Calculus and Analytic Geometry (Thomas) Notes

Steven Labalme

September 2, 2021

Contents

List of Figures

List of Tables

Chapter 1

The Rate of Change of a Function

1.1 Introduction

7/3:

- Discusses the importance of calculus, when it should be used, and why one should study it.
- **Analytic geometry**: "Uses algebraic methods and equations to study geometric problems. Conversely, it permits us to visualize algebraic equations in terms of geometric curves" (Thomas, 1972, p. 2).

1.2 Coordinates

- "The basic idea in analytic geometry is the establishment of a one-to-one correspondence between the points of a plane on the one hand and pairs of numbers (x, y) on the other hand" (Thomas, 1972, p. 2).
- Such a correspondence is most commonly established as follows.

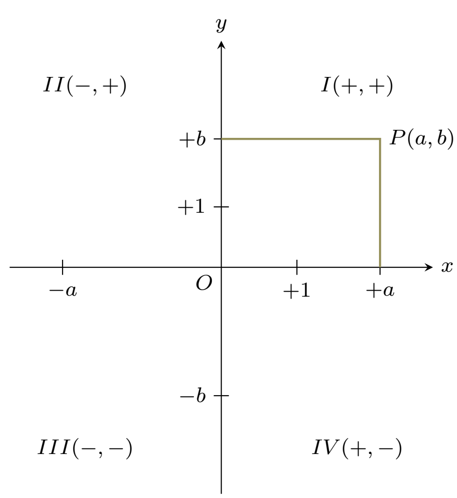


Figure 1.1: Cartesian coordinates.

  - "A horizontal line in the plane, extending indefinitely to the left and to the right, is chosen as the x-axis or axis of **abscissas**. A reference point O on this line and a unit of length are then chosen. The axis is scaled off in terms of this unit of length in such a way that the number zero is attached to O, the number $+a$ is attached to the point which is a units to the right of O, and $-a$ is attached to the symmetrically located point to the left of O. In this way, a one-to-one

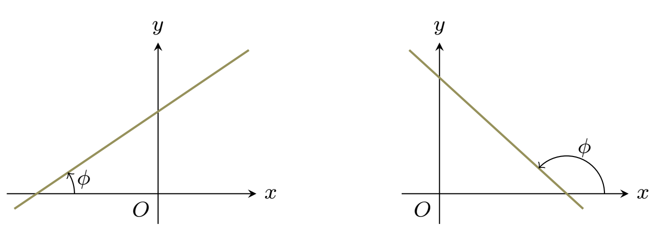
correspondence is established between points of the x-axis and the set of all **real numbers**" (Thomas, 1972, pp. 2–3).
  - "Now through O take a second, vertical line in the plane, extending indefinitely up and down. This line becomes the y-axis, or axis of **ordinates**. The unit of length used to represent $+1$ on the y-axis need not be the same as the unit of length used to represent $+1$ on the x-axis. The y-axis is scaled off in terms of the unit of length adopted for it, with the positive number $+b$ attached to the point b units above O and negative number $-b$ attached to the symmetrically located point b units below O" (Thomas, 1972, p. 3).
  - "If a line parallel to the y-axis is drawn through the point marked a on the x-axis, and a line parallel to the x-axis is drawn through the point marked b on the y-axis, their point of intersection P is to be labeled $P(a, b)$. Thus, given the pair of real numbers a and b, we find one and only one point with abscissa a and ordinate b, and this point we denote by $P(a, b)$" (Thomas, 1972, p. 3).
  - "Conversely, if we start with any point P in the plane, we may draw lines through it parallel to the coordinate axes. If these lines intersect the x-axis at a and the y-axis at b, we then regard the pair of numbers (a, b) as corresponding to the point P. We say that the coordinates of P are (a, b)" (Thomas, 1972, p. 3).
  - "The two axes divide the plane into four quadrants, called the first quadrant, second quadrant, and so on, and labeled I, II, III, IV in [Figure 1.1]. Points in the first quadrant have both coordinates positive, and in the second quadrant the x-coordinate (abscissa) is negative and the y-coordinate (ordinate) is positive. The notations $(-, -)$ and $(+, -)$ in quadrants III and IV of [Figure 1.1] represent the signs of the coordinates of points in these quadrants" (Thomas, 1972, p. 3).

1.3 Increments

- **Increments**: The values $\Delta x = x_2 - x_1$ and $\Delta y = y_2 - y_1$ concerning a particle, the initial position of which is $P_1(x_1, y_1)$ and the terminal position of which is $P_2(x_2, y_2)$.
- If the unit of measurement for both axes is the same, then we may express distances in the plane in terms of this unit using the Pythagorean theorem.

1.4 Slope of a Straight Line

- Let L be a straight line not parallel to the y-axis intersecting distinct points $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$. Then L has a **rise**, **run**, and **slope**.
- **Rise**: The increment Δy.
- **Run**: The increment Δx.
- **Slope**: The rate of rise per run $m = \frac{\text{rise}}{\text{run}} = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1}$. *Also known as* **inclination**.



(a) Negative x-intercept. (b) Positive x-intercept.

Figure 1.2: The slope and the angle of inclination.

- If we chose different distinct points, the slope would be same because the triangles in the Cartesian plane would be similar.
- Δy is proportional to Δx with m as the proportionality factor.
- On interpolation: If we're given the values of a function at (x_1, y_1) and (x_2, y_2), then we may approximate the function by a straight line L passing through those two points and approximate the value $f(x)$ for any $x_1 \leq x \leq x_2$.
- If the scales on both axes are equal, then the slope of L is equal to the tangent of the **angle of inclination** that L makes with the positive x-axis. That is, $m = \tan\phi$ (see Figure 1.2).

- **Parallel** (lines): Two lines with equal inclinations ($m_1 = m_2$).
- **Perpendicular** (lines): Two lines with inclinations that differ by 90° ($m_1 = -\frac{1}{m_2}$).
  - Note that we can prove the relation between the slopes using the angles of inclination as follows.

$$\begin{aligned} m_1 &= \tan\phi_1 \\ &= \tan(\phi_2 + 90^\circ) \\ &= -\cot\phi_2 \\ &= -\frac{1}{\tan\phi_2} \\ &= -\frac{1}{m_2} \end{aligned}$$

1.5 Equations of a Straight Line

7/5:
- How do you know if $P(x, y)$ is a point on the line P_1P_2 through distinct points $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$?
  - If $x_1 = x_2$, then P_1P_2 is vertical and P lies on P_1P_2 iff $x = x_1$.
  - If $x_1 \neq x_2$, then the slope of P_1P_2 $m_{P_1P_2} = \frac{y_2 - y_1}{x_2 - x_1}$. Thus, P lies on P_1P_2 iff $P = P_1$ or, for the line PP_1 through P and P_1, $m_{P_1P_2} = m_{PP_1} = \frac{y - y_1}{x - x_1}$. In other words, the coordinates x, y of P must satisfy $y - y_1 = m_{P_1P_2}(x - x_1)$.
- Thomas (1972) calls the above equation the **point-slope form**.
- **Variable**: "A symbol, such as x, which may take on any value in some set of numbers" (Thomas, 1972, p. 10).
- **Slope-intercept form**: $y = mx + b$.
- **General form**: $Ax + By + C = 0$.
  - Such an equation (one that contains only first powers of x and y and constants) is said to be **linear in *x* and *y***.
  - "Every straight line in the plane is represented by a linear equation and, conversely, every linear equation represents a straight line" (Thomas, 1972, p. 10).
- ***y*-intercept**: The constant b in the above equation.
- Let L be a line with the equation $Ax + By + C = 0$. The shortest distance d from a point $P_1(x_1, y_1)$ not on L to L is given by

$$d = \frac{|Ax_1 + By_1 + C|}{\sqrt{A^2 + B^2}}$$

  - Derive by finding a line perpendicular to L through P_1.

1.6 Functions and Graphs

- **Domain** (of a variable x): "The set of numbers over which x may vary" (Thomas, 1972, p. 12).
- Defines **open intervals**, **half-open intervals**, and **closed intervals**.

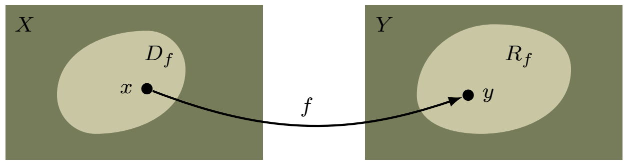


Figure 1.3: A function f maps the domain D_f onto the range R_f. The image of x is $y = f(x)$.

- **Function**: For two nonempty sets X, Y, the collection f of ordered pairs (x, y) with $x \in X$ and $y \in Y$ that assigns to every $x \in X$ a unique $y \in Y$. *Also known as* **mapping** (from X to Y), $\boldsymbol{y = f(x)}$, $\boldsymbol{f : x \to y}$[1].
  - When using the latter notation, it is understood that the domain is $\mathbb{R}$ unless this is impossible (e.g., $f : x \to \frac{1}{x}$ must exclude 0 from the domain).
- **Domain** (of a function f): "The collection of all first elements x of the pairs (x, y) in f" (Thomas, 1972, p. 13). *Also known as* $\boldsymbol{D_f}$.
- **Range** (of a function f): "The set of all second elements y of the pairs (x, y) in f" (Thomas, 1972, p. 13). *Also known as* $\boldsymbol{R_f}$.
- **Image** (of x): The value y to which a function maps x.
- Thomas (1972) considers functions from the reals to the reals, but also more abstract functions.
  - For example, it considers the function from all triangles (a set of decidedly nonnumerical objects) to their enclosed areas (the set of positive real numbers).

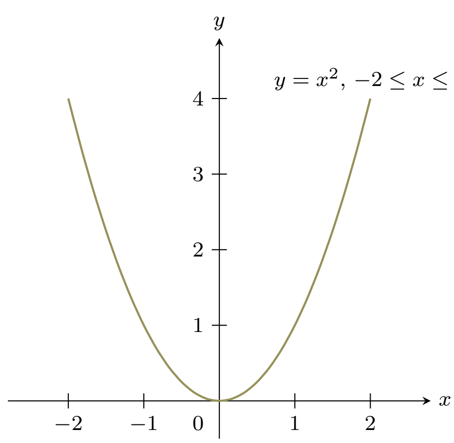


Figure 1.4: Graph of a function.

- **Graph** (of a function): "The set of points which correspond to members of the function" (Thomas, 1972, p. 14).

[1] "eff sends ex into wy"

  - For example, let X be the closed interval $[-2, 2]$. To each $x \in X$, assign the number x^2. This describes the function
$$f = \{(x, y) : -2 \leq x \leq 2,\ y = x^2\}$$
The graph of f can be seen in Figure 1.4.

- **Independent variable**: The first variable x in the ordered pair (x, y). *Also known as* **argument**.
- **Dependent variable**: The second variable y in the ordered pair (x, y).
- **Real-valued function of a real variable**: "A function f whose domain and range are sets of real numbers" (Thomas, 1972, p. 14).
  - As a general rule, *function* indicates a real-valued function of a real variable for the first seven chapters of Thomas (1972).
- f can be represented by...
  - A table of corresponding values (this will be incomplete, though).
  - Corresponding numerical scales, as on a slide rule (this will be incomplete, though).
  - A simple formula, such as $f(x) = x^2$ (this may be less exact than ordered pairs, but it is more easily understood/applicable/complete).
  - A graph (for any value x in the domain, begin x units from the origin along the x-axis, move vertically until intersecting the curve, and then move horizontally until intersecting the image y on the y-axis).
- Some mappings cannot be expressed in terms of algebraic operations on x.
  - For example, the **greatest-integer function** "maps any real number x onto that unique integer which is the largest among all integers that are less than or equal to x" (Thomas, 1972, p. 15).
    - The image of x is represented by $[x]$, and the function by $f : x \to [x]$.
    - An example of a **step function**.
    - It exhibits points of **discontinuity**.
- Note: The fact that a one-to-one mapping exists between the points in the interval $(0, 1]$ and $[1, \infty)$ (namely, $f : x \to \frac{1}{x}$) proves that there are equally many points in both intervals.
- The absolute value function can be geometrically interpreted in the context of distance from a point. As such, it is useful in describing **neighborhoods**.

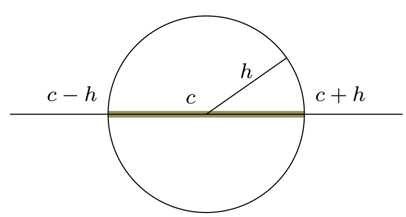


Figure 1.5: The symmetric neighborhood $N_h(c)$, centered at c, with radius h.

- **Symmetric neighborhood** (of a point c): "The open interval $(c - h, c + h)$, where h may be any positive number" (Thomas, 1972, p. 17). *Also known as* $\boldsymbol{N_h(c)}$.
- **Radius** (of a symmetric neighborhood): The value h (see Figure 1.5).
- **Neighborhood** (of a point c): The open interval $(c-h, c+k)$, where h, k may be any positive numbers.
  - Like requiring that $|x - c|$ is small.

- **Deleted neighborhood** (of a point c): "A neighborhood of c from which c itself has been removed" (Thomas, 1972, p. 17).
  - Like requiring that $|x - c| > 0$.
- **Intersection** (of the neighborhoods $(c-h_1, c+k_1)$ and $(c-h_2, c+k_2)$): The neighborhood $(c-h, c+k)$, where $h = \min(h_1, h_2)$ and $k = \min(k_1, k_2)$.
  - "The intersection of two neighborhoods of c is a neighborhood of c, and the intersection of two deleted neighborhoods of c is a deleted neighborhood of c" (Thomas, 1972, p. 18).
- Let A be a neighborhood of c. Then denote the deleted neighborhood equivalent to A with c removed by A^-.

7/6:

- "A function is determined by the domain and by any rule that tells what image in the range is to be associated with each element of the domain" (Thomas, 1972, p. 18).
  - Thus, we can think of a "function machine" that takes in elements of the domain and computes the image based on the rule.
  - Thomas (1972) visualizes function machines as flow charts.
  - In theory, a function machine could store every pair (x, y) in its memory to be recalled later. Since machines have limited memory, a fractional set of pairs could also be stored and the values in between calculated by interpolation.
  - In practice, though, calculating as we go is usually best.
- Two restrictions should be inferred to apply to the domain of a function, even if they are unstated: First, never divide by 0. Second, do not consider complex outputs (yet).
- Sometimes we have functions of more than one independent variable.
  - For example, the volume $v = \frac{1}{3}\pi r^2 h$ of a right circular cone is uniquely determined only when r, h are given definite, positive, nonzero values.
  - "Its domain is the set of all pairs (r, h) with $r > 0$, $h > 0$. Its range is the set of positve numbers $v > 0$" (Thomas, 1972, p. 20).
  - r, h are independent variables. v is a dependent variable.
- "More generally, suppose that some quantity y is uniquely determined by n other quantities $x_1, x_2, x_3, \dots, x_n$. The set of all ordered $(n+1)$-tuples $(x_1, x_2, x_3, \dots, x_n, y)$ that can be obtained by substituting permissible values of the variables $x_1, x_2, \dots, x_n$ and the corresponding values of y is a function whose domain is the set of all allowable n-tuples $(x_1, x_2, x_3, \dots, x_n)$ and whose range is the set of all possible values of y corresponding to this domain. If values can be assigned independently to each of the x's, we call them independent variables and say that y is a function of the x's. We also write $y = f(x_1, x_2, x_3, \dots, x_n)$ to indicate that y is a function of the n x's, just as we write $y = f(x)$ to indicate that y is a function of one independent variable x" (Thomas, 1972, p. 20).
  - Note, though, that functions of a single variable will be the primary concern of this book.
- **Signum function**: The function

$$\operatorname{sgn} x = \begin{cases} -1 & x < 0 \\ 0 & x = 0 \\ 1 & x > 0 \end{cases}$$

1.7 Ways of Combining Functions

7/7:

- The domains of the sum, product, difference, and quotient of two functions f and g are the intersections of D_f and D_g.
  - Note that for the quotient, we must also exclude points where $g(x) = 0$.
- There is a distinction between $x \cdot \frac{1}{x}$ and 1 (namely, the fact that the latter includes 0 in its domain while the former excludes it).
- Translation.

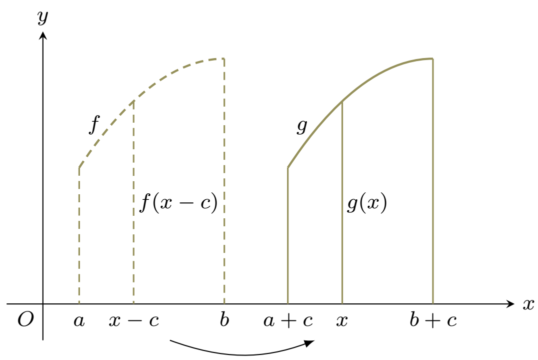
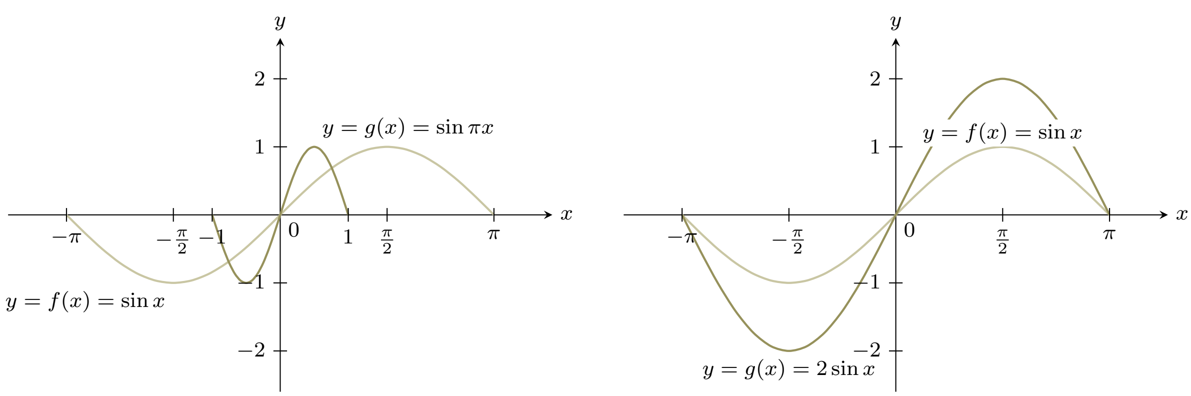


Figure 1.6: Translation.

  - "Take the ordinate of f at $x - c$ and shift it to the right c units to get the ordinate of g at x" (Thomas, 1972, p. 23).
  - $g(x) = f(x - c)$ implies that the graph of g is that of f translated c units to the right.
  - If $D_f = [a, b]$ ($f(x)$ is only defined when $a \leq x \leq b$), then $f(x-c)$ is only defined when $a \leq x-c \leq b$. This implies that $g(x)$ is only defined when $a + c \leq x \leq b + c$; hence, $D_g = [a + c, b + c]$.
- Change of scale.



(a) Horizontal scaling. (b) Vertical scaling.

Figure 1.7: Change of scale.

  - Suppose $g(x) = f(kx)$. Then to transform the graph of f into that of g, "compress or stretch the x-axis by shrinking (if $k > 1$) or stretching (if $k < 1$) every interval of length k on the x-axis into an interval of length 1" (Thomas, 1972, p. 24).

  - Let $g(x) = f(kx)$. If $D_f = [a,b]$ ($f(x)$ is only defined when $a \leq x \leq b$), then $f(kx)$ is only defined when $a \leq kx \leq b$. This implies that $g(x)$ is only defined when $a/k \leq x \leq b/k$; hence, $D_g = [a/k, b/k]$.
  - Be wary when $k = 0$.
- Suppose $g(x) = k \cdot f(x)$. Then to transform the graph of f into that of g, "stretch the f curve vertically (if $k > 1$), or compress it (if $k < 1$) [or] change the scale on the y-axis so that points labeled $1, 2, 3, \ldots$ for the graph of f are relabled $k, 2k, 3k, \ldots$ for the graph of $[g]$" (Thomas, 1972, p. 25).
  - This kind of stretching may cause R_g to differ from R_f by some factor k, but it will not affect D_f and D_g.

1.8 Behavior of Functions

- Linear functions (refer to Figure 1.8 throughout the following).

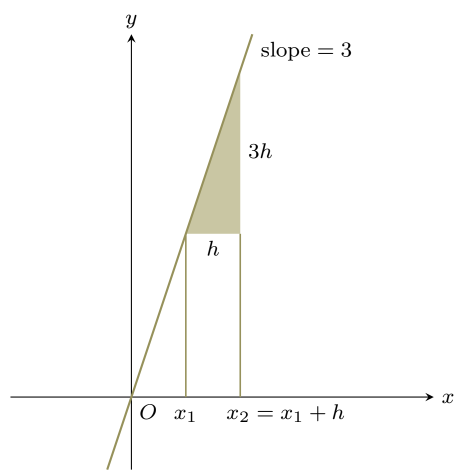


Figure 1.8: Behavior of linear functions.

