Section 2.8

(3/19/08)

Linear approximations and differentials

Overview: *In this section we return to the main concept of this chapter: the approximations of graphs by tangent lines. We discuss additional examples of tangent-line approximations and show how tangent line estimates of errors can be calculated using the notation of differentials. We finish the section with a description of some of the key steps in the history of the derivative.*

Topics:

- ***Tangent-line approximations***
- ***Differentials***
- ***Error estimates***
- ***A closer look at linear approximations***
- ***History of the derivative***

Tangent-line approximations

The tangent line $y = f(a) + f'(a)(x - a)$ to $y = f(x)$ at $x = a$ approximates the graph near $x = a$ and consequently,

$$f(x) \approx f(a) + f'(a)(x - a) \quad \text{for } x \approx a. \tag{1}$$

The function $y = f(a) + f'(a)(x - a)$ here, whose graph is the tangent line, is called the LINEAR APPROXIMATION or LOCAL LINEARIZATION of $y = f(x)$ at $x = a$.

Example 1 Find the linear approximation $y = L(x)$ of $f(x) = 15 - 4x + x^3$ at $x = 2$.

SOLUTION By formula **(1)** with a = 2, the linear approximation is $L(x) = f(2) + f'(2)(x - 2)$. We calculate $f(2) = 15 - 4(2) + 2^3 = 15 - 8 + 8 = 15$. Then we use the differentiation formulas from Section 2.4 to write $f'(x) = \dfrac{d}{dx}(5 - 4x + x^3) = -4 + 3x^2$, and calculate $f'(2) = -4 + 3(-2)^2 = -4 + 12 = 8$. The linear approximation is $L(x) = 25 + 8(x - 2)$. The graphs of $y = f(x)$ and of its approximation $y = L(x)$ are in Figure 1. Notice that $L(x)$ approximates $f(x)$ well for x near 2 but not for x far from 2. □

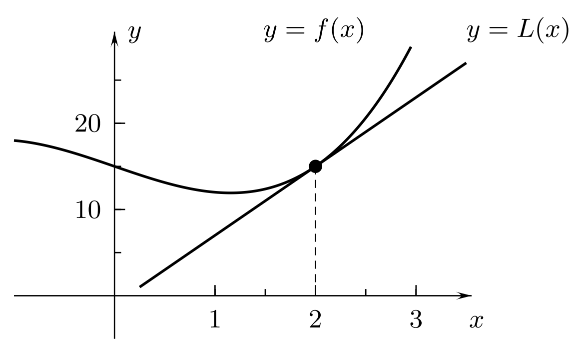


FIGURE 1

Example 2 One hour after leaving Toulouse, France, on a test flight, the French/British supersonic passenger jet Concord had traveled 975 miles and was flying 1520 miles per hour (twice the speed of sound).[(1)] Approximately how far was it from Toulouse six minutes (0.1 hours) later?

SOLUTION We let $s = s(t)$ be the Concord's distance from Toulouse t hours after take off. We are given that $s(1) = 975$ miles and $s'(1) = 1520$ miles per hour, and we want to calculate the approximate value of $s(1.1)$. By **(1)** with $s(t)$ in place of $f(x)$ and $a = 1$, the linear approximation of $s = s(t)$ at $t = 1$ is

$$L(t) = 950 + 1520(t - 1).$$

Therefore, $s(1.1) \approx L(1.1) = 975 + 1520(1.1 - 1) = 1127$ miles. The plane was approximately 1127 miles from Toulouse one hour and six minutes after take-off. □

Differentials

For a function $y = f(x)$ with a derivative at $x = a$, the DIFFERENTIALS dx and df represent corresponding changes in x and y on the tangent line to the graph at $x = a$ (Figure 2).[†] Consequently, if dx is not zero, then dx and df are the run and rise between two points on the tangent line, whose slope is $f'(a)$, so that

$$f'(a) = \frac{\text{Rise}}{\text{Run}} = \frac{df}{dx}. \qquad \textbf{(2)}$$

Notice that even though this equation looks like a statement relating Leibniz and prime notation for the derivative, here df/dx is an actual ratio of numbers rather than just a symbol for the derivative.

