scalar field is again a scalar field. A function whose Laplacian is 0 is called "harmonic."

EXERCISES