  - For example, let $f(x) = 3x$. The graph of f is a straight line through the origin with slope $+3$.
  - Let x_1 be an initial value of x, and let $x_2 = x_1 + h$ be a new value of x obtained by increasing x_1 by h units[2]. The corresponding increase in $f(x)$ is given by

$$f(x_1 + h) - f(x_1) = 3(x_1 + h) - 3x_1 = 3h$$

Thus, we see that f everywhere changes three times as fast as x.

- Quadratic functions (refer to Figure 1.9 throughout the following).
  - Unlike linear functions, quadratic functions do not everywhere have a constant rate of change.
  - Let $f(x) = x^2$. In an analogous manner to the previous example, we find that the change between x_1 and $x_2 = x_1 + h$ is

$$f(x_1 + h) - f(x_1) = (x_1 + h)^2 - x_1^2 = 2x_1h + h^2$$

Thus, we see that the rate of change of f is dependent on both the initial value of x and the amount of increase in x.

[2]Note that h is an alternative notation for Δx.

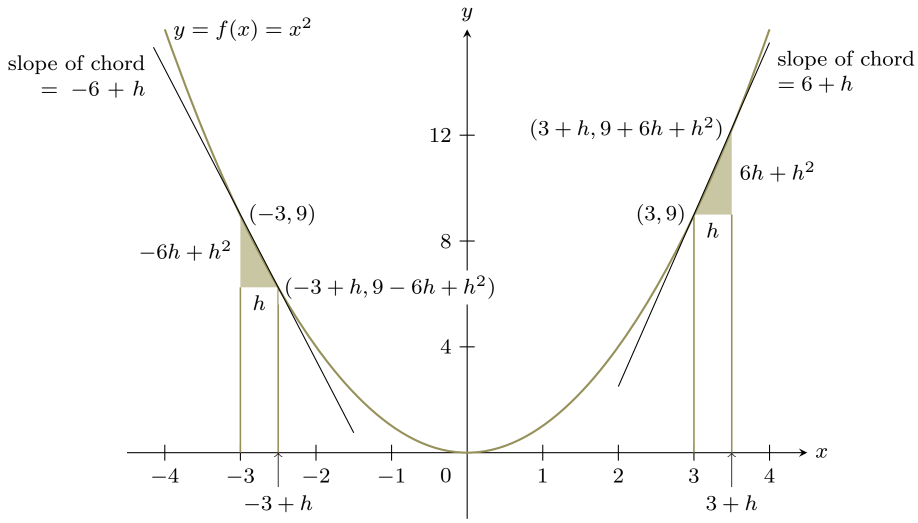


Figure 1.9: Behavior of quadratic functions.

- While a linear function increases at a rate directly proportional to the increase in x, the above demonstrates that "the increase in x^2, as x increases from x_1 to $x_1 + h$, is $2x_1 + h$ times the increase in x" (Thomas, 1972, p. 27).
- We can now define the **average rate of increase**.
- Thus, the average rate of increase of x^2 is $2x_1 + h$.
- Now let's see what happens as h shrinks.

		x_1			
		2	3	-2	-3
h	1	5	7	-3	-5
	0.5	4.5	6.5	-3.5	-5.5
	0.25	4.25	6.25	-3.75	-5.75
	0.1	4.1	6.1	-3.9	-5.9
	0.01	4.01	6.01	-3.99	-5.99
	0.001	4.001	6.001	-3.999	-5.999

Table 1.1: Average rate of change of x^2 versus h.

- From Table 1.1, we see that smaller values of h cause the average rate of change to tend toward $2x_1$. This is the beginning of **differential calculus**.

- **Average rate of increase** (of $f(x)$, per unit of increase in x, from x_1 to $x_1 + h$): The ratio,

$$\frac{f(x_1+h) - f(x_1)}{(x_1+h) - x_1} = \frac{f(x_1+h) - f(x_1)}{h} = \frac{\text{change in } f(x)}{\text{change in } x}$$

- **Differential calculus**: The branch of calculus concerned with the *instantaneous* rate of increase of a function, as opposed to the *average* rate of increase.
- So constant and linear functions are easy to analyze. But for more complicated functions, we need more advanced tools.
- Let's begin exploring the instantaneous rate of change, continuing with the parabola example.

- For x^2, the average rate of change is given by $\frac{f(x_1+h)-f(x_1)}{h} = 2x_1 + h, \ h \neq 0$.
- The $h \neq 0$ is critical — we wish to consider the case where $h = 0$, but we cannot. However, we can consider values of the slope function $m(h) = 2x_1 + h$ in a deleted neighborhood of $h = 0$. By decreasing the radius of the neighborhood, we can get progressively closer to analytically approximating $m(0)$.

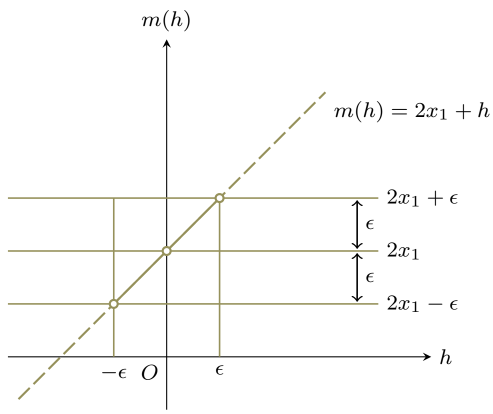


Figure 1.10: Deleted neighborhood of a slope function.

- From Figure 1.10, we can see that $m(h)$ is bounded between $2x_1 + \epsilon$ and $2x_1 - \epsilon$ when $|h|$ is less than ϵ (or any positive number smaller than ϵ)[3].
- In fact, $0 < |h| < \epsilon$, or $-\epsilon < h < \epsilon$, directly implies $2x_1 - \epsilon < 2x_1 + h < 2x_1 + \epsilon$.
- We can now formally define **approximation**, such as what was just described.

- **Approximation** (of $f(x)$ by L to within ϵ on (a, b)): The value L approximates $f(x)$ to within ϵ on the interval (a, b) if $L - \epsilon < f(x) < L + \epsilon$ when $a < x < b$.

1.9 Slope of a Curve

- **Slope of the curve** (at P): The limiting value of the slope of the secant between distinct points P, Q on the curve $y = f(x)$ as Q moves along the curve progressively closer to P. *Also known as* **slope of the tangent to the curve** (at P).
  - A purely geometric definition also exists: "Let C be a curve and P a point on C. If there exists a line L through P such that the measure of one of the angles between L and the secand line PQ approaches zero as Q approaches P along C, then L is said to be tangent to C at P" (Thomas, 1972, p. 30).
  - An advantage of the geometric definition is that it does not depend on the coordinate axes and allows vertical lines.
  - However, in most cases, we will stick with the algebraic definition.
- Thomas (1972) considers the average rate of increase equation for a cubic function, informally allowing Δx to tend towards 0 to derive a slope function.

1.10 Derivative of a Function

- We now formalize our notion of a slope function.

[3]Note that ϵ is used to denote an arbitrary (often arbitrarily small) positive number.

- We know that the slope m_{sec} of the secant from $P(x,y)$ to a point on the curve $y=f(x)$ at $(x+\Delta x, f(x+\Delta x))$ is given by

$$m_{\text{sec}} = \frac{f(x+\Delta x)-f(x)}{\Delta x}$$

- Now as Δx tends toward 0, m_{sec} tends toward the slope m_{tan} of the tangent at P. The mathematical symbols which summarize this discussion are

$$m_{\text{tan}} = \lim_{Q\to P} m_{\text{sec}} = \lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x} = \lim_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}$$

- The number given by the last operation in the above equation is clearly related to f. Thus, to indicate relation, we define

$$\boldsymbol{f'(x)} = \lim_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}$$

Also known as $\boldsymbol{y'}$, $\boldsymbol{\frac{\mathrm{d}y}{\mathrm{d}x}}$, $\boldsymbol{D_x y}$.

- This limit may sometimes fail to exist. However, at each point where it does exist, f is said to have a **derivative**, or to be **differentiable**. Similarly, $f'(x)$ is said to be the **derivative** (of f at x).
- Differential calculus is concerned with two problems.
  1. "Given a function f, determine those values of x (in the domain of f) at which the function possesses a derivative" (Thomas, 1972, p. 32).
  2. "Given a function f and an x at which the derivative exists, find $f'(x)$" (Thomas, 1972, p. 32).
- **Derived function**: "The set of all pairs of numbers $(x, f'(x))$ that can be formed by this process" (Thomas, 1972, p. 33). *Also known as* **derivative** (of f).
  - "The domain of f' is a subset of the domain of f" (Thomas, 1972, p. 33).
  - Symbolically, $D_{f'}\subset D_f$. However, for most functions considered in this book, $D_{f'} = D_f$ with maybe a few exceptions.
- On computing $f'(x)$ by eliminating the division by 0 and then substituting: "We may say that after the division by Δx has been carried out and the expression has been reduced to a form... which 'makes sense' (that is, does not involve division by zero) when Δx is taken equal to zero, then the limit as Δx approaches zero does exist and may be found by simply replacing Δx by zero in this reduced form" (Thomas, 1972, p. 34).
- Essentially, what we are doing when we eliminate the division by zero is we are expanding the domain of the function, the limit of which we are taking, to include a point of interest (Thomas (1972) elaborates quite a bit on this point).

1.11 Velocity and Rates

- Mainly just applies average and instantaneous rates of change to the physical problem of distance and velocity. However...
- "Derivatives are important in economic theory, where they are usually indicated by the adjective **marginal**" (Thomas, 1972, p. 37).
  - "Suppose that in order to produce $x+\Delta x$ tons of steel weekly, it would cost $y+\Delta y$ dollars. The increase in cost per unit increase in output would be $\Delta y/\Delta x$. The limit of this ratio, as Δx tends to zero, is called the **marginal cost**" (Thomas, 1972, p. 37).
  - There also exists **marginal revenue** $\mathrm{d}P/\mathrm{d}x$ and **marginal profit** $\mathrm{d}T/\mathrm{d}x$.

- Note: "The *average* rate of change of y per unit change in x, $\Delta y/\Delta x$, when multiplied by the number of units change in x, Δx, gives the actual change in y:

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

The *instantaneous* rate of change of y per unit change in x, $f'(x)$, multiplied by the number of units change in x, Δx, gives the change that would be produced in y if the point (x, y) were to move along the tangent line instead of moving along the curve; that is

$$\Delta y_{\text{tan}} = f'(x)\,\Delta x$$

One reason calculus is important is that it enables us to find quantitatively how a change in one of two related variables affects the second variable" (Thomas, 1972, p. 38).

Chapter 4

Applications

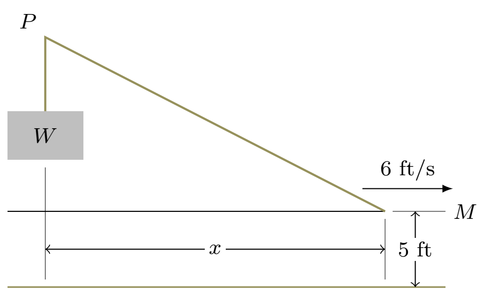
4.1 Increasing or Decreasing Functions: The Sign of $\mathrm{d}y/\mathrm{d}x$

7/8:
- **Increasing** (function f on $[a,b]$): A function f such that $f(x_1) > f(x_2)$ when $x_1 > x_2$ for all x_1, x_2 in the interval $[a,b]$. *Also known as* **rising**.
- **Decreasing** (function f on $[a,b]$): A function f such that $f(x_1) < f(x_2)$ for $a \leq x_2 < x_1 \leq b$. *Also known as* **falling**.
  - Sometimes, we consider functions that increase or decrease on open or half-open intervals.
- **Increasing** (function f at a point c): A function f such that in some neighborhood N of c, $x > c \Rightarrow f(x) > f(c)$ and $x < c \Rightarrow f(x) < f(c)$ for all $x \in N$.
- **Decreasing** (function f at a point c): A function f such that in some neighborhood N of c, $x > c \Rightarrow f(x) < f(c)$ and $x < c \Rightarrow f(x) > f(c)$ for all $x \in N$.
  - As a strange example, $\operatorname{sgn} x$ is increasing at $x = 0$.
- A function may oscillate sufficiently fast at a point to be neither increasing nor decreasing.
  - For example, for
  $$f(x) = \begin{cases} x \sin \frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$$
  "no matter how small a neighborhood of zero N may be, there are x's in N for which $f(x)$ is positive and those for which it is negative. This function oscillates infinitely often between positive and negative values in ever neighborhood of $x = 0$" (Thomas, 1972, p. 107).
- When $\mathrm{d}y/\mathrm{d}x > 0$, y is increasing. When $\mathrm{d}y/\mathrm{d}x < 0$, y is decreasing. When $\mathrm{d}y/\mathrm{d}x = 0$, y may be increasing (consider $y = x^3$), decreasing (consider $y = -x^3$), or neither (consider $y = x^2$).
- There is a relation between increasing and decreasing points, and positive and negative slopes, respectively, of the tangent lines to those points.
- Knowing where a function is increasing or decreasing can help in sketching the curve.

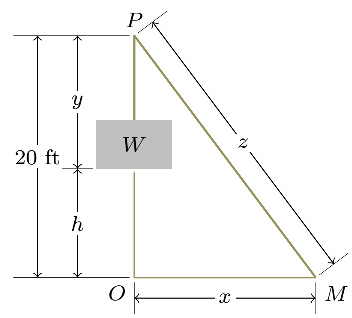
4.2 Related Rates

- In certain physical settings, we must consider not only quantities but the rates at which those quantities are changing to answer questions.

- For a **problem in related rates**, it is typical that "(a) certain variables are related in a definite way for all values of t under consideration, (b) the values of some or all of these variables and the rates of change of some of them are given at some particular instant, and (c) it is required to find the rate of change of one or more of them at this instant" (Thomas, 1972, p. 110).
  - "The variables may then all be considered to be functions of time, and if the equations which relate them for all values of t are differentiated with respect to t, the new equations so obtained will tell how their rates of change are related" (Thomas, 1972, p. 110).
- We explore three examples to illustrate the most common techniques used.
- Suppose (see Figure 4.1a) there is a "rope running through a pulley at P, bearing a weight W at one end. The other end is held in a man's hand M at a distance of 5 feet above the ground as he walks in a straight line at the rate of 6 [ft/s]" (Thomas, 1972, p. 108). Additionally (see Figure 4.1b), "suppose that the pulley is 25 ft above the ground, the rope is 45 ft long, and at a given instant the distance x is 15 ft and the man is walking away from the pulley. How fast is the weight being raised at this particular instant?" (Thomas, 1972, p. 109).



(a) The man and the pulley.



(b) Construction of the pulley.

Figure 4.1: Related rates: The pulley.

  - We begin by assessing what is given and what we want to find.
    We are given...
    (a) Relationships between the variables which are to hold for all instants of time:
    $$y + z = 45 \qquad h + y = 20 \qquad 20^2 + x^2 = z^2$$
    (b) Quantities at a given instant in time, which we may take to be $t = 0$:
    $$x = 15 \qquad \frac{\mathrm{d}x}{\mathrm{d}t} = 6$$
    We want to find...
    $$\frac{\mathrm{d}h}{\mathrm{d}t}$$
    at the instant $t = 0$.
  - We obtain a relationship between x (whose rate is given) and h (whose rate we want).
    $$y = 20 - h$$
    $$z = 45 - (20 - h) = 25 + h$$
    $$20^2 + x^2 = (25 + h)^2$$

– We now implicitly differentiate the above equation with respect to t and solve for $\mathrm{d}h/\mathrm{d}t$.

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(20^2+x^2\right) &= \frac{\mathrm{d}}{\mathrm{d}t}(25+h)^2\\ 0+2x\frac{\mathrm{d}x}{\mathrm{d}t} &= 2(25+h)\frac{\mathrm{d}h}{\mathrm{d}t}\\ \frac{\mathrm{d}h}{\mathrm{d}t} &= \frac{x}{25+h}\frac{\mathrm{d}x}{\mathrm{d}t}\end{aligned}$$

– We see that we will need the value of h at $t=0$. This can be found via the equation $20^2+x^2=(25+h)^2$ since we know the value of x at $t=0$.

$$\begin{aligned}(25+h)^2 &= 20^2+(15)^2\\ h &= 0\end{aligned}$$

– Since we now have every value that we have set equal to $\mathrm{d}h/\mathrm{d}t$, all that is left is to plug and chug.

$$\begin{aligned}\frac{\mathrm{d}h}{\mathrm{d}t} &= \frac{x}{25+h}\frac{\mathrm{d}x}{\mathrm{d}t}\\ &= \frac{15}{25+0}\cdot 6\\ &= \frac{18}{5}\ \mathrm{ft/s}\end{aligned}$$

- Suppose (see Figure 4.2) there is a "ladder 26 ft long which leans against a vertical wall. At a particular instant, the foot of the ladder is 10 ft out from the base of the wall and is being drawn away from the wall at the rate of 4 [ft/s]. How fast is the top of the ladder moving down the wall at this instant?" (Thomas, 1972, p. 110).

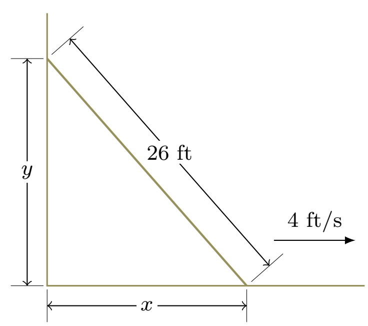


Figure 4.2: Related rates: The ladder.

– Symbolically, the problem is asking this: given

$$x^2+y^2=26^2 \qquad x=10 \qquad \frac{\mathrm{d}x}{\mathrm{d}t}=4$$

find

$$\frac{\mathrm{d}y}{\mathrm{d}t}$$

– As before, differentiate and solve for $\mathrm{d}y/\mathrm{d}t$.

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(x^2+y^2\right) &= \frac{\mathrm{d}}{\mathrm{d}t}\left(26^2\right)\\ 2x\frac{\mathrm{d}x}{\mathrm{d}t}+2y\frac{\mathrm{d}y}{\mathrm{d}t} &= 0\\ \frac{\mathrm{d}y}{\mathrm{d}t} &= -\frac{x}{y}\frac{\mathrm{d}x}{\mathrm{d}t}\end{aligned}$$

- Now find y and substitute.

$$\begin{aligned}10^2 + y^2 &= 26^2 \\ y &= 24\end{aligned}$$

$$\begin{aligned}\frac{\mathrm{d}y}{\mathrm{d}t} &= -\frac{x}{y}\frac{\mathrm{d}x}{\mathrm{d}t} \\ &= -\frac{10}{24}\cdot 4 \\ &= -\frac{5}{3}\text{ ft/s}\end{aligned}$$

- Note that the negative sign indicates that y is decreasing; that the top of the ladder is moving *down* at 5/3 ft/s (or up at $-5/3$ ft/s).

- Suppose there is an inverted right "conical reservoir [of height 10 ft and base radius 5 ft] into which water runs at the constant rate of 2 ft^3 per minute. How fast is the water level rising when it is 6 ft deep?" (Thomas, 1972, p. 111).
  - Let h be the height (in ft) of the reservoir, r be the base radius (in ft) of the reservoir, x be the radius (in ft) of the section of the cone at the water line at time t (in min), y be the depth (in ft) of water in the tank at time t (in min), and v be the volume (in ft^3) of water in the tank at time t (in min).
  - Thus, the problem is asking this: given

$$v = \frac{1}{3}\pi x^2 y \qquad\qquad \frac{x}{y} = \frac{r}{h}$$

$$h = 10 \qquad r = 5 \qquad y = 6 \qquad \frac{\mathrm{d}v}{\mathrm{d}t} = 2$$

find

$$\frac{\mathrm{d}y}{\mathrm{d}t}$$

  - Like with the pulley, we need to find an equation relating just v and y. Use a substitution based on similar triangles.

$$\begin{aligned}v &= \frac{1}{3}\pi x^2 y \\ &= \frac{1}{3}\pi\left(\frac{ry}{h}\right)^2 y \\ &= \frac{\pi r^2}{3h^2}y^3\end{aligned}$$

  - Differentiate, solve, and substitute.

$$\begin{aligned}\frac{\mathrm{d}v}{\mathrm{d}t} &= \frac{\pi r^2}{h^2}y^2\frac{\mathrm{d}y}{\mathrm{d}t} \\ \frac{\mathrm{d}y}{\mathrm{d}t} &= \frac{h^2}{\pi r^2 y^2}\frac{\mathrm{d}v}{\mathrm{d}t} \\ &= \frac{10^2}{\pi 5^2 6^2}\cdot 2 \\ &= \frac{2}{9\pi} \approx 0.071\text{ ft/min}\end{aligned}$$

4.3 Significance of the Sign of the Second Derivative

- Note: if $\mathrm{d}y/\mathrm{d}x$ fails to exist at some point P, *but* $\mathrm{d}x/\mathrm{d}y = 0$, the tangent to P is vertical.
  - On obtaining $\mathrm{d}x/\mathrm{d}y$[1]:
$$\frac{\mathrm{d}x}{\mathrm{d}y} = \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^{-1}$$
- "The sign of the second derivative tells whether the graph of $y = f(x)$ is concave upward (y'' positive) or downward (y'' negative)" (Thomas, 1972, p. 113).
- **Point of inflection**: "A point where the curve changes the direction of its concavity from downward to upward or vice versa [that is not a **cusp**]" (Thomas, 1972, p. 114). *Also known as* **inflection point**.
  - Inflection points occur where y'' changes sign. This can happen when $y'' = 0$ or when y'' fails to exist.
- **Cusp**: A sharp corner on a graph (a place where y' fails to exist).