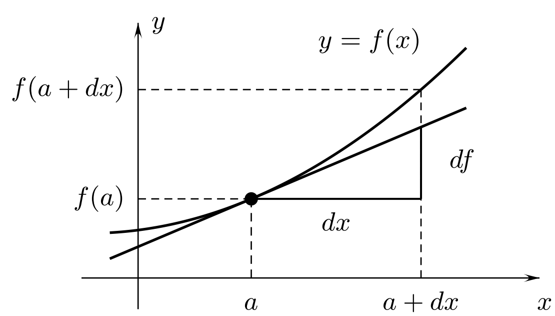


FIGURE 2

Multiplying both sides of equation **(2)** by dx gives a formula for df in terms of dx:

$$df = f'(a)\,dx. \qquad \textbf{(3)}$$

Since dx and df are the run and rise between two points on the tangent line, as in Figure 2, formula **(3)** expresses the approximation of the graph by the tangent line. Also, since a is a constant, we can view dx as the independent variable and df as the dependent variable in this equation.

Example 3 Express df in terms of dx for $f(x) = x^3$ at $x = 2$.

SOLUTION $f'(x) = 3x^2$ and $f'(2) = 3(2)^2 = 12$ so at $x = 2$, $df = 12dx$ by **(3)**. □

[(1)] Data from "Test flights of supersonic aircraft" by J. F. Renaudie, *Flight Test Instrumentation*, AGARD Conference Proceedings No. 32, London: Technical Editing and Reproduction Ltd., 1967, p. 39.

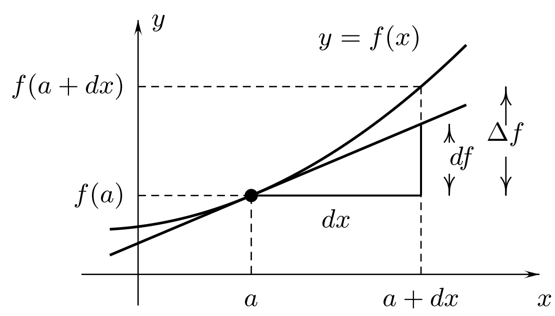
[†] The symbols dx and df were used in calculus of the seventeenth and eighteenth centuries to designate "infinitesimal" changes in x and f (see the Historical Notes in Section 42.8) and are used in this way by some contemporary mathematicians and other scientists.

Error estimates

If we let Δf denote the change $f(a+dx) - f(a)$ in the value of $f(x)$ from $x = a$ to $x = a + dx$, as in Figure 3, then for small dx,

$$\Delta f = f(a+dx) - f(a) \approx df. \tag{4}$$

FIGURE 3



This statement is convenient for expressing error estimates based on tangent line approximations. Suppose a quantity x is measured to be a and a formula is used to calculate a related quantity $f(a)$. We let $a + dx$ denote the (unknown) exact value of x. Then the exact value of $f(x)$ is $f(a+dx)$ and the error in the value of $f(x)$ that results from using a is $|\Delta f| = |f(a+dx) - f(a)|$. Therefore,

$$[\text{The error in } f(x)] = |\Delta f| = |f(a+dx) - f(a)| \approx |df| = |f'(a)dx|. \tag{5}$$

The next example shows how this notation is used.

Example 4 The width w of a cube is measured to be 10 centimeters with an error ≤ 0.03 centimeters. The value $w = 10$ is used to find the approximate volume $V = 10^3 = 1000$ cubic centimeters. Use differentials to estimate the maximum error in the calculated volume.

SOLUTION The volume of a cube of width w is $V = w^3$ and $V' = 3w^2$. Therefore the differentials dV and dw are related by the equation, $dV = 3w^2\,dw$. Setting $w = 10$ gives $dV = 300\,dw$. The error from using the measured $w = 10$ instead of the (actual, unknown) width $10 + dw$ is approximately $|dV| = 300|dw|$. Since we are given that $|dw| \leq 0.03$, we conclude that

$$\begin{aligned}[\text{The error in } V] &\approx |dV| = |300\,dw| \\ &\leq 300(0.03) = 9 \text{ cubic centimeters.}\end{aligned}$$

The maximum error is approximately 9 cubic centimeters. □

A closer look at linear approximations

The proof of the general Chain Rule for differentiating composite functions in Section 4.1 will require the following result concerning linear approximations.

Lemma 1 *Supose that $y = f(x)$ has a derivative at $x = a$ and $L(x) = f(a) + f'(a)(x - a)$ is its linear approximation at $x = a$. Then*

$$f(x) = L(x) + E(x)(x - a) \tag{6}$$

where $y = E(x)$ is defined in an open interval containing a and $E(x) \to 0$ as $x \to a$.

Since $E(x) \to 0$ as $x \to a$ in **(6)**, we describe equation **(6)** by saying that $f(x) - L(x)$ "tends to zero faster than $x - a$ as $x \to a$."

Proof of Lemma 1: According to Definition 1 of the derivative from Section 2.3,

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}. \tag{7}$$

Let I be an open interval in the domain of f that contains a. For $x \neq a$ in I, we let $E(x)$ denote the "error"

$$E(x) = \frac{f(x) - f(a)}{x - a} - f'(a)$$

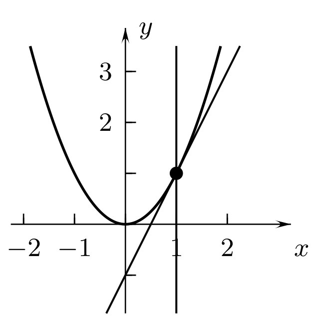
that is made when $f'(a)$ is approximated by the difference quotient in **(7)**. Multiplying both sides of this equation by $x - a$ and solving for $f(x)$ gives for $x \neq a$ in I,

$$f(x) = f(a) + f'(a)(x - a) + E(x)(x - a). \tag{8}$$

We define $E(a)$ to be 0, so that **(8)** is valid for all x in I. This equation then gives **(6)** since $L(x) = f(x) + f'(a)(x - a)$. Moreover, $E(x) \to 0$ as $x \to a$ because of **(7)**. **QED**

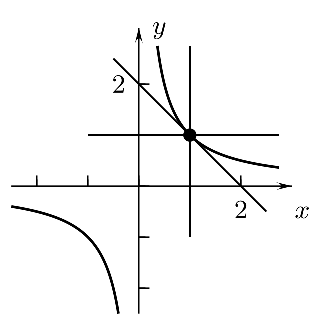
History of the derivative

Euclid (ca. 300 BC) stated that a line is tangent to a circle if it intersects the curve at one and only one point.[2] He also used this definition for ellipses, but it could not be applied without modification to the other conic sections—parabolas and hyperbolas—since there are two such lines at each point on a parabola and three at each point on a hyperbola. For the parabola $y = x^2$ in Figure 4, the two lines are the tangent line and the line parallel to the y-axis. For the hyperbola $y = 1/x$ in Figure 5, the three lines are the tangent line and the lines parallel to the x- and y-axes.



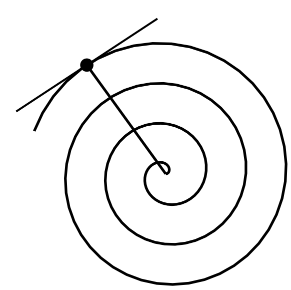
$y = x^2$

FIGURE 4



$y = 1/x$

FIGURE 5



An Archimedean spiral

FIGURE 6

---

[2] *The Thirteen Books of Euclid* by T. Heath, Vol. 2, New York, NY: Dover Publications, Inc., 1956, p. 1.