4.4 Curve Plotting

- When sketching curves given the equation, use the following procedure.
  A. "Calculate $\mathrm{d}y/\mathrm{d}x$ and $\mathrm{d}^2y/\mathrm{d}x^2$" (Thomas, 1972, p. 115).
  B. "Find the values of x for which $\mathrm{d}y/\mathrm{d}x$ is positive and for which it is negative. Calculate y and $\mathrm{d}^2y/\mathrm{d}x^2$ at the points of transition between positive and negative values of $\mathrm{d}y/\mathrm{d}x$. These may give maximum or minimum points on the curve" (Thomas, 1972, p. 115).
  C. "Find the values of x for which $\mathrm{d}^2y/\mathrm{d}x^2$ is positive and for which it is negative. Calculate y and $\mathrm{d}y/\mathrm{d}x$ at the points of transition between positive and negative values of $\mathrm{d}^2y/\mathrm{d}x^2$. These may give points of inflection of the curve" (Thomas, 1972, p. 115).
  D. "Plot a few additional points. In particular, points which lie between the transition points already determined or points which lie to the left and to the right of all of them will ordinarily be useful. The nature of the curve for large values of $|x|$ should also be indicated" (Thomas, 1972, p. 115).
  E. "Sketch a smooth curve through the points found above, unless there are discontinuitites in the curve or its slope. Have the curve pass through its points rising or falling as indicated by the sign of $\mathrm{d}y/\mathrm{d}x$, and concave upward or downward as indicated by the sign of $\mathrm{d}^2y/\mathrm{d}x^2$" (Thomas, 1972, p. 115).
- As you plot points, consider sketching their tangents, too.
- Consider making a table with columns of significant x values, their assigned y, y', and y'' values, and any important remarks before starting to draw.
- If $f(x) = \frac{P(x)}{Q(x)}$, solve $Q(x) = 0$ to find vertical asymptotes.

4.5 Maxima and Minima: Theory

- **Relative maximum** (of f): A point $(a, f(a))$ of a function f such that $f(a) \geq f(a+h)$ for all positive and negative values of h sufficiently near zero. *Also known as* **local maximum**.
- **Relative minimum** (of f): A point $(b, f(b))$ of a function f such that $f(b) \leq f(x)$ for all x in some neighborhood of a. *Also known as* **local minimum**.

[1]This is another place where Leibniz's notation is particularly useful.

- **Absolute maximum** (of f): A point $(a, f(a))$ of a function f such that $f(a) \geq f(x)$ for all $x \in D_f$.
- **Absolute minimum** (of f): A point $(b, f(b))$ of a function f such that $f(b) \geq f(x)$ for all $x \in D_f$.
- We now prove a relationship between f' and the maxima and minima of f.

  Theorem 4.1. Let the function f be defined for $a \leq x \leq b$ and have a relative maximum or minimum at $x = c$, where $a < c < b$. If the derivative $f'(x)$ exists as a finite number at $x = c$, then $f'(c) = 0$.

  Proof. If $f'(c)$ were positive, then f would be increasing. But f is neither increasing nor decreasing at c because f has a local maximum or minimum at c. Hence, $f'(c)$ cannot be positive. Likewise, $f'(c)$ cannot be negative. Therefore, $f'(c) = 0$. □

  - Note that the theorem does not pertain to cases where $f'(c)$ does not exist, nor does it pertain to cases where c is at one of the endpoints of the interval $[a, b]$.
  - Also note that the converse of the theorem does not hold.
- The inverse of an increasing function is increasing. This also holds for decreasing functions.
- If f is only defined on $[a, b]$, then $f'(a)$ and $f'(b)$ do not exist (because the limit is different on both sides). However,

  $$f'(a^+) = \lim_{h \to 0^+} \frac{f(x+h) - f(x)}{h} \qquad\qquad f'(b^-) = \lim_{h \to 0^-} \frac{f(x+h) - f(x)}{h}$$

  may exist. These are called the **right-hand derivative** and **left-hand derivative** respectively.

  - It's imprecise to say that f is differentiable at these endpoints, but many mathematicians will allow it[2] as they expect us to just know that what is really meant is it is differentiable on (a, b) and one-side differentiable at the endpoints.
- "If the domain of f is the bounded, closed interval $[a, b]$ and if $f'(a^+)$ and $f'(b^-)$ exist, then it is easy to verify that f has a local [maximum or minimum] at a if [$f'(a^+) < 0$ or $f'(a^+) > 0$, respectively] and f has a local [minimum or maximum] at b if [$f'(b^-) < 0$ or $f'(b^-) > 0$]" (Thomas, 1972, p. 121).
- Maxima and minima are more generally referred to as **critical points** or **extrema**.
- Candidates for extrema exist at points where (1) the derivative is zero, (2) the derivative fails to exist, and (3) the domain of the function has an end.

4.6 Maxima and Minima: Problems

7/9:

- Basically optimization: Maximizing or minimizing functions using differential calculus.
- We explore a number of problems to illustrate the most common techniques used.
- "Find two positive numbers whose sum is 20 and such that their product is as large as possible" (Thomas, 1972, p. 122).
  - We want $x + y = 20$ and $x \cdot y = \max$. Thus, we need to use calculus on xy. But xy is a function of multiple variables. However, using the substitution, $y = 20 - x$, we can change xy into $20x - x^2$ and optimize that.
  - Since x, y are positive, $x \geq 0$ and $20 - x = y \geq 0 \Rightarrow x \leq 20$. Thus, the domain of $20x - x^2$ is $[0, 20]$.

[2] Really?

  - Critical points exist where $0 = \mathrm{d}y/\mathrm{d}x = 20 - 2x$, or where $x = 10$, and at the endpoints. Since $\mathrm{d}^2y/\mathrm{d}x^2 = -2$ (is negative), we know that the point at $x = 10$ is a maximum. Since $20(10) - (10)^2 > 20(0) - (0)^2 = 20(20) - 20^2$, the point at $x = 10$ is *the* maximum that we're looking for.
  - Thus, $x = 10$, $y = 20 - 10 = 10$ are the two numbers whose sum is 20 and whose product is as large as possible.

- "A square sheet of tin a inches on a side is to be used to make an open-top box by cutting a small square of tin from each corner and bending up the sides. How large a square should be cut from each corner for the box to have as large a volume as possible?" (Thomas, 1972, p. 122).
  - Suppose we cut a square of $x \times x$ in^2 from each corner of the tin sheet. Then the base of the box, when folded up, would be $a - 2x \times a - 2x$ in^2, and the height would be x in. Thus, the volume v of the box as a function of the side length x of one of the squares removed is

$$v(x) = x(a - 2x)^2$$

  - Since we cannot remove a negative area, $x \geq 0$. Furthermore, since we cannot remove more area than exists, $a - 2x \geq 0 \Rightarrow x \leq a/2$. Thus, $D_v = [0, a/2]$.
  - Critical points exist where $0 = \mathrm{d}y/\mathrm{d}x = 12x^2 - 8ax + a^2 = (2x - a)(6x - a)$, or where $x = \frac{a}{2}, \frac{a}{6}$, and at the left endpoint (the right endpoint is already one of the critical points indicated by the first derivative). Since

$$\left.\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\right|_{x=\frac{a}{6}} = -4a \qquad\qquad \left.\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\right|_{x=\frac{a}{2}} = 4a$$

    we know that only the point at $x = \frac{a}{6}$ is a maximum. In comparison with the point at $x = 0$, since $v\left(\frac{a}{6}\right) > v(0)$, the point at $x = \frac{a}{6}$ is *the* maximum.
  - Thus, every corner square removed should have dimensions $\frac{a}{6} \times \frac{a}{6}$ in^2.

- "An oil can is to be made in the form of a right circular cylinder and is to contain one quart of oil. What dimensions of the can will require the least amount of material?" (Thomas, 1972, p. 123).
  - There are 57.75 in^3 in a quart. Thus, we require that $57.75 = V = \pi r^2 h$.
  - We interpret "least amount of material" to mean "cylinder with the smallest surface area." Thus, we seek to minimize $A = 2\pi r^2 + 2\pi rh$.
  - We need to substitute either r or h from the volume equation into the area equation. Since $h = \frac{V}{\pi r^2}$ is comparatively simpler algebraically than $r = \sqrt{\frac{V}{\pi h}}$, choose to substitute out the h in the area equation, giving

$$A = 2\pi r^2 + \frac{2V}{r}$$

  - Considering the physical situation, we find that $r \in (0, \infty)$.
  - We differentiate A with respect to r and find the points where $\mathrm{d}A/\mathrm{d}r$ is equal to zero. Note that in this case, there are no endpoints to consider.

$$\begin{aligned} 0 &= \frac{\mathrm{d}A}{\mathrm{d}r} \\ &= 4\pi r - 2Vr^{-2} \\ 4\pi r^3 &= 2V \\ r &= \sqrt[3]{\frac{V}{2\pi}} \end{aligned}$$

- At this value of r,

$$\left.\frac{\mathrm{d}^2 A}{\mathrm{d}r^2}\right|_{r=\sqrt[3]{V/2\pi}} = 4\pi + 4Vr^{-3}\big|_{r=\sqrt[3]{V/2\pi}} = 12\pi$$

  Thus, the point at $r = \sqrt[3]{\frac{V}{2\pi}}$ is a relative minimum. Since $\frac{\mathrm{d}^2 A}{\mathrm{d}r^2}$ is positive for all $r \in D_A$, the point at $r = \sqrt[3]{\frac{V}{2\pi}}$ is the absolute minimum.
- Therefore, the radius and height of the desired oil can are given by

$$r = \sqrt[3]{\frac{V}{2\pi}} \approx 2.09 \text{ in} \qquad\qquad h = 2\sqrt[3]{\frac{V}{2\pi}} \approx 4.19 \text{ in}$$

- Note that there is an alternate method of solving problems of this type: Related rates.
  - We have $V = \pi r^2 h$ and $A = 2\pi r^2 + 2\pi rh$. Since V is a constant, $\mathrm{d}V/\mathrm{d}r = 0$. Since we want to find critical points of A, we set $\mathrm{d}A/\mathrm{d}r = 0$.
  - We find that

$$\frac{\mathrm{d}A}{\mathrm{d}r} = 4\pi r + 2\pi\left(h + r\frac{\mathrm{d}h}{\mathrm{d}r}\right)$$ [3]

  - We can find $\mathrm{d}h/\mathrm{d}r$ by implicitly differentiating $V = \pi r^2 h$.

$$\frac{\mathrm{d}V}{\mathrm{d}r} = 2\pi rh + \pi r^2 \frac{\mathrm{d}h}{\mathrm{d}r}$$

    Thus,

$$\begin{aligned} 0 &= 2\pi rh + \pi r^2 \frac{\mathrm{d}h}{\mathrm{d}r} \\ \frac{\mathrm{d}h}{\mathrm{d}r} &= -\frac{2h}{r} \end{aligned}$$

  - Substituting, we find that

$$\frac{\mathrm{d}A}{\mathrm{d}r} = 4\pi r - 2\pi h$$

    which implies (since we only care about the above equation when $\mathrm{d}A/\mathrm{d}r = 0$) that

$$h = 2r$$

  - Bringing back $V = \pi r^2 h$, we can now find that

$$\begin{aligned} V &= \pi r^2 (2r) \\ r &= \sqrt[3]{\frac{V}{2\pi}} \end{aligned} \qquad\qquad \begin{aligned} V &= \pi\left(\frac{h}{2}\right)^2 h \\ h &= 2\sqrt[3]{\frac{V}{2\pi}} \end{aligned}$$

  - Since

$$\frac{\mathrm{d}^2 A}{\mathrm{d}r^2} = 4\pi + \frac{4\pi h}{r}$$

    the second derivative is positive for all permissible values of r, h. Thus, the values that we have found are minimums.

- "A wire of length L is to be cut into two pieces, one of which is bent to form a circle and the other to form a square. How should the wire be cut if the sum of the areas enclosed by the two pieces is to be a maximum?" (Thomas, 1972, p. 125).

[3] Remember product rule implicit differentiation!

- Note: choose your variable names wisely. One could choose $x + y = L \Rightarrow \frac{x^2}{16} + \frac{y^2}{4\pi} = \max$, *or* one could choose $2\pi r + 4x = L \Rightarrow \pi r^2 + x^2 = \max$.
- We use related rates, as in the second answer to the previous problem. Because of the similarity in method, I will transcribe the intro math alone, without prose.

$$\begin{aligned} L &= 2\pi r + 4x & A &= \pi r^2 + x^2 \\ \frac{\mathrm{d}L}{\mathrm{d}r} &= 2\pi + 4\frac{\mathrm{d}x}{\mathrm{d}r} & \frac{\mathrm{d}A}{\mathrm{d}r} &= 2\pi r - \pi x \\ \frac{\mathrm{d}x}{\mathrm{d}r} &= -\frac{\pi}{2} & \frac{\mathrm{d}^2A}{\mathrm{d}r^2} &= 2\pi + \frac{\pi^2}{2} \end{aligned}$$

$$x = 2r \Rightarrow \begin{cases} r = \frac{L}{2\pi+8} \\ x = \frac{L}{\pi+4} \end{cases}$$

- Now is where it gets interesting. Since the second derivative is always positive, the r, x values above represent the *minimum* area, not the *maximum*. Thus, in this case, it is actually *necessary* to consider the endpoints.
- Let $r \in [0, L/2\pi]$. Thus,

$$A = \frac{1}{16}L^2,\ r = 0 \qquad A = \frac{1}{4\pi + 16}L^2,\ r = \frac{L}{2\pi + 8} \qquad A = \frac{1}{4\pi}L^2,\ r = \frac{L}{2\pi}$$

  It is now clear that $A = \max$ at the right endpoint. Thus, to maximize the enclosed area, dedicate the entirety of the wire to making a circle.

- "Fermat's principle in optics states that light travels from a point A to a point B along that path for which the time of travel is a minimum. Let us find the path that a ray of light will follow in going from a point A in a medium where the velocity of light is c_1 to a point B in a second medium where the velocity of light is c_2, when both points lie in the xy-plane and the x-axis separates the two media" (Thomas, 1972, p. 125).

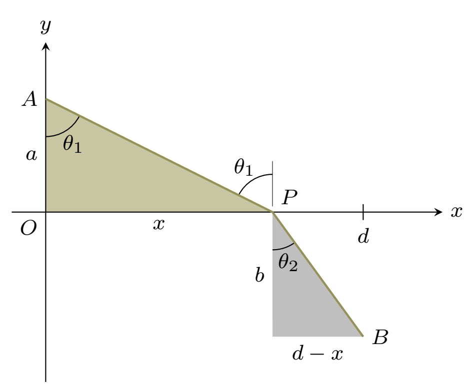


Figure 4.3: Optimization: Fermat's principle (optics).

- WLOG, let point A lie on the positive y-axis, as in Figure 4.3.
- In either medium, light will travel in a straight line since "shortest time" and "shortest path" are equivalent statements. Thus, the path will consist of a straight line segment from A to some point P along the x-axis, and then another straight line segment from P to B.
- Since $v = \frac{\Delta x}{\Delta t}$ relates velocity, distance, and time, and we are looking to minimize time, we observe that

$$t_{AP} = \frac{\sqrt{a^2 + x^2}}{c_1} \qquad t_{PB} = \frac{\sqrt{b^2 + (d - x)^2}}{c_2}$$

– We want to collectively minimize $t = t_{AP} + t_{PB}$, a function of x. Thus, we find

$$\frac{\mathrm{d}t}{\mathrm{d}x} = \frac{x}{c_1\sqrt{a^2+x^2}} - \frac{d-x}{c_2\sqrt{b^2+(d-x)^2}}$$

– As it so happens, from Figure 4.3, we can see that

$$\frac{\mathrm{d}t}{\mathrm{d}x} = \frac{\sin\theta_1}{c_1} - \frac{\sin\theta_2}{c_2}$$

– Thus, $\mathrm{d}t/\mathrm{d}x = 0$ when

$$\frac{\sin\theta_1}{c_1} = \frac{\sin\theta_2}{c_2}$$

and when θ_1, θ_2 are in domains that makes sense for the physical problem.

– Note that "instead of determining this value of x explicitly, it is customary to characterize the path followed by the ray of light by leaving the equation for $\mathrm{d}t/\mathrm{d}x = 0$ in the above form[4]" (Thomas, 1972, p. 126).

- "Suppose a manufacturer can sell x items per week at a price $P = 200 - 0.01x$ cents, and that it costs $y = 50x + 20000$ cents to produce the x items. What is the production level for maximum profits?" (Thomas, 1972, p. 127).
  – The total revenue per week on x items is $xP = 200x - 0.01x^2$.
  – The total profit T per week on x items is $T = xP - y = -0.01x^2 + 150x - 20000$.
  – T maximizes when

$$\begin{aligned} 0 &= \frac{\mathrm{d}T}{\mathrm{d}x} \\ &= -0.02x + 150 \\ x &= 7500 \text{ units} \end{aligned}$$

  – These units should be sold at \$1.25 per item.

- We conclude by outlining a general procedure to be followed for optimization questions.
  1. "When possible, draw a figure to illustrate the problem and label those parts that are important in the problem. Constants and variables should be clearly distinguished" (Thomas, 1972, p. 126).
  2. "Write an equation for the quantity that is to be a maximum or a minimum. If this quantity is denoted by y, it is desirable to express it in terms of a single independent variable x. This may require some algebraic manipulation to make use of auxiliary conditions of the problem" (Thomas, 1972, p. 127).
  3. "If $y = f(x)$ is the quantity to be a maximum or a minimum, find those values of x for which… $f'(x) = 0$" (Thomas, 1972, p. 127).
  4. "Test each value of x for which $f'(x) = 0$ to determine wheether it provides a maximum or minimum or neither. The usual tests are:
     (a) If $\mathrm{d}^2y/\mathrm{d}x^2$ is positive when $\mathrm{d}y/\mathrm{d}x = 0$, then y is a minimum.
     If $\mathrm{d}^2y/\mathrm{d}x^2$ is negative when $\mathrm{d}y/\mathrm{d}x = 0$, then y is a minimum.
     If $\mathrm{d}^2y/\mathrm{d}x^2 = 0$ when $\mathrm{d}y/\mathrm{d}x = 0$, then the test fails.

[4] This is known as the law of refraction or Snell's law. More can be read about this law on Sears (1949, p. 27).

(b) If

$$\frac{\mathrm{d}y}{\mathrm{d}x} \text{ is } \begin{cases} \text{positive} & \text{for } x < x_c \\ \text{zero} & \text{for } x = x_c \\ \text{negative} & \text{for } x > x_c \end{cases}$$

then a maximum occurs at x_c. But if $\mathrm{d}y/\mathrm{d}x$ changes from negative to zero to positive as x advances through x_c, there is a minimum. If $\mathrm{d}y/\mathrm{d}x$ does not change its sign, neither a maximum or a minimum need occur" (Thomas, 1972, p. 127).

5. "If the derivative fails to exist at some point, examine this point as possible maximum or minimum" (Thomas, 1972, p. 127).
6. "If the function $y = f(x)$ is defined for only a limited range of values $a \leq x \leq b$, examine $x = a$ and $x = b$ for possible extreme values of y" (Thomas, 1972, p. 127).

- Note that it is sometimes acceptable to forego the second-derivative test (as it was in the last problem, above): "It is often obvious from the formulas, or from physical conditions, that we have a continuous and everywhere-differentiable function that does not attain its maximum at an end point. Hence it has at least one maximum at an interior point, at which its derivative must be zero. So if we find just one zero for the derivative, we have the maximum without any appeal to second-derivative or other tests" (Thomas, 1972, p. 127).

4.7 Rolle's Theorem

- Rolle's Theorem formalizes the idea that between two point where a smooth[5] curve crosses the x-axis, there should be at least one point where the tangent to the curve is flat.
- We now formally state and prove Rolle's Theorem.