The definition of tangent line that Apollonius (ca. 225 BC) used in his treatise *Conics* could be applied to any conic section. He defined a tangent line to be a line such that no other straight line could fall between it and the curve.(3)

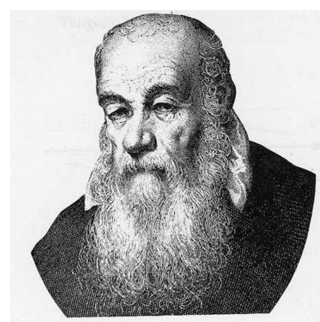
Another early approach to the study of tangent lines was based on the principle that an object's motion at each point as it moves along a curve is in the direction of the tangent line at that point. The Greek scientist and mathematician Archimedes (ca. 287–222 BC) apparently used this idea to find the lines tangent to a spiral traversed by a point that moves along a ray at a constant speed while the ray rotates about its end at a constant speed (Figure 6). Such a spiral is now known as an Archimedean spiral. This approach to tangent lines was also used in the 1640's by a number of scholars, including an Italian physicist Evangelista Torricelli (1608–1647). They used this approach to determine tangent lines to a variety of curves, including those given in modern notation by the equations $y = ax^n$ with various rational exponents n.

In the 1630s, Descartes devised the METHOD OF EQUAL ROOTS for finding lines perpendicular to certain curves by studying the solutions of polynomial equations. This approach was later used extensively to find tangent lines.

The need for mathematical descriptions of velocity contributed to the development of the concept of the derivative. The philosopher Aristotle (384–322 BC) believed that an object falling in air would have a constant velocity proportional to its weight. This view was commonly held through the Middle Ages, although it was challenged by various scholars who observed that the velocity is not constant but increases with time. In the fourteenth century, scholastic philosophers at Merton College, Oxford, studied motion with constant acceleration and deduced what is now known as the *Merton rule:* an object with constant acceleration travels the same distance as it would if it had constant velocity equal to the average of its initial and final velocities. In the seventeenth century, Galileo Galilei (1564–1642) and others discovered that in a void, all falling objects have the same constant acceleration, so that their motion may be determined by using the Merton rule. This result, however did not resolve the question of motion subject to air resistance.

Evangelista Torricelli

(1608–1647)



Galileo Galilei

(1564–1642)

The approaches to the study of tangent lines and velocity that led to the modern concept of the derivative were based on one notion or another of the "infinitely small." Some mathematicians considered a curve to be composed of infinitely short line segments and contended that the tangent line at a point is the extension of the infinitely short line segment at that point. Others stated that tangent lines are the lines through pairs of infinitely close points on the curve. Variable velocities were studied by considering the infinitely small distance that an object travels in an infinitely short period of time.

(3) *Apollonius of Perga* by T. Heath, Cambridge: W. Heffer & Sons, 1961, p. 22.

The algebraic calculations based on these ideas involved INFINITESIMALS—quantities that are supposed to be infinitely small and yet not zero. Questions concerning the existence of the infinitely small and the infinitely large had been actively debated by philosophers in previous centuries, and the notion of infinitely small nonzero numbers was questioned to some extent by those who used it. However, the concept of infinitesimals was accepted because of its effectiveness as a mathematical tool.

Fermat used infinitesimals to find tangent lines to curves as early as 1636, and by 1660 a number of others, including the Scottish mathematician James Gregory (1638–75) and the English theologian Isaac Barrow (1630–1677), had done similar calculations.

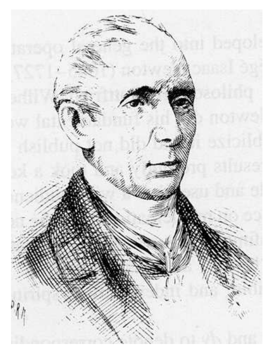
Techniques using infinitesimals were developed into the general operation now known as differentiation by Barrow's protégé Isaac Newton (1642–1727) in the 1660's and independently by the German philosopher Gottfried Wilhelm Leibniz (1646–1716) in the 1670's. Although Newton did his fundamental work earlier than Liebniz, he did not publish his results until after 1700. Leibniz published his results promptly and took a keen interest in making his techniques comprehensible and useful to a wide audience. Consequently, he exerted a much greater influence on mathematicians of the next 150 years than did Newton. Leibniz's use of infinitesimals was perhaps related to his philosophy, which contended that the "best of all possible worlds" in which we live consists of infinitesimal, indivisible, and indestructible spiritual atoms called "monads."