Theorem 4.2 (Rolle's Theorem)**.** Let the function f be defined and continuous on the closed interval $[a, b]$ and differentiable in the open interval (a, b). Furthermore, let $f(a) = f(b) = 0$. Then there is at least one number c between a and b where $f'(x)$ is zero; that is, $f'(c) = 0$ for some c in (a, b).

Proof. We use casework.

Case 1 ($f(x) = 0$ for all $x \in [a, b]$): Thus, $f'(x) = 0$ for all $x \in (a, b)$, and the theorem holds in this case.

Case 2 ($f(x) \neq 0$ for all $x \in [a, b]$): Thus, $f(x)$ is positive or negative somewhere on the interval. In any case, it will have a maximum positive or minimum negative value $f(c)$ at some point $x = c$ on the interval. As a positive or negative value, $f(c) \neq 0$. Thus, $f(c) \neq f(a)$ and $f(c) \neq f(b)$. Therefore, by Theorem 4.1, $f'(c) = 0$, and the theorem holds in this case, too. □

- As a corollary: "Suppose a and b are two real numbers such that (a) $f(x)$ is continuous on $[a, b]$ and its first derivative $f'(x)$ exists on (a, b), (b) $f(a)$ and $f(b)$ have opposite signs, and (c) $f'(x)$ is different from zero for all values of x in (a, b). Then there is one and only one real root of the equation $f(x) = 0$ between a and b" (Thomas, 1972, p. 130).

4.8 The Mean Value Theorem

- We now look to generalize Rolle's Theorem.
- This theorem considers a "function $y = f(x)$ which is continuous on $[a, b]$ and which has a nonvertical tangent at each point between $A(a, f(a))$ and $B(b, f(b))$, although the tangent may be vertical at one or both of the end points A and B" (Thomas, 1972, p. 131).

[5]A function with a cusp could clearly disobey this rule.

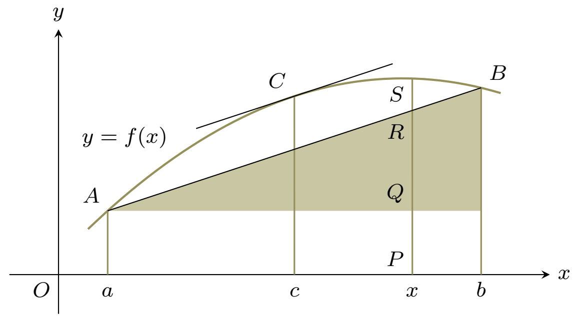


Figure 4.4: The mean value theorem.

- "Geometrically, the Mean Value Theorem states that if the function f is continuous on $[a, b]$ and differentiable on (a, b), then there is at least one number c in (a, b) where the tangent to the curve is parallel to the chord through A and B" (Thomas, 1972, p. 131).
  - This intuitively makes sense — consider moving the chord AB vertically upward or downward until it intersects only 1 point of the curve (as opposed to 0 or 2 [or, in theory, more than 2]). In Figure 4.4, this happens at C.
  - Note that this one point will occur where the vertical distance between AB and the curve is maximized.
- The idea of vertical distance is actually key to analytically proving the Mean Value Theorem.
  - The vertical distance between the chord and the curve is equal to the length of RS in Figure 4.4.
  - Also from Figure 4.4, it is clear that $RS = PS - PR$.
  - Now the length of PS is equal to $f(x)$, and the length of PR is equal to the following.

$$PR = f(a) + \frac{f(b) - f(a)}{b - a}(x - a)$$

  - Thus, the length $F(x)$ of RS at x is given by

$$F(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a}(x - a)$$

  - We prove the Mean Value Theorem by applying Rolle's Theorem to $F(x)$.
  - Indeed, Rolle's Theorem guarantees that $F'(c) = 0$ for some $c \in (a, b)$. Since $F'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$, $f'(c) = \frac{f(b)-f(a)}{b-a}$. Thus, the slope of f is indeed equal to the slope of the chord AB for at least one point in (a, b).
- We now formally state the Mean Value Theorem.

  Theorem 4.3 (The Mean Value Theorem)**.** Let $y = f(x)$ be continuous on $[a, b]$ and be differentiable in the open interval (a, b). Then there is at least one number c between a and b such that

$$f(b) - f(a) = f'(c)(b - a)$$

- Obviously, we can use differential calculus to pinpoint the values of c that satisfy the Mean Value Theorem.
- Applied to kinematics, the Mean Value Theorem tells us that on any interval where position is described by a continuous, differentiable function of time, the instantaneous velocity is equal to the average velocity at least once.

- With the Mean Value Theorem, we can prove some corollaries.

  Corollary 4.1. *If a function F has a derivative which is equal to zero for all values of x in an interval (a, b), that is, if $F'(x) = 0$ for $x \in (a, b)$, then the function is constant throughout the interval: $F(x) =$ constant for $x \in (a, b)$.*

  Proof. Suppose for the sake of contradiction that x_1, x_2 are two distinct elements of the interval (a, b) for which $F(x_1) \neq F(x_2)$. WLOG, let $x_1 < x_2$. Since F is differentiable everywhere on (a, b), the Mean Value Theorem applies. Therefore, there exists some number $c \in (a, b)$ such that $F(x_1) - F(x_2) = F'(c)(x_1 - x_2)$. Since $F'(c) = 0$ everywhere on the interval by the hypothesis, $F(x_1) = F(x_2)$, a contradiction. Thus, the value of F at x_1 is the same as its value at x_2 for all x_1, x_2 in (a, b). □

  Corollary 4.2. *If F_1 and F_2 are two functions each of which has its derivative equal to $f(x)$ for $a < x < b$, that is, if $\mathrm{d}F_1/\mathrm{d}x = \mathrm{d}F_2/\mathrm{d}x = f(x)$ for $a < x < b$, then $F_1(x) - F_2(x) =$ constant for all $x \in (a, b)$.*

  Proof. Apply Corollary 4.1 to $F(x) = F_1(x) - F_2(x)$ (the derivative of this function F is equal to zero everywhere on the interval since $F'(x) = F_1'(x) - F_2'(x) = f(x) - f(x) = 0$). □

  Corollary 4.3. *Let f be continuous on $[a, b]$ and differentiable on (a, b). If $f'(x)$ is positive throughout (a, b), then f is an increasing function on $[a, b]$, and if $f'(x)$ is negative throughout (a, b), then f is decreasing on $[a, b]$.*

  Proof. Let x_1 and x_2 be any two numbers in $[a, b]$, such that $x_1 < x_2$. By the Mean Value Theorem, $f(x_2) - f(x_1) = f'(c)(x_2 - x_1)$ for some $c \in (x_1, x_2)$. Since $(x_2 - x_1) > 0$, $f'(c)(x_2 - x_1)$ has the same sign as $f'(c)$. Thus, $f(x_2) > f(x_1)$ if $f'(x)$ is positive on (a, b), and $f(x_2) < f(x_1)$ if $f'(x)$ is negative on (a, b). □

4.9 Extension of the Mean Value Theorem

- We extend the Mean Value Theorem to prove a result about using the tangent line to a function to approximate future values of it.

  Theorem 4.4 (Extended Mean Value Theorem)**.** Let $f(x)$ and its first derivative $f'(x)$ be continuous on the closed interval $[a, b]$, and suppose its second derivative $f''(x)$ exists in the open interval (a, b). Then there is a number c_2 between a and b such that the following holds.

  $$f(b) = f(a) + f'(a)(b - a) + \frac{1}{2}f''(c_2)(b - a)^2$$

  Proof. Let K be the number defined by the following equation.

  $$f(b) = f(a) + f'(a)(b - a) + K(b - a)^2 \tag{4.1}$$

  Let $F(x)$ be the function defined by replacing every instance of b in Equation 4.1 with x and subtracting the right-hand side from the left; that is,

  $$F(x) = f(x) - f(a) - f'(a)(x - a) + K(x - a)^2 \tag{4.2}$$

  By Equation 4.2, we have $F(a) = 0$. By Equation 4.1, we have $F(b) = 0$. Moreover, F and F' are continuous on $[a, b]$, and

  $$F'(x) = f'(x) - f'(a) - 2K(x - a)$$

  Therefore, F satisfies the hypotheses of Rolle's Theorem, i.e., there exists a number $c_1 \in (a, b)$ such that $F'(c_1) = 0$. This, coupled with the facts that $F'(a) = 0$ and F'' is continuous on (a, c_1), proves that there exists a number $c_2 \in (a, c_1)$ such that $F''(c_2) = 0$. Since

  $$F''(x) = f''(x) - 2K$$

  we have $K = \frac{1}{2}f''(c_2)$. Substituting this result into Equation 4.1 gives the desired result. □

- There is an even more general version of the Extended Mean Value Theorem available as an exercise. This serves as the beginning of the calculus of sequences and series.
- Let's consider an application of the Extended Mean Value theorem: "Use the linearization of $f(x) = \sqrt{x}$ at $a = 4$ to approximate $\sqrt{5}$, and estimate the size of the error in the approximation" (Thomas, 1972, p. 135).
  - The linearization of f at a is $L_a(x) = f(a) + f'(a)(x - a)$.
  - Thus, for $f(x) = \sqrt{x}$, $L_4(x) = 2 + \frac{1}{4}(x - 4)$. Therefore, we find $\sqrt{5} \approx 2.25$.
  - By the Extended Mean Value Theorem, $f(b) - L_a(b) = \frac{1}{2}f''(c_2)(b - a)^2$ for some $c_2 \in (a, b)$.
  - Applied to this problem, the above means that the error equals $-\frac{1}{8x^{3/2}}(5-4)^2$ for some $x \in (4, 5)$. Inputting the bounds on the interval into the error function, we find that the error is between -0.011 and -0.016, i.e., $\sqrt{5} = 2.25 - e$ for some $e \in (0.011, 0.016)$.
  - Thus, even on the upper end of possible error, we have less than 1% error.
  - Practically, we'd correct our estimate to between 2.234 and 2.239, which is extremely close to the actual three-decimal value of 2.236.
- Note that "if the hypotheses of the Extended Mean Value Theorem are satisfied on $[a, b]$, then they also hold on $[a, x]$ for any $x \in (a, b)$" (Thomas, 1972, p. 135).
  - For values of x close to a, we can reasonably approximate f with the quadratic function

    $$Q_a(x) = f(a) + f'(a)(x - a) + \frac{1}{2}f''(a)(x - a)^2$$

    This is also closely related to the beginning of sequences and series.
- Further evidence that the sign of the second derivative indicates concavity: If f'' is continuous and positive at $x = a$, then it is positive in any sufficiently small neighborhood of a. Then by the Extended Mean Value Theorem, the graph of f near a lies above the tangent at a.
  - A similar argument holds for when f'' is continuous and negative at a.
  - This algebraic argument analytically supersedes our previous geometric argument.

Chapter 8

Hyperbolic Functions

8.1 Introduction

6/24: • **Hyperbolic functions**: Certain combinations of e^x and e^{-x} that are used to solve certain engineering problems (the hanging cable) and are useful in connection with differential equations.

8.2 Definitions and Identities

- Let

$$\cosh u = \frac{1}{2}(\mathrm{e}^u + \mathrm{e}^{-u}) \qquad\qquad \sinh u = \frac{1}{2}(\mathrm{e}^u - \mathrm{e}^{-u})$$

  - These combinations of exponentials occur sufficiently frequently that we give a special name to them.
  - Although the names may seem random, $\sinh u$ and $\cosh u$ do share many analogous properties with $\sin u$ and $\cos u$.
  - Pronounced to rhyme with "gosh you" and as "cinch you," respectively.
- Like $x = \cos u$ and $y = \sin u$ are associated with the point (x, y) on the unit circle $x^2 + y^2 = 1$, $x = \cosh u$ and $y = \sinh u$ are associated with the point (x, y) on the unit hyperbola $x^2 - y^2 = 1$.
  - Note that $x = \cosh u$ and $y = \sinh u$ are associated with the *right-hand* branch of the unit hyperbola.
  - Also note that sine and cosine are sometimes referred to as the **circular functions**.
- Analogous to sine and cosine, we have the identity

$$\cosh^2 u - \sinh^2 u = 1$$

- We define the remaining hyperbolic trig functions as would be expected.

$$\tanh u = \frac{\sinh u}{\cosh u} = \frac{\mathrm{e}^u - \mathrm{e}^{-u}}{\mathrm{e}^u + \mathrm{e}^{-u}} \qquad\qquad \operatorname{sech} u = \frac{1}{\cosh u} = \frac{2}{\mathrm{e}^u + \mathrm{e}^{-u}}$$

$$\coth u = \frac{\cosh u}{\sinh u} = \frac{\mathrm{e}^u + \mathrm{e}^{-u}}{\mathrm{e}^u - \mathrm{e}^{-u}} \qquad\qquad \operatorname{csch} u = \frac{1}{\sinh u} = \frac{2}{\mathrm{e}^u - \mathrm{e}^{-u}}$$

- Since $\cosh u + \sinh u = \mathrm{e}^u$, we can replace any combination of exponentials with hyperbolic sines and cosines and vice versa.
- Note that the hyperbolic functions are *not* periodic.
  - This does mean, though, that they have more easily defined properties at infinity.
- "Practically all the circular trigonometric identities have hyperbolic analogies" (Thomas, 1972, p. 267).

8.3 Derivatives and Integrals

6/25:

- Derivatives of the hyperbolic functions:

$$\frac{\mathrm{d}}{\mathrm{d}x}(\sinh u) = \cosh u\,\frac{\mathrm{d}u}{\mathrm{d}x} \qquad \frac{\mathrm{d}}{\mathrm{d}x}(\cosh u) = \sinh u\,\frac{\mathrm{d}u}{\mathrm{d}x}$$
$$\frac{\mathrm{d}}{\mathrm{d}x}(\tanh u) = \operatorname{sech}^2 u\,\frac{\mathrm{d}u}{\mathrm{d}x} \qquad \frac{\mathrm{d}}{\mathrm{d}x}(\operatorname{sech} u) = -\operatorname{sech} u \tanh u\,\frac{\mathrm{d}u}{\mathrm{d}x}$$
$$\frac{\mathrm{d}}{\mathrm{d}x}(\coth u) = -\operatorname{csch}^2 u\,\frac{\mathrm{d}u}{\mathrm{d}x} \qquad \frac{\mathrm{d}}{\mathrm{d}x}(\operatorname{csch} u) = -\operatorname{csch} u \coth u\,\frac{\mathrm{d}u}{\mathrm{d}x}$$

  - Note that these are almost exact analogs of the formulas for the corresponding circular functions, the exception being that the negative signs are not associated with the cofunctions but with the latter three.

- We now introduce the hanging cable problem, deriving the differential equation that represents the condition for equilibrium of forces acting on a section AP of a hanging cable (Figure 8.1).

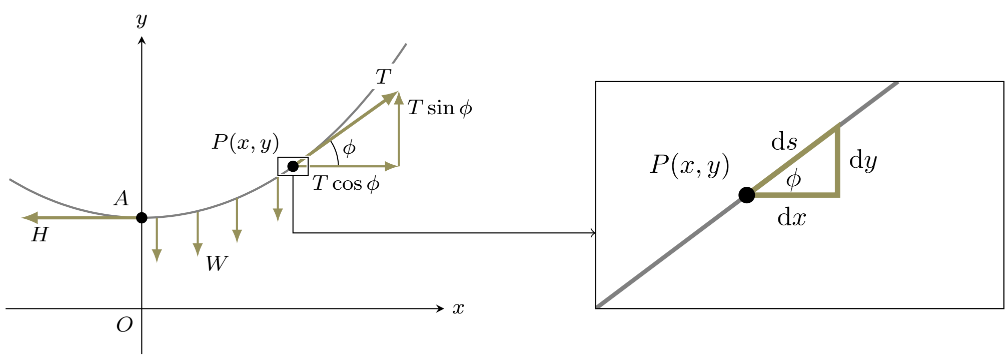


Figure 8.1: A section AP of a hanging cable.

- Let point A to be the lowest point in the arc of the hanging cable, and let it be at $(0, y_0)$ in the Cartesian plane.

- Continue along the right arc of the cable until arriving at some point $P(x, y)$.

- We wish to consider only segment AP, so we need to anchor points A and P as if the rest of the cable were still there. Now every infinitesimal sliver of the cable is being pulled (downward) slightly by gravity, but significantly (tangentially) by the rest of the cable. Thus, we can compensate at point A by pulling it tangentially left with some force H, and at point P by pulling it tangentially up and to the right with some force T.

- Since the cable is at equilibrium, the three forces acting on the cable as a whole (T, H, and W) are balanced. Thus,

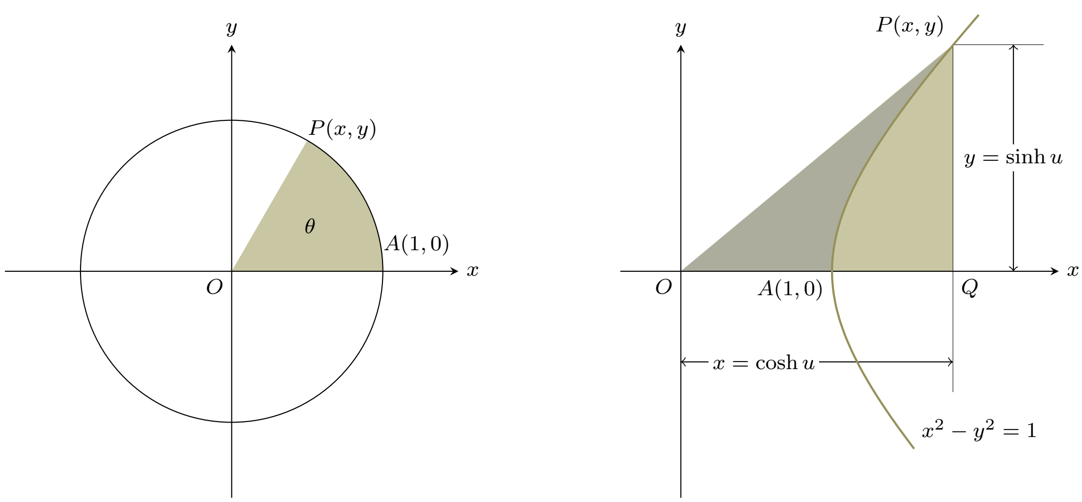
$$T \sin\phi = W$$
$$T \cos\phi = H$$

- Combining these two equations gives an important result:

$$\frac{T\sin\phi}{T\cos\phi} = \frac{W}{H}$$
$$\tan\phi = \frac{W}{H}$$

- Now $\tan\phi$ is a particularly important piece of the puzzle, because it actually equals $\frac{\mathrm{d}y}{\mathrm{d}x}$ (see the zoomed-in section of Figure 8.1). Thus,
$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{W}{H}$$
- Contrary to how it may look, W is actually not a constant — the weight of section AP is dependent on P (i.e., is dependent on how long the section is). If we assume that the cable has a uniform weight per length ratio w and that section AP is s units long, then we have $W = ws$. Thus,
$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{ws}{H}$$
- s is just arc length. Thus, $s = \int_A^P \sqrt{1 + (\mathrm{d}y/\mathrm{d}x)^2}\,\mathrm{d}x$. However, because we cannot have an integral in a differential equation, we differentiate to find $\mathrm{d}s = \sqrt{1 + (\mathrm{d}y/\mathrm{d}x)^2}\,\mathrm{d}x$.
  - Note that this expression for $\mathrm{d}s$ makes sense because, by the zoomed-in section of Figure 8.1, $\mathrm{d}s = \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2} = \sqrt{(\mathrm{d}x^2/\mathrm{d}x^2 + \mathrm{d}y^2/\mathrm{d}x^2)\mathrm{d}x^2} = \sqrt{1 + (\mathrm{d}y/\mathrm{d}x)^2}\,\mathrm{d}x$.
- If we differentiate $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{ws}{H}$, we can get $\mathrm{d}s$ into the equation and substitute, as follows, to yield the final differential equation.
$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{w}{H}\frac{\mathrm{d}s}{\mathrm{d}x}$$
$$\boxed{\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{w}{H}\sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}}$$

8.4 Geometric Meaning of the Hyperbolic Radian



(a) Circular radians. (b) Hyperbolic radians.

Figure 8.2: Geometric meaning of radians.

- For circular sine and cosine, the "meaning of the variable θ in the equations $x = \cos\theta$, $y = \sin\theta$ as they relatee to the point $P(x, y)$ on the unit circle $x^2 + y^2 = 1$ [is] the radian measure of the angle AOP in [Figure 8.2a], that is $\theta = \frac{\text{arc } AP}{\text{radius } OA}$" (Thomas, 1972, p. 271).
  - However, since $A = \frac{1}{2}r^2\theta = \frac{\theta}{2}$ for $r = 1$, θ also equals twice the area of the sector AOP.