Isaac Newton

(1642–1727)

Gottfried Leibniz

(1646–1716)



Augustin Cauchy

(1789–1857)

Leibniz used what he called "differentials" dx and dy to denote corresponding infinitesimal changes in quantities x and y related by an equation. If the equation were

$$y = x^2 \tag{9}$$

he would write the same equation with x replaced by $x + dx$ and with y replaced by $y + dy$ to obtain $y + dy = (x + dx)^2$. Expanding the square and subtracting the original equation **(9)** would give

$$dy = 2x(dx) + (dx)^2. \tag{10}$$

Leibniz would then drop the square $(dx)^2$ of the differential from equation **(10)**. He would justify this by stating that the square of the differential is negligible in comparison with the differentials themselves (as a point is negligible in comparison with a line, he once said). Without the square, the equation would be

$$dy = 2x(dx) \tag{11}$$

which corresponds to the modern formula $dy/dx = 2x$.

To compare this use of infinitesimals with the modern reasoning based on limits, we let Δx denote a nonzero (and noninfinitesimal) change in x and let Δy be the corresponding change in y. By the same algebraic calculations as above, we obtain the equation $\Delta y = 2x(\Delta x) + (\Delta x)^2$. Then, instead of throwing away the $(\Delta x)^2$ term, we divide the equation by Δx and take the limit as $\Delta x \to 0$ to obtain

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} (2x + \Delta x) = 2x.$$

The obvious criticism of the calculations with infinitesimals which led from equation **(10)** to equation **(11)** is that the differential dx is a nonzero quantity, and yet its square, which is dropped from the calculation, is somehow zero. The most famous attack on the use of infinitesimals was published in 1734 by an English bishop, George Berkeley, in a treatise titled *The analyst, or a discourse addressed to an infidel mathematician, wherein it is examined whether the object, principles, and inferences of the modern analyst are more distinctly conceived or more evidently deduced than religious mysteries and points of faith.* (The "infidel mathematician" was supposedly the astronomer Edmund Halley.) In this treatise, Berkeley ridiculed infinitesimals as "the ghosts of departed quantities" and discussed the apparent contradictions involved in their use.

At one time or another both Newton and Leibniz gave presentations of their calculus that were fairly close to the modern treatments involving limits. Newton did so, for example, in his theory of prime and ultimate ratios, and Leibniz in a reply to a critic, Nieuwentijt.

Nevertheless, Newton and Leibniz and the other prominent mathematicians of the late seventeenth and eighteenth centuries generally relied on the use of infinitesimals. Their intuition and insight into the problems they studied enabled them to develop most of what is now known as elementary calculus and a great deal of more advanced mathematics, even though their arguments did not meet modern standards of mathematical reasoning.

It was not until the nineteenth century that calculus was widely studied as a systematic application of the concept of limit. This change in point of view was stimulated by the work of the Swiss mathematician L. Euler (1707–1783), the French mathematicians J. Le Rond D'Alembert (1717–1783), S. L'Hullier (1750–1840), A. Cauchy (1789–1857), and the Czechoslovakian priest B. Bolzano (1781–1848).

Interactive Examples 2.8

Interactive solutions are on the web page http//www.math.ucsd.edu/˜ashenk/.†

1. Find the linear approximation $y = L(x)$ of $f(x) = \sqrt{x}$ at $x = 16$.

2. Give an equation relating df and dx at $x = 10$ for $f(x) = x^3 + 3x^2$.

3. What is the approximate value of $y = P(x)$ at $x = 1002$ if $P(1000) = \frac{1}{2}$ and $P'(1000) = \frac{1}{6}$?

4. By weighing a square plate of known density, it is determined that its area is 100 square centimeters with an error ≤ 0.05 square centimeters. **(a)** What is the width of the plate if its area is exactly 100 square centimeters? **(b)** Use differentials to estimate the maximum error in the answer to part (a).

†In the published text the interactive solutions of these examples will be on an accompanying CD disk which can be run by any computer browser without using an internet connection.

CONCEPTS:

1. Suppose that $y = L(x)$ is the linear approximation of $y = f(x)$ at $x = a$ and $f'(x_0) = 10$. What is $L'(x)$?

2. What is the linear approximation at $x = 5$ of $h = Af + Bg$ if the linear approximations at $x = 5$ of f and g are $L_1(x) = a + b(x-5)$ and $L_2(x) = cx + d(x-5)$, respectively?

3. What is the linear approximation at $x = 5$ of $h = fg$ if the linear approximations at $x = 5$ of f and g are $L_1(x) = a + b(x-5)$ and $L_2(x) = cx + d(x-5)$, respectively?

4. What is the linear approximation at $x = 5$ of $g = f^{10}$ if the linear approximation at $x = 5$ of f is $L_1(x) = a + b(x-5)$?

BASICS:

5.O **(a)** Find the linear approximation $y = L(x)$ of $f(x) = 8x - x^4$ at $x = 1$. C**(b)** Generate the graphs of f and L together in the window $-0.25 \le x \le 2.5$, $-2 \le y \le 10$ and copy them on your paper.

In Problems 6 through 13, find the linear approximation L of the function f at $x = a$.

6.O $f(x) = x^{1/4}$; $a = 16$

7.A $f(x) = x^2 + x^3$; $a = 10$

8. $f(x) = x^{10}$; $a = 1$

9. $f(x) = x^{-1/2}$; $a = 4$

10.O $f(x) = (x^2+1)(x^3+x+1)$; $a = 1$

11.A $f(x) = \dfrac{x+1}{x-3}$; $a = 5$

12. $f(x) = (x^2+6)^3$; $a = 2$

13. $f(x) = x\sqrt{x+5}$ $a = 4$

In Problems 14 through 22 give equations relating df and dx at $x = a$.

14.O $f(x) = x^{1/3}$; $a = 8$

15.A $f(x) = 1000 - 2000x$; $a = 8000$

16.A $f(x) = \sqrt{x}$; $a = 100$

17. $f(x) = x^{-1} - x^{-5}$; $a = 1$

18. $f(x) = 1 + x + x^2 + x^3$; $a = 2$

19.A $f(x) = 18x^{-1} - 27x^{-2}$; $a = 1$

20.O $f(x) = \dfrac{x^2}{x^4 - 15}$; $a = 2$

21. $f(x) = \sqrt{1 + x + x^2 + x^3}$; $a = 1$

22.O What is the approximate value of $T(10.02)$ if $y = T(x)$ is such that $T(10) = 7$ and $T'(10) = 4$?

23.O The width of a square is measured to be $w = 5$ inches and this number is used to calculate its area $A = w^2 = 5^2 = 25$ square inches. Use differentials to estimate the maximum possible error in the calculated area if the width is measured with an accuracy of 0.01 inches.

24.O What is the approximate value of $Z = Z(u)$ at $u = 1.09$ if $Z(1) = 100$ and $Z'(1) = -20$?

25.A What is the approximate value of $y = P(x)$ at $x = 5.005$ if $P(5) = 6000$ and $P'(5) = -100$?

26. What is the approximate value of $y = Q(x)$ at $x = 0.995$ if $Q(1) = 500$ and $Q'(1) = -1000$?

27.O On July 1, 2000 the population of Nevada was 2.066 million and was increasing at the rate of 0.103 million per year.(4) What was the approximate population of Nevada on July 1, 2001?