- To understand the meaning of the variable u, calculate the area of the sector AOP in Figure 8.2b as an analog to circular area.

$$\begin{aligned} A_{AOP} &= A_{OQP} - A_{AQP} \\ &= \frac{1}{2}bh - \int_A^P y \, \mathrm{d}x \end{aligned}$$

$$y = \sinh u, \quad x = \cosh u \Rightarrow \frac{\mathrm{d}x}{\mathrm{d}u} = \sinh u \Rightarrow \mathrm{d}x = \sinh u \, \mathrm{d}u$$

$$\begin{aligned} &= \frac{1}{2}xy - \int_A^P \sinh^2 u \, \mathrm{d}u \\ &= \frac{1}{2}\cosh u \sinh u - \frac{1}{2}\int_A^P (\cosh 2u - 1) \, \mathrm{d}u \\ &= \frac{1}{2}\sinh u \cosh u - \frac{1}{2}\left[\frac{1}{2}\sinh 2u - u\right]_{A(u=0)}^{P(u=u)} \\ &= \frac{1}{2}\sinh u \cosh u - \left(\frac{1}{4}\sinh 2u - \frac{1}{2}u\right) \\ &= \frac{1}{2}\sinh u \cosh u - \left(\frac{1}{2}\sinh u \cosh u - \frac{1}{2}u\right) \\ &= \frac{1}{2}u \end{aligned}$$

  - This implies that u also equals twice the area of the sector AOP (the hyperbolic sector, that is).
  - This means, for example, that "cosh 2 and sinh 2 may be interpreted as the coordinates of P when the area of the sector AOP is just equal to the area of a square having OA as side" (Thomas, 1972, p. 272).

8.5 The Inverse Hyperbolic Functions

6/26:
- The inverse of $x = \sinh y$ is $y = \sinh^{-1} x$.
  - Since there is a one-to-one correspondence between x and y values for the inverse hyperbolic sine function, there is no need to define a principal branch (as there was with circular sine).
- The inverse of $x = \cosh y$ is $y = \cosh^{-1} x$, where $y \geq 0$ and $x \geq 1$.
  - Since there is a two-to-one correspondence between x and y values this time, we choose the positive values to be the principal branch and let the negative values be defined by the function $y = -\cosh^{-1} x$.
- Note that the only other inverse hyperbolic trig function that needs a principal branch is (rather appropriately) $\operatorname{sech} x$. Likewise, the positive values make up the principal branch.
- Like the hyperbolic trig functions have exponential forms, the inverse hyperbolic trig functions have logarithmic forms.

- For example,

$$\begin{aligned}
y &= \sinh^{-1} x \\
\sinh y &= x \\
\frac{1}{2}(\mathrm{e}^y - \mathrm{e}^{-y}) &= x \\
\mathrm{e}^y - \frac{1}{\mathrm{e}^y} &= 2x \\
\mathrm{e}^{2y} - 1 &= 2x\mathrm{e}^y \\
0 &= \mathrm{e}^{2y} - 2x\mathrm{e}^y - 1 \\
\mathrm{e}^y &= x \pm \sqrt{x^2+1} \\
y &= \ln\left(x + \sqrt{x^2+1}\right)
\end{aligned}$$

- Like the inverse circular trig functions, the inverse hyperbolic functions are quite useful as the end results of integration of radicals. First, however, we must derive their derivatives.

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\sinh^{-1} u\right) = \frac{1}{\sqrt{1+u^2}}\frac{\mathrm{d}u}{\mathrm{d}x} \qquad \frac{\mathrm{d}}{\mathrm{d}x}\left(\cosh^{-1} u\right) = \frac{1}{\sqrt{u^2-1}}\frac{\mathrm{d}u}{\mathrm{d}x}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\tanh^{-1} u\right) = \frac{1}{1-u^2}\frac{\mathrm{d}u}{\mathrm{d}x}, \quad |u|<1 \qquad \frac{\mathrm{d}}{\mathrm{d}x}\left(\operatorname{sech}^{-1} u\right) = \frac{-1}{u\sqrt{1-u^2}}\frac{\mathrm{d}u}{\mathrm{d}x}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\coth^{-1} u\right) = \frac{1}{1-u^2}\frac{\mathrm{d}u}{\mathrm{d}x}, \quad |u|>1 \qquad \frac{\mathrm{d}}{\mathrm{d}x}\left(\operatorname{csch}^{-1} u\right) = \frac{-1}{|u|\sqrt{1+u^2}}\frac{\mathrm{d}u}{\mathrm{d}x}$$

8.6 The Hanging Cable

- We seek to derive the solution to the following differential equation associated with the hanging cable problem, as described in Section 8.3.

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{w}{H}\sqrt{1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}$$

- Since the above equation involves the second derivative, we will have to deal with two constants of integration. Since it doesn't matter where the hanging cable lies in the Cartesian plane, we can choose its location such that the final answer will be as simple as possible.
  - "By choosing the y-axis to be the vertical line through the lowest point of the cable, one condition becomes $\frac{\mathrm{d}y}{\mathrm{d}x} = 0$ when $x = 0$" (Thomas, 1972, p. 277).
  - "Then we may still move the x-axis up or down to suit our convenience. That is, we let $y = y_0$ when $x = 0$, and we may choose y_0 so as to give us the simplest form in our final answer" (Thomas, 1972, p. 277).
- Let's begin solving the original equation. Since it involves y' and y'' but not y, let $y' = p$ and start by integrating with respect to p.

$$\begin{aligned}
\frac{\mathrm{d}p}{\mathrm{d}x} &= \frac{w}{H}\sqrt{1+p^2} \\
\frac{\mathrm{d}p}{\sqrt{1+p^2}} &= \frac{w}{H}\,\mathrm{d}x \\
\int \frac{\mathrm{d}p}{\sqrt{1+p^2}} &= \int \frac{w}{H}\,\mathrm{d}x \\
\sinh^{-1} p &= \frac{w}{H}x + C_1
\end{aligned}$$

- Since $p = \frac{\mathrm{d}y}{\mathrm{d}x} = 0$ and $x = 0$, $C_1 = 0$. Thus,

$$\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}x} &= \sinh\left(\frac{w}{H}x\right) \\
\mathrm{d}y &= \sinh\left(\frac{w}{H}x\right)\mathrm{d}x \\
\int \mathrm{d}y &= \int \sinh\left(\frac{w}{H}x\right)\mathrm{d}x \\
y &= \frac{H}{w}\cosh\left(\frac{w}{H}x\right) + C_2
\end{aligned}$$

- Since $y = y_0$ when $x = 0$, $C_2 = y_0 - \frac{H}{w}$. Thus, let $y_0 = \frac{H}{w}$. Therefore, we are finished:

$$y = \frac{H}{w}\cosh\left(\frac{w}{H}x\right)$$

Chapter 9

Methods of Integration

9.1 Basic Formulas

6/30:

- Useful, abstract info (that I already know) on what makes a student good at integating, e.g., integrating is an exercise in trial-and-error, but there are ways to increase your likelihood of being successful.

9.2 Powers of Trigonometric Functions

- When integrating power functions, look for integral/derivative relationships, which may allow you to substitute u and $\mathrm{d}u$ at the same time.
  - For example, when confronted with $\int \sin^n ax \cos ax \,\mathrm{d}x$, note that $\cos ax$ is almost the derivative of $\sin ax$, and choose $u = \sin ax$ and $\frac{\mathrm{d}u}{a} = \cos ax \,\mathrm{d}x$ to yield $\frac{1}{a} \int u^n \,\mathrm{d}u$.
- When integrating power functions, it may be possible to split the exponent into a product ($u^n = u^a u^b$ where $a + b = n$) and work off of properties of one of the functions raised to a smaller exponent (u^a may have properties that u^n lacks).
  - For example, when confronted with $\int \sin^3 x \,\mathrm{d}x$, recall that $\sin^2 x$ has Pythagorean properties, and split the exponent.

$$\begin{aligned}\int \sin^3 x \,\mathrm{d}x &= \int \sin^2 x \sin x \,\mathrm{d}x \\ &= \int \left(1 - \cos^2 x\right) \sin x \,\mathrm{d}x\end{aligned}$$

    Now we can use the previous property, since $\sin x$ and $\cos x$ have an integral/derivative relationship.

$$\begin{aligned}&= -\int \left(1 - u^2\right) \mathrm{d}u \\ &= \int \left(u^2 - 1\right) \mathrm{d}u\end{aligned}$$

  - Note that this technique is applicable whenever an odd power of sine or cosine is to be integrated. For higher powers, consider the following.

$$\int \cos^{2n+1} x \,\mathrm{d}x = \int \left(\cos^2 x\right)^n \cos x \,\mathrm{d}x = \int \left(1 - \sin^2 x\right)^n \cos x \,\mathrm{d}x = \int \left(1 - u^2\right)^n \mathrm{d}u$$

    Remember that $\left(1 - u^2\right)^n$ can be expanded via the binomial theorem.

- When integrating a composite trigonometric function, consider reducing it to a radical of powers of sines and cosines.
  - For example, $\sec x \tan x = \frac{\sin x}{\cos^2 x}$.
- When integrating positive integer powers of $\tan x$, use either the base cases or the **reduction formula**.
  - Begin by deriving a reduction formula.

$$\begin{aligned}\int \tan^n x \, \mathrm{d}x &= \int \tan^{n-2} x \left(\sec^2 x - 1\right) \mathrm{d}x \\ &= \int \tan^{n-2} \sec^2 x \, \mathrm{d}x - \int \tan^{n-2} x \, \mathrm{d}x \\ &= \frac{\tan^{n-1} x}{n-1} - \int \tan^{n-2} x \, \mathrm{d}x\end{aligned}$$

  Since the reduction formula decreases the exponent by 2, work out two base cases.

$$\int \tan^0 x \, \mathrm{d}x = \int \mathrm{d}x = x + C$$

$$\int \tan^1 x \, \mathrm{d}x = \int \frac{\sin x}{\cos x} \, \mathrm{d}x = -\int \frac{\mathrm{d}u}{u} = -\ln|\cos x| + C$$

  - Note that the impetus for initially deriving such a formula was the search for a way to get $\sec^2 x$ into the integrand, which can be done by splitting the exponent.
  - This method can easily be adjusted to suit negative powers of $\tan x$ (positive powers of $\cot x$).
- When integrating even powers of $\sec x$, either use the secant reduction formula, or split the exponent.
  - We derive the following formula.

$$\begin{aligned}\int \sec^{2n} x \, \mathrm{d}x &= \int \sec^{2n-2} x \sec^2 x \, \mathrm{d}x \\ &= \int \left(1 + \tan^2 x\right)^{n-1} \sec^2 x \, \mathrm{d}x \\ &= \int \left(1 + u^2\right)^{n-1} \mathrm{d}u\end{aligned}$$

- When integrating secant (or cosecant) alone, produce $\frac{u'}{u}$ by multiplying the integrand by a clever form of 1.
  - For example,

$$\begin{aligned}\int \sec x \, \mathrm{d}x &= \int \sec x \cdot \frac{\sec x + \tan x}{\sec x + \tan x} \, \mathrm{d}x \\ &= \int \frac{\sec^2 x + \sec x \tan x}{\sec x + \tan x} \\ &= \ln|\sec x + \tan x| + C\end{aligned}$$

9.3 Even Powers of Sines and Cosines

- When integrating the product of sines and cosines raised to powers where at least one exponent is a positive odd integer, split the exponent and use u-substitution.
  - In effect, we wish to evaluate $\int \sin^m x \cos^n x \, \mathrm{d}x$ where at least one of m, n is a positive odd integer.

– For example, when confronted with $\int \cos^{\frac{2}{3}} x \sin^5 x \, \mathrm{d}x$, split the exponent of $\sin^5 x$ and choose $u = \cos x$ and $-\mathrm{d}u = \sin x \, \mathrm{d}x$.

$$\int \cos^{\frac{2}{3}} x \sin^5 x \, \mathrm{d}x = \int \cos^{\frac{2}{3}} x \left(1 - \cos^2 x\right)^2 \sin x \, \mathrm{d}x = \int u^{\frac{2}{3}} \left(u^2 - 1\right) \mathrm{d}u$$

- When integrating the product of sines and cosines raised to powers where both exponents are even integers, begin by transforming it into a sum of either just sines *or* just cosines raised to even integers. Then split the exponents and use one of the following formulas. It may be necessary to use these formulas multiple times. Use them until the problem has been reduced to a sum with only odd exponents.

$$\sin^2 u = \frac{1}{2}(1 - \cos 2u) \qquad \cos^2 u = \frac{1}{2}(1 + \cos 2u)$$

  – Note that "these identities may be derived very quickly by adding or subtracting the equations [$\cos^2 u + \sin^2 u = 1$ and $\cos^2 u - \sin^2 u = \cos 2u$] and by dividing by two" (Thomas, 1972, p. 287).

  – For example, when confronted with $\int \sin^2 x \cos^4 x \, \mathrm{d}x$, begin by changing it to a case with only powers of cosine (chose to eliminate the sine function because it is raised to a lower exponent and, thus, will need less binomial expansion).

$$\begin{aligned} \int \sin^2 x \cos^4 x \, \mathrm{d}x &= \int \left(1 - \cos^2 x\right) \cos^4 x \, \mathrm{d}x \\ &= \int \cos^4 x \, \mathrm{d}x - \int \cos^6 x \, \mathrm{d}x \end{aligned}$$

Now split the exponents.

$$= \int \left(\cos^2 x\right)^2 \mathrm{d}x - \int \left(\cos^2 x\right)^3 \mathrm{d}x$$

Employ the above formulas and use binomial expansion. If necessary, repeat (split the exponents, employ the above formulas, use binomial expansion) until only odd exponents remain (remember that 1 is an odd exponent).

$$\begin{aligned} &= \int \left(\frac{1}{2}(1 + \cos 2x)\right)^2 \mathrm{d}x - \int \left(\frac{1}{2}(1 + \cos 2x)\right)^3 \mathrm{d}x \\ &= \frac{1}{4} \int \left(1 + 2\cos 2x + \cos^2 2x\right) \mathrm{d}x \\ &\quad - \frac{1}{8} \int \left(1 + 3\cos 2x + 3\cos^2 2x + \cos^3 2x\right) \mathrm{d}x \\ &= \frac{1}{4} \int \left(1 + 2\cos 2x + \frac{1}{2}(1 + \cos 4x)\right) \mathrm{d}x \\ &\quad - \frac{1}{8} \int \left(1 + 3\cos 2x + \frac{3}{2}(1 + \cos 4x) + \cos^3 2x\right) \mathrm{d}x \end{aligned}$$

These integrals may now be handled using previously discussed techniques.

9.4 Integrals With Terms $\sqrt{a^2 - u^2}$, $\sqrt{a^2 + u^2}$, $\sqrt{u^2 - a^2}$, $a^2 + u^2$, $a^2 - u^2$

- When integrating a radical that resembles the derivative of an inverse trig function, we may factor out the issue so as to make the integral resemble one of the known formulas.

- For example, when confronted with $\int \frac{\mathrm{d}u}{a^2+u^2}$, divide the a^2 term out of the denominator and integrate with respect to $\frac{u}{a}$[1].

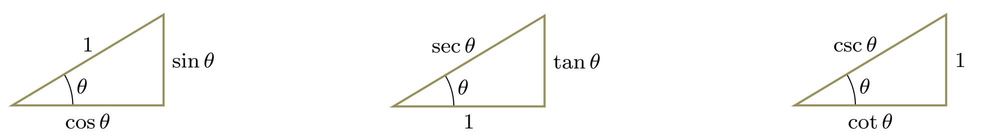
$$\begin{aligned}\int \frac{\mathrm{d}u}{a^2+u^2} &= \frac{1}{a^2}\int \frac{\mathrm{d}u}{1+\left(\frac{u}{a}\right)^2}\\ &= \frac{1}{a^2}\int \frac{a\,\mathrm{d}\left(\frac{u}{a}\right)}{1+\left(\frac{u}{a}\right)^2}\\ &= \frac{1}{a}\tan^{-1}\frac{u}{a}+C\end{aligned}$$

- However, this method is partially flawed in that it relies on having memorized the derivatives of the inverse trig functions, i.e., it is not terribly analytical. This shortcoming will now be addressed with a new, more general technique.

- The new method leans heavily on the following three identities.

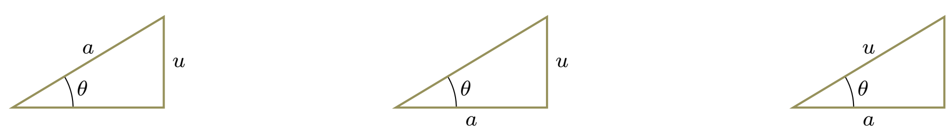
$$1-\sin^2\theta=\cos^2\theta \qquad 1+\tan^2\theta=\sec^2\theta \qquad \sec^2\theta-1=\tan^2\theta$$

  - With the help of these identities, it is possible to...
    1. use $u=a\sin\theta$ to replace a^2-u^2 with $a^2\cos^2\theta$;
    2. use $u=a\tan\theta$ to replace a^2+u^2 with $a^2\sec^2\theta$;
    3. use $u=a\sec\theta$ to replace u^2-a^2 with $a^2\tan^2\theta$.



(a) $\cos^2\theta+\sin^2\theta=1$. (b) $1+\tan^2\theta=\sec^2\theta$. (c) $\cot^2\theta+1=\csc^2\theta$.

Figure 9.1: Geometric rationale for the trigonometric identities.



(a) $\sqrt{a^2-u^2}=a\cos\theta$, $u=a\sin\theta$. (b) $\sqrt{u^2+a^2}=a\sec\theta$, $u=a\tan\theta$. (c) $\sqrt{u^2-a^2}=a\tan\theta$, $u=a\sec\theta$.

Figure 9.2: Geometric rationale for the trigonometric substitutions.

  - These identities and substitutions can be easily remembered by thinking of the Pythagorean theorem in conjunction with Figures 9.1 and 9.2, respectively.

- We may now evaluate inverse trig integrals analytically.

[1]Thomas (1972) uses differentials with more complex functions than a single variable quite often. It's not something I've seen before, but it's something I should get used to (and it does make sense if you think about it — it's just an extension of the underlying concept of separation of variables integration).

- For example, when confronted with $\int \frac{\mathrm{d}u}{a^2+u^2}$, choose $u = a\tan\theta$ and $\mathrm{d}u = a\sec^2\theta\,\mathrm{d}\theta$.

$$\begin{aligned}\int \frac{\mathrm{d}u}{a^2+u^2} &= \int \frac{a\sec^2\theta\,\mathrm{d}\theta}{a^2+(a\tan\theta)^2} \\ &= \int \frac{a\sec^2\theta}{a^2\left(1+\tan^2\theta\right)}\,\mathrm{d}\theta \\ &= \frac{1}{a}\int \frac{\sec^2\theta}{\sec^2\theta}\,\mathrm{d}\theta \\ &= \frac{1}{a}\int \mathrm{d}\theta \\ &= \frac{1}{a}\theta + C\end{aligned}$$

At this point, solve $u = a\tan\theta$ for θ and substitute.

$$= \frac{1}{a}\tan^{-1}\frac{u}{a} + C$$

- Some integrals will simplify to have a plus/minus in the denominator, leading to two possible solutions. However, there are sometimes ways to isolate a single solution.
  - For example, $\int \frac{\mathrm{d}u}{\sqrt{a^2-u^2}} = \int \frac{a\cos\theta\mathrm{d}\theta}{\pm a\cos\theta} = \pm\theta + C$. However, when we consider the fact that $\theta = \sin^{-1}\frac{u}{a}$, we know that $\theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ (because inverse sine is not arcsine, and inverse sine is only defined over the principal branch of sine). Thus, since $\theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, $\cos\theta \in [0,1]$, i.e., is always positive. Thus, we choose $\int \frac{\mathrm{d}u}{\sqrt{a^2-u^2}} = +\theta + C = \sin^{-1}\frac{u}{a} + C$ as our one solution.
  - For example, $\int \frac{\mathrm{d}u}{\sqrt{u^2-a^2}}$ equals $\ln\left|\frac{u}{a} + \frac{\sqrt{u^2-a^2}}{a}\right| + C$ or $-\ln\left|\frac{u}{a} - \frac{\sqrt{u^2-a^2}}{a}\right| + C$ depending on whether $\tan\theta$ is positive or negative. But it can be shown algebraically that the two solutions are actually equal:

$$\begin{aligned}-\ln\left|\frac{u}{a} - \frac{\sqrt{u^2-a^2}}{a}\right| &= \ln\left|\frac{a}{u-\sqrt{u^2-a^2}}\right| \\ &= \ln\left|\frac{a\left(u+\sqrt{u^2-a^2}\right)}{\left(u-\sqrt{u^2-a^2}\right)\left(u+\sqrt{u^2-a^2}\right)}\right| \\ &= \ln\left|\frac{a\left(u+\sqrt{u^2-a^2}\right)}{a^2}\right| \\ &= \ln\left|\frac{u}{a} + \frac{\sqrt{u^2-a^2}}{a}\right|\end{aligned}$$

  Thus, we have $\ln\left|\frac{u}{a} + \frac{\sqrt{u^2-a^2}}{a}\right| + C$ as the one solution[2].