28.A At the beginning of year 2000, twenty-seven percent of Americans were considered to be health risks because of obesity and that percentage was increasing at the rate of 0.75 percent per year.(5) Based on this data, approximately what percent could be expected to be at risk because of obesity at the beginning of 2004?

29. At 2:00 PM a car is 100 miles south of Atlanta, Georgia and is traveling south 60 miles per hour. Approximately how far is it from Atlanta at 2:06 PM?

30. A silk thread with a one-square-millimeter cross section is 36 inches long when there is no tension

(4)Data adapted from *Nevada Population Estimates* Reno, NV: The Nevada State Demographer's office, 2001, p. 1.

(5)Data adapted from *Scientific American*, April, 2001, Washington, DC: Scientific American, Inc. , 2001, p. 30.

on it and its length increases by $y(F)$ inches when it is stretched by a force of F pounds, where $y(7) = 3.6$ inches and $y'(7) = 0.9$ inches per pound.[6] Approximately how much does the thread stretch when the force is 7.5 pounds?

31.[A] Exactly ten gallons of soup is measured to weigh 90 pounds with an error ≤ 0.1 pound. **(a)** What is the density ρ of the soup, measured in pounds per gallon, if it weighs exactly 90 pounds? **(b)** Use differentials to estimate the maximum error in the answer to part (a).

32. The radius of a sphere is measured to be $r = 10$ centimeters with an error ≤ 0.01 centimeters. and this number is used to calculate the surface area $A = 4\pi r^2$ square centimeters of the sphere. **(a)** What is the surface area of the sphere if its radius is exactly 10 centimeters? **(b)** Use differentials to estimate the maximum error in the answer to part (a).

33.[O] One leg of a right triangle is known to be exactly 3 meters long and the other leg is measured to be 4 meters with an error ≤ 0.05 meters. **(a)** What is the length of the hypotenuse of the triangle if the second leg is exactly 4 meters long? **(b)** Use differentials to estimate the maximum error in the answer to part (a).

EXPLORATION:

34.[A] Give a formula for dV in terms of r and dr, where V is the volume of a sphere of radius r and describe the formula in terms of the surface area of the sphere.

35.[A] **(a)** Standard barometric pressure $P = P(h)$ is 14.7 (pounds per square inch) at sea level and its rate of change with respect to height h above sea level is -0.00049 (pounds per square inch per foot) at sea level. Give the linear approximation of $P = P(h)$ at $h = 0$.[7] What is the error in this approximation **(b)** at 10,000 feet, and **(c)** at 20,000 feet if $P(10{,}000) = 10.2$ pounds per square inch and $P(20{,}0000 = 6.4$ pounds per square inch?

C **36.** **(a)** An object that weighs one pound on the surface of the earth weighs $f(h) = \dfrac{16}{(4+h)^2}$ pounds when it is h thousand miles above the surface of the earth. Give the linear approximation L of f at $h = 0$. **(b)** Generate the graphs of $w = f(h)$ and $w = L(h)$ in the window $-0.5 \leq h \leq 3, -0.25 \leq w \leq 1.25$. **(c)** Explain why the linear approximation is a bad mathematical model for heights above 2000 miles.

37. Show that Lemma 1 above holds for $f(x) = x^{-1}$ with $a = 1$ by finding the function $E(x)$.

38. Because the tangent line to the graph of $f(x) = x^2$ at $x = 1$ is $L(x) = 1 + 2(x - 1)$, the linear function $M(x) = 1 + 3(x - 1)$ is not the linear approximation of f at $x = 1$. Show that the error $f(x) - M(x)$ made in approximating f by M does not tend to zero faster than $x - 1$ as $x \to 1$. (See Lemma 1).

(End of Section 2.8)

[6] Data adapted from *Handbook of Engineering Materials* by D. Miner and J. Seastone, New York NY: John Wiley & Sons, 1955, pp. 3-122.

[7] Data adapted from *The World Book Encyclopedia*, Volume 1, Chicago, Il: Childcraft International, Inc., 1977, p. 8