- Some integrals will have extraneous constants that can be combined with C to simplify the *indefinite* integral.
  - Continuing with the above example,

$$\begin{aligned}\int \frac{\mathrm{d}u}{\sqrt{u^2-a^2}} &= \ln\left|\frac{u}{a} + \frac{\sqrt{u^2-a^2}}{a}\right| + C \\ &= \ln\left|u + \sqrt{u^2-a^2}\right| - \ln|a| + C \\ &= \ln\left|u + \sqrt{u^2-a^2}\right| + C\end{aligned}$$

[2] Note that we could choose to use the other solution, but we choose this one because it's "simpler" (it uses addition instead of subtraction).

- When integrating an inverse trig integral with excess polynomial terms, look to transform it into a (power of a) trig integral problem.
  - For example, when confronted with $\int \frac{x^2\,\mathrm{d}x}{\sqrt{9-x^2}}$, treat it as a case of a^2-u^2, but substitute the trig expression into the x^2 term in the numerator, too.

$$\int \frac{x^2\,\mathrm{d}x}{\sqrt{9-x^2}} = \int \frac{9\sin^2\theta \cdot 3\cos\theta\,\mathrm{d}\theta}{3\cos\theta} = 9\int \sin^2\theta\,\mathrm{d}\theta$$

  This integral may now be handled using previously discussed techniques.

- Many inverse trig integrals can also be evaluated hyperbolically, making use of the following three identities.

$$\cosh^2\theta - 1 = \sinh^2\theta \qquad 1-\tanh^2\theta = \operatorname{sech}^2\theta \qquad 1+\sinh^2\theta = \cosh^2\theta$$

  - With the help of these identities, it is possible to...
    1. use $u = a\tanh\theta$ to replace a^2-u^2 with $a^2\operatorname{sech}^2\theta$;
    2. use $u = a\sinh\theta$ to replace a^2+u^2 with $a^2\cosh^2\theta$;
    3. use $u = a\cosh\theta$ to replace u^2-a^2 with $a^2\sinh^2\theta$.

9.5 Integrals With ax^2+bx+c

- When integrating composite functions where the inner function is a binomial, look to factor said binomial.
  - The general quadratic $f(x) = ax^2+bx+c$, $a \neq 0$, can be reduced to the form $a\left(u^2+B\right)$ by completing the square and choosing $u = x+\frac{b}{2a}$ and $B = \frac{4ac-b^2}{4a^2}$:

$$\begin{aligned} ax^2+bx+c &= a\left(x^2+\frac{b}{a}x\right)+c \\ &= a\left(x^2+\frac{b}{a}x+\frac{b^2}{4a^2}\right)+c-\frac{b^2}{4a} \\ &= a\left(\left(x+\frac{b}{2a}\right)^2+\frac{4ac-b^2}{4a^2}\right) \end{aligned}$$

- When integrating the square root of a binomial, or some similarly tricky function of a binomial, we can transform the binomial into a form such that it will suit one of the inverse trig integrals.
  - Since it would lead to complex numbers, we disregard cases where $a\left(u^2+B\right)$ is negative, i.e., we focus on cases where (1) a is positive, and (2) a, B are both negative.
  - That being said, if it is an odd root ($\sqrt[3]{x}$, $\sqrt[5]{x}$, etc.), the sign doesn't matter.
  - For example, when confronted with $\int \frac{(x+1)\,\mathrm{d}x}{\sqrt{2x^2-6x+4}}$, begin by factoring the binomial[3].

$$2x^2-6x+4 = 2\left(x^2-3x\right)+4 = 2\left(x-\frac{3}{2}\right)^2-\frac{1}{2} = 2\left(u^2-a^2\right)$$

  Note that $u = x-\frac{3}{2}$ and $a = \frac{1}{2}$. We can now return to the integral, which we shall reformulate in terms of u in its entirety.

$$\int \frac{(x+1)\,\mathrm{d}x}{\sqrt{2x^2-6x+4}} = \int \frac{\left(u+\frac{5}{2}\right)\mathrm{d}u}{\sqrt{2\left(u^2-a^2\right)}}$$

[3]Note that, in place of inspection, we could use the general form factorization derived above.

Split it into two separate integrals and factor out the constants.

$$= \frac{1}{\sqrt{2}} \int \frac{u\,\mathrm{d}u}{\sqrt{u^2 - a^2}} + \frac{5}{2\sqrt{2}} \int \frac{\mathrm{d}u}{\sqrt{u^2 - a^2}}$$

The right integral is a straight-up inverse trig integral. The left one, however, needs something special. It could be dealt with as previously discussed by substituting $u = a\tan\theta$ for all instances of u and evaluating it is a more complex trig integral in θ. However, for the sake of showing a different technique, we will choose $z = u^2 - a^2$ and $\frac{1}{2}\mathrm{d}z = u\,\mathrm{d}u$ and treat it as a power function in z.

$$\begin{aligned}
&= \frac{1}{2\sqrt{2}} \int \frac{\mathrm{d}z}{\sqrt{z}} + \frac{5}{2\sqrt{2}} \ln\left|u + \sqrt{u^2 - a^2}\right| + C_2 \\
&= \frac{1}{2\sqrt{2}} \int z^{-\frac{1}{2}}\,\mathrm{d}z + \frac{5}{2\sqrt{2}} \ln\left|u + \sqrt{u^2 - a^2}\right| + C_2 \\
&= \frac{1}{\sqrt{2}} z^{\frac{1}{2}} + C_1 + \frac{5}{2\sqrt{2}} \ln\left|u + \sqrt{u^2 - a^2}\right| + C_2
\end{aligned}$$

Return all of the substitutions and combine the constants of integration.

$$\begin{aligned}
&= \sqrt{\frac{u^2 - a^2}{2}} + \frac{5}{2\sqrt{2}} \ln\left|u + \sqrt{u^2 - a^2}\right| + C \\
&= \sqrt{\frac{x^2 - 3x + 2}{2}} + \frac{5}{2\sqrt{2}} \ln\left|x - \frac{3}{2} + \sqrt{x^2 - 3x + 2}\right| + C
\end{aligned}$$

9.6 Integration by the Method of Partial Fractions

7/1:

- **Method of Partial Fractions**: The process of "split[ting] a fraction into a sum of fractions having simpler denominators" (Thomas, 1972, p. 294).
- If we wish to split a rational fraction $\frac{f(x)}{g(x)}$ into a sum of simpler fractions, then...
  - "The degree of $f(x)$ should be less than the degree of $g(x)$. If this is not the case, we first perform a long division, then work with the remainder term. This remainder can always be put into the required form" (Thomas, 1972, p. 294).
  - "The factors of $g(x)$ should be known. Theoretically, any polynomial $g(x)$ with real coefficient can be expressed as a product of real linear and quadratic factors. In practice, it may be difficult to perform the factorization" (Thomas, 1972, p. 294).
- If $x - r$ is a linear factor of $g(x)$ and $(x - r)^m$ is the highest power of $x - r$ that divides $g(x)$, then, to this factor, assign the sum of m partial fractions

$$\frac{A_1}{x - r} + \frac{A_2}{(x - r)^2} + \cdots + \frac{A_m}{(x - r)^m}$$

  - If $x^2 + px + q$ is a quadratic factor[4] of $g(x)$ and $\left(x^2 + px + q\right)^n$ is the highest power of $x^2 + px + q$ that divides $g(x)$, then, to this factor, assign the sum of n partial fractions

$$\frac{B_1x + C_1}{x^2 + px + q} + \frac{B_2x + C_2}{\left(x^2 + px + q\right)^2} + \cdots + \frac{B_nx + C_n}{\left(x^2 + px + q\right)^n}$$

  - Continue, as necessary, to higher degree factors of $g(x)$ (although this caveat is not addressed by Thomas (1972)).

[4] A binomial factor, the factoring of which into linear factors would introduce complex numbers.

  - Notice how the degree of the polynomial in the numerator of the partial fractions will be at most one less than the degree of the denominator.

- Any rational function integrated via the method of partial fractions can be reduced to the problem of evaluating the following two types of integrals.

$$\int \frac{\mathrm{d}x}{(x-r)^h} \qquad\qquad \int \frac{(ax+b)\,\mathrm{d}x}{(x^2+px+q)^k}$$

  - The left integral, with the substitution $z = x - r$ and $\mathrm{d}z = \mathrm{d}x$, becomes a power integral.
  - The right integral, after completing the square in the denominator, substituting, and splitting into two fractions by the numerator, becomes a pair of inverse trig substitution integrals.

- With the method of partial fractions, there is a new way to integrate $\sec\theta$.

$$\begin{aligned}
\int \sec\theta\,\mathrm{d}\theta &= \int \frac{\mathrm{d}\theta}{\cos\theta} \\
&= \int \frac{\cos\theta\,\mathrm{d}\theta}{\cos^2\theta} \\
&= \int \frac{\mathrm{d}x}{1-x^2} \\
&= \int \frac{0.5}{1+x}\,\mathrm{d}x + \int \frac{0.5}{1-x}\,\mathrm{d}x \\
&= \frac{1}{2}\ln|1+x| - \frac{1}{2}\ln|1-x| + C \\
&= \frac{1}{2}\ln\left|\frac{1+x}{1-x}\right| + C \\
&= \ln\sqrt{\frac{1+\sin\theta}{1-\sin\theta}} + C
\end{aligned}$$

  - Note that $\ln\sqrt{\frac{1+\sin\theta}{1-\sin\theta}} + C$ is equivalent to the previously derived form:

$$\begin{aligned}
\sqrt{\frac{1+\sin\theta}{1-\sin\theta}} &= \sqrt{\frac{(1+\sin\theta)^2}{1-\sin^2\theta}} \\
&= \left|\frac{1+\sin\theta}{\cos\theta}\right| \\
&= |\sec\theta + \tan\theta|
\end{aligned}$$

9.7 Integration by Parts

- This is the second general method of integration (the first being substitution).
- It relies on the following formulas for indefinite and definite integrals, respectively.

$$\int u\,\mathrm{d}v = uv - \int v\,\mathrm{d}u + C \qquad\qquad \int_{(1)}^{(2)} u\,\mathrm{d}v = uv\Big]_{(1)}^{(2)} - \int_{(1)}^{(2)} v\,\mathrm{d}u$$

  - The indefinite integral formula can be derived from the differential of a product rule as follows.

$$\begin{aligned}
\mathrm{d}(uv) &= u\,\mathrm{d}v + v\,\mathrm{d}u \\
u\,\mathrm{d}v &= \mathrm{d}(uv) - v\,\mathrm{d}u \\
\int u\,\mathrm{d}v &= uv - \int v\,\mathrm{d}u + C
\end{aligned}$$

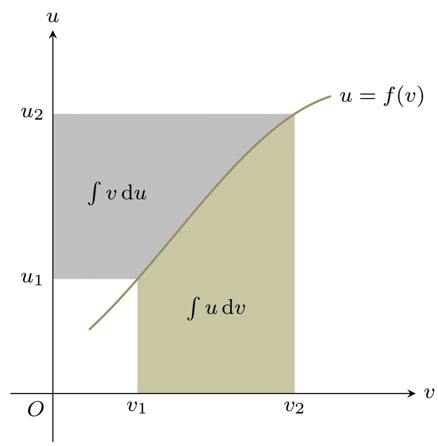


Figure 9.3: Geometric rationale for definite integration by parts.

- The definite integral formula can be thought of as an adjustment of the above, or it can be conceived geometrically: In Figure 9.3, $\int_{(1)}^{(2)} u\,\mathrm{d}v$ is the yellow area, which is clearly equivalent to the total area[5] $uv\big]_{(1)}^{(2)}$ minus the grey area $\int_{(1)}^{(2)} v\,\mathrm{d}u$.

- Since $\int \mathrm{d}v = v + C_1$, $\int u\,\mathrm{d}v$ actually equals $u(v+C_1) - \int (v+C_1)\,\mathrm{d}u$. However, since

$$\begin{aligned} u(v+C_1) - \int (v+C_1)\,\mathrm{d}u &= uv + C_1 u - \int v\,\mathrm{d}u - \int C_1\,\mathrm{d}u \\ &= uv - \int v\,\mathrm{d}u \end{aligned}$$

  it is customary to drop the first constant of integration.

  - That being said, it is sometimes useful — when evaluating $\int \ln(x+1)\,\mathrm{d}x = \ln(x+1)(x+C_1) - \int \frac{x+C_1}{x+1}\,\mathrm{d}x$, being able to choose $C_1 = 1$ greatly simplifies the second integral.

- When integrating an inverse trig function, consider using integration by parts.

  - For example, when confronted with $\int \tan^{-1} x\,\mathrm{d}x$, integration by parts turns it into an inverse trig derivative problem.

$$\int \tan^{-1} x\,\mathrm{d}x = x\tan^{-1} x - \int \frac{x\,\mathrm{d}x}{1+x^2}$$

- When attempting integration by parts, don't be afraid to use it multiple times.

  - For example, when confronted with $\int x^2 \mathrm{e}^x\,\mathrm{d}x$, use integration by parts twice.

$$\begin{aligned} \int x^2\mathrm{e}^x\,\mathrm{d}x &= x^2\mathrm{e}^x - \int 2x\mathrm{e}^x\,\mathrm{d}x \\ &= x^2\mathrm{e}^x - \left(2x\mathrm{e}^x - \int 2\mathrm{e}^x\,\mathrm{d}x\right) \end{aligned}$$

- When attempting integration by parts, look for the original integral showing up again — if it does, combine like terms.

- When integrating powers of $\cos x$, consider using a reduction formula.

[5] Note that $uv\big]_{(1)}^{(2)} = u_2v_2 - u_1v_1$, the latter of which, as the difference of two rectangles, clearly represents the total shaded area.

– Begin by deriving a reduction formula (this will involve splitting the exponent!).

$$\begin{aligned}
\int \cos^n x \,\mathrm{d}x &= \int \cos^{n-1} x \cos x \,\mathrm{d}x \\
&= \cos^{n-1} x \sin x + (n-1) \int \cos^{n-2} x \sin^2 x \,\mathrm{d}x \\
&= \cos^{n-1} x \sin x + (n-1) \int \cos^{n-2} x \,\mathrm{d}x - (n-1) \int \cos^n x \,\mathrm{d}x \\
(1+(n-1)) \int \cos^n x \,\mathrm{d}x &= \cos^{n-1} x \sin x + (n-1) \int \cos^{n-2} x \,\mathrm{d}x \\
\int \cos^n x \,\mathrm{d}x &= \frac{\cos^{n-1} x \sin x}{n} + \frac{n-1}{n} \int \cos^{n-2} x \,\mathrm{d}x
\end{aligned}$$

Since the reduction formula decreases the exponent by 2, work out two base cases.

$$\int \cos^0 x \,\mathrm{d}x = x + C$$
$$\int \cos^1 x \,\mathrm{d}x = \sin x + C$$

- When integrating powers of $\sin x$, consider using a reduction formula.

$$\int \sin^n x \,\mathrm{d}x = -\frac{\sin^{n-1} x \cos x}{n} + \frac{n-1}{n} \int \sin^{n-2} x \,\mathrm{d}x$$

9.8 Integration of Rational Functions of $\sin x$ and $\cos x$, and Other Trigonometric Integrals

- "It has been discovered that the substitution $z = \tan \frac{x}{2}$ enables one to reduce the problem of integrating any rational function of $\sin x$ and $\cos x$ to a problem involving a rational function of z. This in turn can be integrated by the method of partial fractions" (Thomas, 1972, p. 300).
  - However, this substitution should be used only as a last resort — the associated algebra is often quite cumbersome.
- To increase the ease of use for this substitution, it will help to derive three results.

$$\cos x = \frac{1-z^2}{1+z^2} \qquad \sin x = \frac{2z}{1+z^2} \qquad \mathrm{d}x = \frac{2\,\mathrm{d}z}{1+z^2}$$

- "The following types of integrals. . . arise in connection with alternating-current theory, heat transfer problems, bending of beams, cable stress analysis in suspension bridges, and many other places where trigonometric series (or Fourier series) are applied to problems in mathematics, science, and engineering" (Thomas, 1972, p. 301).

$$\int \sin mx \sin nx \,\mathrm{d}x \qquad \int \sin mx \cos nx \,\mathrm{d}x \qquad \int \cos mx \cos nx \,\mathrm{d}x$$

- When confronted with one of these integrals, integration by parts may be used. However, using one of the following three identities will be more simple.

$$\begin{aligned}
\sin mx \sin nx &= \frac{1}{2}(\cos(m-n)x - \cos(m+n)x) \\
\sin mx \cos nx &= \frac{1}{2}(\sin(m-n)x + \sin(m+n)x) \\
\cos mx \cos nx &= \frac{1}{2}(\cos(m-n)x + \cos(m+n)x)
\end{aligned}$$

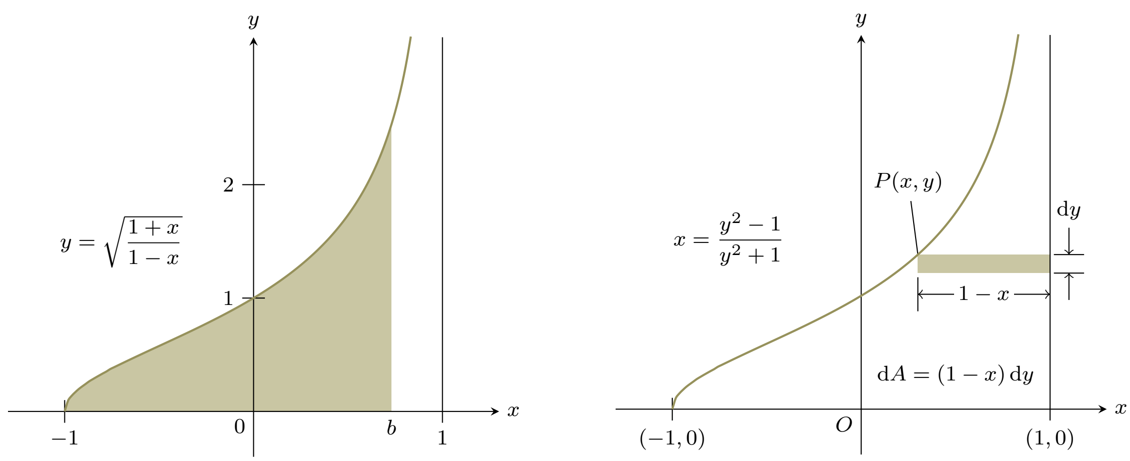
- Note that "these identities follow at once from $\cos(A+B) = \cos A\cos B - \sin A\sin B$, $\cos(A-B) = \cos A\cos B + \sin A\sin B$, and $\sin(A+B) = \sin A\cos B + \cos A\sin B$, $\sin(A-B) = \sin A\cos B - \cos A\sin B$" (Thomas, 1972, p. 301).

9.9 Further Substitutions

- "Some integrals involving fractional powers of the variable x may be simplified by substituting $x = z^n$, where n is the least common multiple of the denominators of the exponents" (Thomas, 1972, p. 302).
  - For example, $\int \frac{\sqrt{x}\,\mathrm{d}x}{1+\sqrt[4]{x}}$ can be simplified by taking $x = z^4$.
- "Even when it is not clear at the start that a substitution will work, it is advisable to try one that seems reasonable and pursue it until it either gives results or appears to make matters worse. In the latter case, try something else! Sometimes a chain of substitutions $u = f(x)$, $v = g(u)$, $z = h(v)$, and so on, will produce results when it is by no means obvious that this will work" (Thomas, 1972, p. 302).
- "The criterion of success is whether the new integrals so obtained appear to be simpler than the original integral. Here it is handy to remember that any rational function of x can be integrated by the method of partial fractions and that any rational function of $\sin x$ and $\cos x$ can be integrated by using the substitution $z = \tan\frac{x}{2}$. If we can reduce a given integral to one of these types, we then know how to finish the job" (Thomas, 1972, p. 302).
- To evaluate a definite integral after a (series of) substitution(s), either return the substitution(s) and keep the bounds or keep the substitution(s) and determine new bounds based on the new variable of integration.
  - For example, if $z^2 = \frac{1+x}{1-x}$ and $x \in [-1,1]$, then $z \in [0,\infty)$. And if $z = \tan\theta$, then $\theta \in \left[0, \frac{\pi}{2}\right]$.

9.10 Improper Integrals

- **Improper integral**: An integral of the form $\int_a^b f(x)\,\mathrm{d}x$ where some $x \in [a,b]$ is infinite, and/or one or both of a, b are infinite in magnitude.



(a) With respect to x.

(b) With respect to y.

Figure 9.4: Defining improper integrals.

- Say we wish to evaluate $\int_{-1}^{1}\sqrt{\frac{1+x}{1-x}}\,\mathrm{d}x$, knowing that the integrand approaches ∞ as $x \to 1$ (see Figure 9.4a). Well, if the upper bound b is some value slightly *less than* 1, we *can* evaluate the integral. Thus evaluating the original integral becomes a problem of evaluating

$$\lim_{b\to 1^-}\int_{-1}^{b}\sqrt{\frac{1+x}{1-x}}\,\mathrm{d}x = \lim_{b\to 1^-}\left(\sin^{-1}x - \sqrt{1-x^2}\Big]_{-1}^{b}\right) = \lim_{b\to 1^-}\left(\sin^{-1}b - \sqrt{1-b^2} + \frac{\pi}{2}\right)$$

  - Sometimes such a limit will converge. Sometimes it will not (it will diverge). Either way, it answers the question of the nature of the area under the curve (by yielding some finite value, or the infinite one).
  - Note that the integral works out just the same if we sum vertical elements instead (see Figure 9.4b), evaluating the following.

$$\lim_{c\to\infty}\int_0^c (1-x)\,\mathrm{d}y = \lim_{c\to\infty}\int_0^c \frac{2\,\mathrm{d}y}{y^2+1}$$

- When integrating a function $f(x)$ on $[a,b]$ where $f(x) \to \infty$ at some x-value $c \in (a,b)$, split the integral.

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{c\to c^-}\int_a^c f(x)\,\mathrm{d}x + \lim_{c\to c^+}\int_c^b f(x)\,\mathrm{d}x$$

- On determining whether or not an improper integral with a nonintegrable integrand exists, we can sometimes compare it with an integral that we know.
  - For example,

$$\int_1^b \mathrm{e}^{-x^2}\,\mathrm{d}x \leq \int_1^b \mathrm{e}^{-x}$$

    for all $b \geq 1$ since $0 < \mathrm{e}^{-x^2} \leq \mathrm{e}^{-x}$ for all $x \geq 1$. Thus, $\int_1^\infty \mathrm{e}^{-x^2}\,\mathrm{d}x$ evaluates to some finite value.

- Note that some improper integrals diverge by oscillation.
  - For example, $\int_0^\infty \cos x\,\mathrm{d}x$ diverges in this manner.

9.11 Numerical Methods for Approximating Definite Integrals

- One could use the **trapezoidal rule**.
- A better choice, though, is **Simpson's rule**.

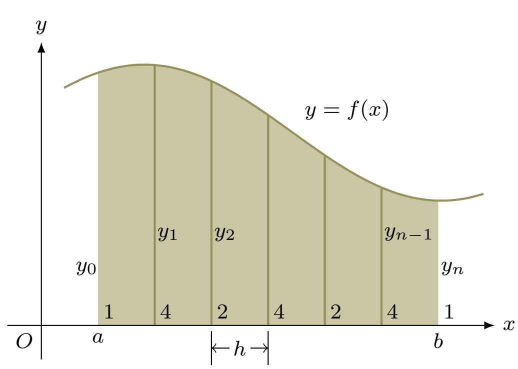


Figure 9.5: Simpson's rule.

- Simpson's rule approximates the curve via parabolas, which can be uniquely defined by three points and have a nice formula for the area underneath them.
- We now derive Simpson's rule.
  - Suppose we wish to approximate the area under the part of the curve from x_i to x_{i+2} in Figure 9.5. We know that there exists some parabola intersecting (x_i, y_0), (x_{i+1}, y_1), and (x_{i+2}, y_2). However, for the sake of simplifying the algebra, we choose to consider the parabola intersecting $(-h, y_0)$, $(0, y_1)$, and (h, y_2) (the area under both parabolas will be equivalent since $x_{i+1} - x_i = h$). Let this translated parabola be called Ax^2+Bx+C for some $A, B, C \in \mathbb{R}$. Then the area underneath this parabola A_p can be described by the following.

$$\begin{aligned} A_p &= \int_{-h}^{h} \left(Ax^2 + Bx + C\right) \mathrm{d}x \\ &= \frac{2Ah^3}{3} + 2Ch \end{aligned}$$

  - We know that the area underneath this parabola is dependent only on h, y_0, y_1, and y_2. Thus, we look to express A, C from the integral in terms of h, y_0, y_1, y_2. To accomplish this, we will use the facts that

$$\begin{aligned} y_0 &= Ah^2 - Bh + C \\ y_1 &= C \\ y_2 &= Ah^2 + Bh + C \end{aligned}$$

  We can now see that $C = y_1$, so all that's left is to solve for A. This can be done by adding the first and third equations, substituting, and solving as follows.

$$\begin{aligned} y_0 + y_2 &= 2Ah^2 + 2C \\ 2Ah^2 &= y_0 + y_2 - 2y_1 \\ A &= \frac{y_0 - 2y_1 + y_2}{2h^2} \end{aligned}$$

  - Thus, we can reformulate the area under the parabola as follows.

$$\begin{aligned} A_p &= \frac{2h^3}{3} \cdot \frac{y_0 - 2y_1 + y_2}{2h^2} + 2y_1 h \\ &= \frac{h}{3}(y_0 + 4y_1 + y_2) \end{aligned}$$

  - "Simpson's rule follows from applying this result to successive pieces of the curve $y = f(x)$ between $x = a$ and $x = b$. Each separate piece of the curve, covering an x-subinterval of width $2h$, is approximated by an arc of a parabola through its ends and its mid-point. The area under each parabolic arc is then given by an expression like [the above] and the results are added to give [the following]" (Thomas, 1972, p. 309).

$$\begin{aligned} A_S &= \frac{h}{3}((y_0 + 4y_1 + y_2) + (y_2 + 4y_3 + y_4) + \cdots + (y_{n-2} + 4y_{n-1} + y_n)) \\ &= \frac{h}{3}(y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \cdots + 2y_{n-2} + 4y_{n-1} + y_n) \end{aligned}$$

- Note that the number of subdivisions must be an even integer.

- To find the error in a Simpson's rule approximation, use the fact[6] that "if f is continuous on $[a, b]$ and four times differentiable on (a, b), then there is a number c between a and b such that [the following holds]" (Thomas, 1972, p. 310).

$$\int_a^b f(x)\,\mathrm{d}x = A_S - \frac{b-a}{180} f^{(4)}(c) \cdot h^4$$

[6]A proof of this fact is a topic best left until Analysis. The framework for such a proof may be found on Olmsted (1956, p. 146).

Chapter 14

Vector Functions and Their Derivatives

14.1 Introduction

12/14:

- **Vector function** (of h): A function $\mathbf{F}(h)$ with n components where each component is a function. Essentially, $\mathbf{F} = (f_1, f_2, \dots, f_n)$.
- **Limit** (of $\mathbf{F}(h)$ as $h \to a$): If each component $f_1, \dots, f_n$ of $\mathbf{F}$ has a limit $L_1, \dots, L_n$ as $h \to a$, then

$$\lim_{h \to a} \mathbf{F}(h) = (L_1, \dots, L_n)$$

- **Continuous** (vector function $\mathbf{F}$ at a): A vector function $\mathbf{F}$ where for every $\epsilon > 0$, there corresponds a $\delta > 0$ such that

$$|\mathbf{F}(h) - \mathbf{F}(a)| < \epsilon \quad \text{when} \quad |h - a| < \delta$$

  - Thomas (1972) shows that this is equivalent to the requirement that each component of $\mathbf{F}$ is continuous at a.
- **Derivative** (of a vector function at c): The derivative $\mathbf{F}'(c)$ of a vector function $\mathbf{F}$ at c is given by the equation

$$\mathbf{F}'(c) = \lim_{h \to 0} \frac{\mathbf{F}(c + h) - \mathbf{F}(c)}{h}$$

  - It can be proven that $\mathbf{F}$ is differentiable at c if and only if each of its components are differentiable at c, and that if this condition is met,

$$\mathbf{F}'(c) = (f_1'(c), \dots, f_n'(c))$$

14.2 Velocity and Acceleration

- Results from here on out will generally pertain to 2D questions, but these methods can easily be generalized to higher dimensions.
- Applications of vectors to physics problems.
  - To solve **statics** problems, we only need to know the **algebra** of vectors.
  - To solve **dynamics** problems, we also need to know the **calculus** of vectors.
- **Position vector**: The vector from the origin to a point P that moves along a parametrically defined curve. *Denoted by* $\mathbf{R}$.

- **Velocity vector**: The vector tangent to a point P that moves along a parametrically defined curve and with magnitude $|\mathrm{d}s/\mathrm{d}t|$. *Denoted by* $\mathbf{v}$.

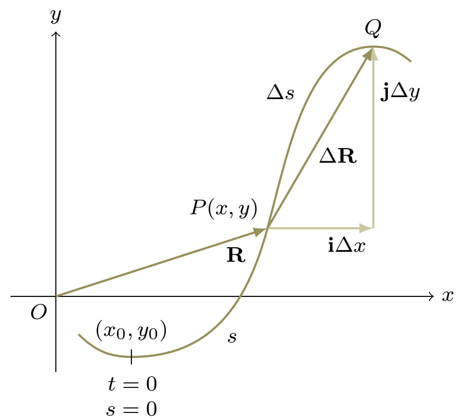


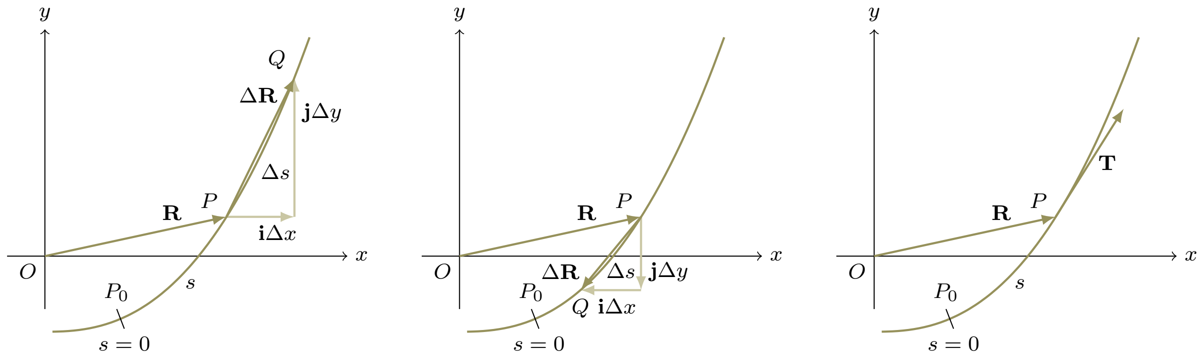
Figure 14.1: Velocity vector.

  - Thomas (1972) semi-rigorously proves from Figure 14.1 that if $\mathbf{R}$ is the position vector, then $\mathrm{d}\mathbf{R}/\mathrm{d}t$ is the velocity vector.
  - Essentially, he proves that
$$\frac{\mathrm{d}\mathbf{R}}{\mathrm{d}t} = \mathbf{i}\frac{\mathrm{d}x}{\mathrm{d}t} + \mathbf{j}\frac{\mathrm{d}y}{\mathrm{d}t}$$
It follows from this that
$$\text{slope of } \frac{\mathrm{d}\mathbf{R}}{\mathrm{d}t} = \frac{\mathbf{j}\text{-component}}{\mathbf{i}\text{-component}} = \frac{\mathrm{d}y/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t} = \frac{\mathrm{d}y}{\mathrm{d}x}$$
$$\left|\frac{\mathrm{d}R}{\mathrm{d}t}\right| = \left|\mathbf{i}\frac{\mathrm{d}x}{\mathrm{d}t} + \mathbf{j}\frac{\mathrm{d}y}{\mathrm{d}t}\right| = \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2} = \left|\frac{\mathrm{d}s}{\mathrm{d}t}\right|$$

- **Acceleration vector**: The derivative of the velocity vector and second derivative of the position vector. *Denoted by* $\mathbf{a}$.
$$\mathbf{a} = \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \mathbf{i}\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + \mathbf{j}\frac{\mathrm{d}^2y}{\mathrm{d}t^2}$$
- Sometimes, we are given a force vector $\mathbf{F} = m\mathbf{a}$ and initial conditions.
  - From these, we can solve for velocity and position vectors via fairly straightforward component integration.
  - Note, however, that constants of integration are now vectors.

14.3 Tangential Vectors

- Let P_0 be a point on a curve. The distance s from P_0 to some point P along the curve is clearly related to the position of P. Thus, we may think of $\mathbf{R}$ as a function of s, and investigate the properties of $\mathrm{d}\mathbf{R}/\mathrm{d}s$.
- **Tangent vector**: The unit vector tangent to a point P along a curve.
  - Since $\Delta\mathbf{R}$ and Δs approach the same quantity as $\Delta s \to 0$, $\Delta\mathbf{R}/\Delta s$ approaches unity, i.e., $|\mathrm{d}\mathbf{R}/\mathrm{d}s| = 1$.
  - Because of the sign change, whether Δs is positive or negative, $\Delta\mathbf{R}/\Delta s$ points in the same general direction for sufficiently small Δs. Indeed, it converges to pointing tangentially.



(a) $\Delta s > 0$.   (b) $\Delta s < 0$.   (c) $\Delta s \to 0$.

Figure 14.2: Tangent vector.

- Thus,

$$\mathbf{T} = \frac{\mathrm{d}\mathbf{R}}{\mathrm{d}s} = \mathbf{i}\frac{\mathrm{d}x}{\mathrm{d}s} + \mathbf{j}\frac{\mathrm{d}y}{\mathrm{d}s}$$

- There are two different ways to find $\mathbf{T}$: Straight differentiation combined with manipulations of differentials, and the chain rule combined with the dot product. We will explore each, in turn, with an example.
- "Find the unit vector $\mathbf{T}$ tangent to the circle $x = a\cos\theta$, $y = a\sin\theta$ at any point $P(x, y)$" (Thomas, 1972, p. 471).
  - From the given equations, we have

$$\mathrm{d}x = -a\sin\theta\,\mathrm{d}\theta \qquad \mathrm{d}y = a\cos\theta\,\mathrm{d}\theta \qquad \begin{aligned} \mathrm{d}s^2 &= \mathrm{d}x^2 + \mathrm{d}y^2 \\ &= a^2(\sin^2\theta + \cos^2\theta)\,\mathrm{d}\theta^2 \\ &= a^2\,\mathrm{d}\theta^2 \\ \mathrm{d}s &= \pm a\,\mathrm{d}\theta \end{aligned}$$

    - We could alternatively obtain $\mathrm{d}s$ by expressing the arc length formula $S = R\theta$ in terms of differentials.
  - "If we measure arc length in the counterclockwise direction, with $s = 0$ at $(a, 0)$, s will be an increasing function of θ, so the +-sign should be taken: $\mathrm{d}s = a\,\mathrm{d}\theta$" (Thomas, 1972, p. 471).
  - Therefore,

$$\begin{aligned} \mathbf{T} &= \mathbf{i}\frac{\mathrm{d}x}{\mathrm{d}s} + \mathbf{j}\frac{\mathrm{d}y}{\mathrm{d}s} \\ &= \mathbf{i}\left(\frac{-a\sin\theta\,\mathrm{d}\theta}{a\,\mathrm{d}\theta}\right) + \mathbf{j}\left(\frac{a\cos\theta\,\mathrm{d}\theta}{a\,\mathrm{d}\theta}\right) \\ &= -\mathbf{i}\sin\theta + \mathbf{j}\cos\theta \end{aligned}$$

- The equations

$$x = a\cos\omega t \qquad y = a\sin\omega t \qquad z = bt$$

  where a, b, ω are positive constants define a circular helix in E^3[1].

[1] Three-dimensional Euclidean space, equivalent to $\mathbb{R}^3$

- Let $P_0 = (a, 0, 0)$, since this is the point on the locus of the parametric equations where $t = 0$. Additionally, let arc length be measured in the direction in which P moves away from P_0 as t increases from 0.
- Using the chain rule to differentiate, we have

$$\begin{aligned}\mathbf{T} &= \mathbf{i}\frac{\mathrm{d}x}{\mathrm{d}s} + \mathbf{j}\frac{\mathrm{d}y}{\mathrm{d}s} + \mathbf{k}\frac{\mathrm{d}z}{\mathrm{d}s} \\ &= \mathbf{i}\left(-a\omega\sin\omega t\,\frac{\mathrm{d}t}{\mathrm{d}s}\right) + \mathbf{j}\left(a\omega\cos\omega t\,\frac{\mathrm{d}t}{\mathrm{d}s}\right) + \mathbf{k}\left(b\,\frac{\mathrm{d}t}{\mathrm{d}s}\right)\end{aligned}$$

- Since $\mathbf{T}$ is a unit vector, we have $1 = |\mathbf{T}| = |\mathbf{T}|^2 = \mathbf{T}\cdot\mathbf{T}$. Thus,

$$\begin{aligned}1 &= \mathbf{T}\cdot\mathbf{T} \\ &= \mathbf{i}\cdot\mathbf{i}\left(-a\omega\sin\omega t\,\frac{\mathrm{d}t}{\mathrm{d}s}\right)^2 + \mathbf{j}\cdot\mathbf{j}\left(a\omega\cos\omega t\,\frac{\mathrm{d}t}{\mathrm{d}s}\right)^2 + \mathbf{k}\cdot\mathbf{k}\left(b\,\frac{\mathrm{d}t}{\mathrm{d}s}\right)^2 \\ &= \left(a^2\omega^2 + b^2\right)\left(\frac{\mathrm{d}t}{\mathrm{d}s}\right)^2 \\ \frac{\mathrm{d}t}{\mathrm{d}s} &= \pm\frac{1}{\sqrt{a^2\omega^2 + b^2}}\end{aligned}$$

- We choose the $+$-sign because s should be a positive function of t.
- Putting this all together, we get

$$\mathbf{T} = \frac{a\omega(-\mathbf{i}\sin\omega t + \mathbf{j}\cos\omega t) + \mathbf{k}b}{\sqrt{a^2\omega^2 + b^2}}$$

14.4 Curvature and Normal Vectors

12/15:
- **Curvature**: The rate of change of the slope angle ϕ between $\mathbf{T}$ and the x-axis with respect to the arc length s. *Denoted by* κ.

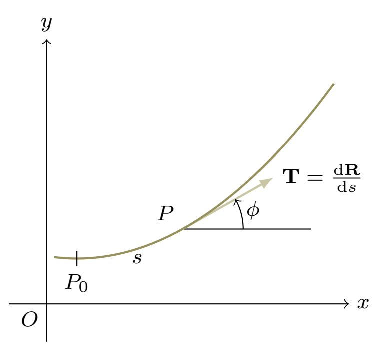


Figure 14.3: Curvature.

  - Measured in radians per unit length.
- From the facts that

$$\kappa = \frac{\mathrm{d}\phi}{\mathrm{d}s} \qquad \tan\phi = \frac{\mathrm{d}y}{\mathrm{d}x} \qquad \mathrm{d}s = \pm\sqrt{\mathrm{d}x^2 + \mathrm{d}y^2}$$

we can derive a formula for κ in terms of the original function $y = f(x)$ as follows.

$$\phi = \tan^{-1}\frac{\mathrm{d}y}{\mathrm{d}x} \qquad\qquad \frac{\mathrm{d}s}{\mathrm{d}x} = \pm\sqrt{1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}$$

$$\frac{\mathrm{d}\phi}{\mathrm{d}x} = \frac{\frac{\mathrm{d}^2y}{\mathrm{d}x^2}}{1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}$$

$$\begin{aligned}\kappa &= \left|\frac{\mathrm{d}\phi}{\mathrm{d}s}\right| = \left|\frac{\mathrm{d}\phi/\mathrm{d}x}{\mathrm{d}s/\mathrm{d}x}\right| \\ &= \frac{\left|\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\right|}{\left[1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2\right]^{3/2}}\end{aligned}$$

- We can similarly derive that

$$\kappa = \frac{\left|\frac{\mathrm{d}^2x}{\mathrm{d}y^2}\right|}{\left[1+\left(\frac{\mathrm{d}x}{\mathrm{d}y}\right)^2\right]^{3/2}}$$

- If the equations for y and x are given parametrically in terms of t, we have

$$\begin{aligned}\kappa &= \frac{\left|\frac{\mathrm{d}x}{\mathrm{d}t}\frac{\mathrm{d}^2y}{\mathrm{d}t^2} - \frac{\mathrm{d}y}{\mathrm{d}t}\frac{\mathrm{d}^2x}{\mathrm{d}t^2}\right|}{\left[\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2\right]^{3/2}} \\ &= \frac{|\dot{x}\ddot{y} - \dot{y}\ddot{x}|}{[\dot{x}^2+\dot{y}^2]^{3/2}}\end{aligned}$$

- Naturally, the curvature of a straight line should be 0. Indeed, we find this from the above equations.
- Naturally, the curvature of a circle should be constant, and should somehow decrease as the radius increases. Indeed, we find from the facts that $s = r\theta$ and $\phi = \theta + \frac{\pi}{2}$ that

$$\kappa = \left|\frac{\mathrm{d}\phi}{\mathrm{d}s}\right| = \left|\frac{\mathrm{d}\theta}{r\,\mathrm{d}\theta}\right| = \frac{1}{r}$$

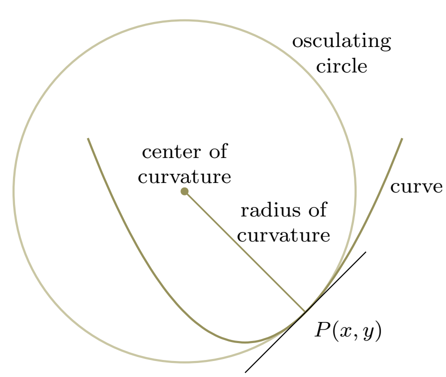


Figure 14.4: Circle, radius, and center of curvature.

- **Circle of curvature** (at P): "The circle that is tangent to a given curve at P, whose center lies on the concave side of the curve and which has the same curvature as the curve has at P" (Thomas, 1972, p. 475). *Also known as* **osculating circle**.
  - Calling the circle of curvature the "osculating circle" refers to the fact that its first and second derivatives at P are equal to the first and second derivatives of the curve at P, meaning that it has a higher degree of contact with the curve at P than any other circle.
- **Radius of curvature** (at P): The radius of the circle of curvature at P. *Denoted by* ρ.

$$\rho = \frac{1}{\kappa} = \frac{\left[1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2\right]^{3/2}}{\left|\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\right|}$$

- **Normal vector**: The unit vector normal to a point P along a curve.
  - Observe that $\mathbf{T}$ can be expressed in terms of the slope angle ϕ:

$$\mathbf{T} = \mathbf{i}\cos\phi + \mathbf{j}\sin\phi$$

  - Since $\mathbf{T}$ can be thought of as a function of ϕ, we can investigate the properties of $\mathrm{d}\mathbf{T}/\mathrm{d}\phi$.
  - Indeed, it is not difficult to show that

$$\frac{\mathrm{d}\mathbf{T}}{\mathrm{d}\phi} = -\mathbf{i}\sin\phi + \mathbf{j}\cos\phi \qquad \left|\frac{\mathrm{d}\mathbf{T}}{\mathrm{d}\phi}\right| = \sqrt{\sin^2\phi + \cos^2\phi} = 1 \qquad \mathbf{T}\cdot\frac{\mathrm{d}\mathbf{T}}{\mathrm{d}\phi} = 0$$

  - Thus,

$$\mathbf{N} = \frac{\mathrm{d}\mathbf{T}}{\mathrm{d}\phi} = -\mathbf{i}\sin\phi + \mathbf{j}\cos\phi$$

- In 3D space, it is harder to define a single normal vector, so we define the...
- **Principal normal vector**: The vector

$$\mathbf{N} = \frac{\mathrm{d}\mathbf{T}/\mathrm{d}s}{|\mathrm{d}\mathbf{T}/\mathrm{d}s|}$$

  - We will soon prove that $\mathbf{T}\cdot\mathrm{d}\mathbf{T}/\mathrm{d}s = 0$.
  - If ϕ is an increasing function of s, then by the chain rule,

$$\frac{\mathrm{d}\mathbf{T}}{\mathrm{d}s} = \frac{\mathrm{d}\mathbf{T}}{\mathrm{d}\phi}\frac{\mathrm{d}\phi}{\mathrm{d}s} = \mathbf{N}\kappa$$

  - Since $\mathbf{N}$ is a unit vector and κ is a constant, κ is equal to the magnitude of $\mathrm{d}\mathbf{T}/\mathrm{d}s$.
  - Thus, we can define the principal normal as above.
- Thomas (1972) uses $\mathrm{d}\mathbf{T}/\mathrm{d}s$ to find both the curvature and principal normal vector of the general circular helix investigated earlier. He also checks limiting cases to rederive the curvature of a circle and of a straight line.
- **Binormal vector**: The vector perpendicular to both $\mathbf{T}$ and $\mathbf{N}$, as defined by

$$\mathbf{B} = \mathbf{T}\times\mathbf{N}$$

14.5 Differentiation of Products of Vectors

- Let $\mathbf{U}$ and $\mathbf{V}$ be vectors whose components are differentiable functions of t.
- Then we can verify by components that

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{U}\cdot\mathbf{V}) = \frac{\mathrm{d}\mathbf{U}}{\mathrm{d}t}\cdot\mathbf{V} + \mathbf{U}\cdot\frac{\mathrm{d}\mathbf{V}}{\mathrm{d}t} \qquad\qquad \frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{U}\times\mathbf{V}) = \frac{\mathrm{d}\mathbf{U}}{\mathrm{d}t}\times\mathbf{V} + \mathbf{U}\times\frac{\mathrm{d}\mathbf{V}}{\mathrm{d}t}$$

  - Note that we can *derive* the above, too, through the Δ-process.
  - Let $\mathbf{W} = \mathbf{U}\times\mathbf{V}$. Then

$$\begin{aligned}
\mathbf{W} + \Delta\mathbf{W} &= (\mathbf{U} + \Delta\mathbf{U}) \times (\mathbf{V} + \Delta\mathbf{V}) \\
&= \mathbf{U}\times\mathbf{V} + \mathbf{U}\times\Delta\mathbf{V} + \Delta\mathbf{U}\times\mathbf{V} + \Delta\mathbf{U}\times\Delta\mathbf{V} \\
\Delta\mathbf{W} &= \mathbf{U}\times\Delta\mathbf{V} + \Delta\mathbf{U}\times\mathbf{V} + \Delta\mathbf{U}\times\Delta\mathbf{V} \\
\frac{\Delta\mathbf{W}}{\Delta t} &= \mathbf{U}\times\frac{\Delta\mathbf{V}}{\Delta t} + \frac{\Delta\mathbf{U}}{\Delta t}\times\mathbf{V} + \frac{\Delta\mathbf{U}}{\Delta t}\times\Delta\mathbf{V} \\
\frac{\mathrm{d}\mathbf{W}}{\mathrm{d}t} &= \mathbf{U}\times\frac{\mathrm{d}\mathbf{V}}{\mathrm{d}t} + \frac{\mathrm{d}\mathbf{U}}{\mathrm{d}t}\times\mathbf{V}
\end{aligned}$$

- Differentiating the triple scalar product:

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{U}\cdot\mathbf{V}\times\mathbf{W}) = \frac{\mathrm{d}\mathbf{U}}{\mathrm{d}t}\cdot\mathbf{V}\times\mathbf{W} + \mathbf{U}\cdot\frac{\mathrm{d}\mathbf{V}}{\mathrm{d}t}\times\mathbf{W} + \mathbf{U}\cdot\mathbf{V}\times\frac{\mathrm{d}\mathbf{W}}{\mathrm{d}t}$$

  - Equivalently,

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = \begin{vmatrix} \frac{\mathrm{d}u_1}{\mathrm{d}t} & \frac{\mathrm{d}u_2}{\mathrm{d}t} & \frac{\mathrm{d}u_3}{\mathrm{d}t} \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} + \begin{vmatrix} u_1 & u_2 & u_3 \\ \frac{\mathrm{d}v_1}{\mathrm{d}t} & \frac{\mathrm{d}v_2}{\mathrm{d}t} & \frac{\mathrm{d}v_3}{\mathrm{d}t} \\ w_1 & w_2 & w_3 \end{vmatrix} + \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ \frac{\mathrm{d}w_1}{\mathrm{d}t} & \frac{\mathrm{d}w_2}{\mathrm{d}t} & \frac{\mathrm{d}w_3}{\mathrm{d}t} \end{vmatrix}$$

- Differentiating $\mathbf{V}\cdot\mathbf{V} = |\mathbf{V}|^2$ gives

$$\begin{aligned}
\mathbf{V}\cdot\frac{\mathrm{d}\mathbf{V}}{\mathrm{d}t} + \frac{\mathrm{d}\mathbf{V}}{\mathrm{d}t}\cdot\mathbf{V} &= 0 \\
2\mathbf{V}\cdot\frac{\mathrm{d}\mathbf{V}}{\mathrm{d}t} &= 0
\end{aligned}$$

  - Thus, for any vector function $\mathbf{V}$, there are three cases: (1) $\mathbf{V} = \mathbf{0}$, (2) $\mathrm{d}\mathbf{V}/\mathrm{d}t = \mathbf{0}$, so $\mathbf{V}$ is constant in both direction and magnitude, and (3) $\mathbf{V}$ and $\mathrm{d}\mathbf{V}/\mathrm{d}t$ are perpendicular.
  - Note that this fact allows to verify that $\mathbf{T}\cdot\mathrm{d}\mathbf{T}/\mathrm{d}s = 0$.

- We can use the calculus of tangential and normal vectors to break velocity and acceleration vectors into tangential and normal components.

$$\begin{aligned}
\mathbf{v} &= \frac{\mathrm{d}\mathbf{R}}{\mathrm{d}t} & \mathbf{a} &= \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} & \mathbf{a} &= \mathbf{T}\frac{\mathrm{d}^2s}{\mathrm{d}t^2} + \frac{\mathrm{d}s}{\mathrm{d}t}\frac{\mathrm{d}\mathbf{T}}{\mathrm{d}s}\frac{\mathrm{d}s}{\mathrm{d}t} \\
&= \frac{\mathrm{d}\mathbf{R}}{\mathrm{d}s}\frac{\mathrm{d}s}{\mathrm{d}t} & &= \mathbf{T}\frac{\mathrm{d}^2s}{\mathrm{d}t^2} + \frac{\mathrm{d}s}{\mathrm{d}t}\frac{\mathrm{d}\mathbf{T}}{\mathrm{d}t} & &= \mathbf{T}\frac{\mathrm{d}^2s}{\mathrm{d}t^2} + \mathbf{N}\kappa\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 \\
&= \mathbf{T}\frac{\mathrm{d}s}{\mathrm{d}t} & & & &= \mathbf{T}\frac{\mathrm{d}^2s}{\mathrm{d}t^2} + \mathbf{N}\left(\frac{\mathbf{v}^2}{\rho}\right)
\end{aligned}$$

  - $\mathbf{v}^2/\rho$ is very similar to v^2/r (think about the circle and radius of curvature)!
  - Another important related equation:

$$|\mathbf{a}| = a_T^2 + a_N^2$$

– Lastly, we can derive a formula for the curvature in terms of velocity and acceleration.

$$\begin{aligned}\mathbf{v}\times\mathbf{a} &= \mathbf{T}\frac{\mathrm{d}s}{\mathrm{d}t}\times\left[\mathbf{T}\frac{\mathrm{d}^2s}{\mathrm{d}t^2}+\mathbf{N}\kappa\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2\right]\\ &= \mathbf{T}\times\mathbf{N}\kappa\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^3\\ |\mathbf{v}\times\mathbf{a}| &= \left|\mathbf{B}\kappa|\mathbf{v}|^3\right|\\ \kappa &= \frac{|\mathbf{v}\times\mathbf{a}|}{|\mathbf{v}|^3}\end{aligned}$$

14.6 Polar and Cylindrical Coordinates

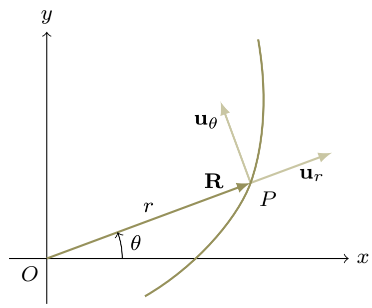


Figure 14.5: Vectors in polar coordinates.

- To analyze polar coordinates, we introduce the unit vectors

$$\mathbf{u}_r=\mathbf{i}\cos\theta+\mathbf{j}\sin\theta \qquad\qquad \mathbf{u}_\theta=-\mathbf{i}\sin\theta+\mathbf{j}\cos\theta$$

- Clearly, we have

$$\frac{\mathrm{d}\mathbf{u}_r}{\mathrm{d}\theta}=\mathbf{u}_\theta \qquad\qquad \frac{\mathrm{d}\mathbf{u}_\theta}{\mathrm{d}\theta}=-\mathbf{u}_r$$

- Additionally, we can see that

$$\mathbf{R}=r\mathbf{u}_r$$

- The velocity vector can easily be expressed in terms of these quantities (and visualized as such geometrically, as in Figure 14.6).

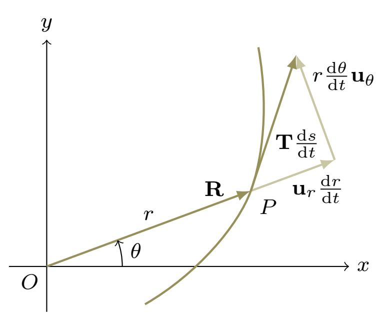


Figure 14.6: Polar velocity vector.

$$\begin{aligned}\mathbf{v} &= \frac{\mathrm{d}\mathbf{R}}{\mathrm{d}t} \\ &= \mathbf{u}_r \frac{\mathrm{d}r}{\mathrm{d}t} + r\frac{\mathrm{d}\mathbf{u}_r}{\mathrm{d}t} \\ &= \mathbf{u}_r \frac{\mathrm{d}r}{\mathrm{d}t} + r\mathbf{u}_\theta \frac{\mathrm{d}\theta}{\mathrm{d}t}\end{aligned}$$

- The acceleration vector can also be expressed in terms of these quantities (the following can be derived by differentiating the above with respect to t and substituting).

$$\mathbf{a} = \mathbf{u}_r \left[\frac{\mathrm{d}^2 r}{\mathrm{d}t^2} - r\left(\frac{\mathrm{d}\theta}{\mathrm{d}t}\right)^2\right] + \mathbf{u}_\theta \left[r\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} + 2\frac{\mathrm{d}r}{\mathrm{d}t}\frac{\mathrm{d}\theta}{\mathrm{d}t}\right]$$

- In three dimensions (esp. for cylindrical coordinates), we have

$$\begin{aligned}\mathbf{R} &= r\mathbf{u}_r + \mathbf{k}z \\ \mathbf{v} &= \mathbf{u}_r \frac{\mathrm{d}r}{\mathrm{d}t} + r\mathbf{u}_\theta \frac{\mathrm{d}\theta}{\mathrm{d}t} + \mathbf{k}\frac{\mathrm{d}z}{\mathrm{d}t} \\ \mathbf{a} &= \mathbf{u}_r \left[\frac{\mathrm{d}^2 r}{\mathrm{d}t^2} - r\left(\frac{\mathrm{d}\theta}{\mathrm{d}t}\right)^2\right] + \mathbf{u}_\theta \left[r\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} + 2\frac{\mathrm{d}r}{\mathrm{d}t}\frac{\mathrm{d}\theta}{\mathrm{d}t}\right] + \mathbf{k}\frac{\mathrm{d}^2 z}{\mathrm{d}t^2}\end{aligned}$$

- Thomas (1972) goes into an lengthy application of the above definitions to deriving Kepler's Laws.

Chapter 15

Partial Differentiation

15.1 Functions of Two or More Variables

12/16:

- **Function** (from D to E^1): A mapping that assigns a unique number w to each point $(x_1, \dots, x_n) \in D \subset E^n$.
  - We write $w = f(x_1, \dots, x_n)$ and say that w is the value of the function f at $(x_1, \dots, x_n)$.
- **Continuous** (function $f(x, y)$): A function $f(x, y)$ such that $w \to w_0 = f(x_0, y_0)$ as $(x, y) \to (x_0, y_0)$.

15.2 The Directional Derivative: Special Cases

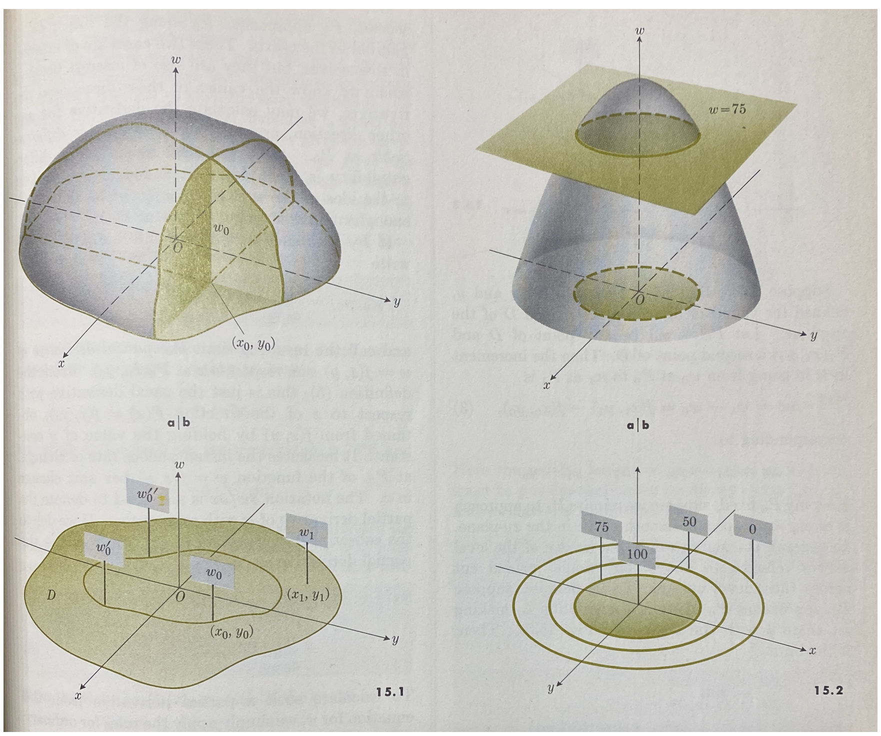


Figure 15.1: Surface plots and contour maps of 2D functions.

- The equation $w = f(x, y)$ can be interpreted as representing a surface in xyw-space, or as a base region D in the xy-plane with a marker bearing a corresponding w-value attached to each point.

  – To introduce order into the second interpretation, we can construct a **contour map** with a number of **contour curves**.

- **Contour curve**: A curve consisting of points $(x, y) \in D$ with equal w-values.
  – The formula for such a curve can be derived by setting $w_0 = f(x, y)$, where $w_0 \in R_f$.
- **Directional derivative** (of $f(x, y)$ at (x_0, y_0) in the ϕ-direction): The limit

$$\frac{\mathrm{d}w}{\mathrm{d}s} = \lim_{\Delta s \to 0} \frac{\Delta w}{\Delta s} = \lim_{P_1 \to P_0} \frac{f(x_1, y_1) - f(x_0, y_0)}{\sqrt{\Delta x^2 + \Delta y^2}}$$

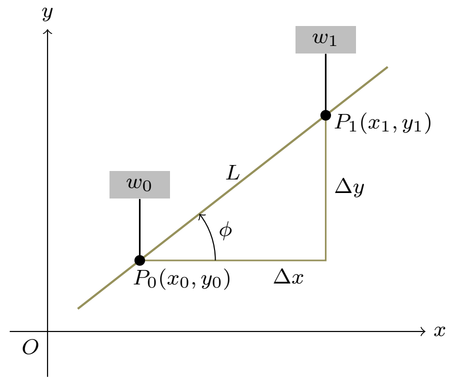


Figure 15.2: The directional derivative.

  – Basically, we let P_1 approach P_0 along a smooth curve (the line L connecting P_1 and P_0 for simplicity and to be definite; L makes an angle ϕ with the x-axis) and watch how $\Delta w = w_1 - w_0 = f(x_1, y_1) - f(x_0, y_0)$, $\Delta x = x_1 - x_0$, and $\Delta y = y_1 - y_0$ change.
  – Note that the directional derivative does depend on the *direction* from which P_1 approaches P_0, not just the absolute distance between P_1 and P_0.
- We now consider two special cases: When "P_1 approaches P_0 along the line $y = y_0$ parallel to the x-axis, [and when] P_1 approaches P_0 along the line $x = x_0$ parallel to the y-axis" (Thomas, 1972, p. 498).
  – These cases are important because if $f(x, y)$ is **differentiable** at P_0, we can calculate the directional derivative in any direction from them.
- **Partial derivative** (of $f(x, y)$ with respect to x at $P_0(x_0, y_0)$): The value

$$f_x(x_0, y_0) = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x, y_0) - f(x_0, y_0)}{\Delta x}$$

  – Essentially, this is the derivative with respect to x of the function $g(x) = f(x, y)$ with y held constant.
  – It measures "the instantaneous rate of change, at P_0, of the function $[f(x, y)]$ per unit change in x" (Thomas, 1972, p. 498).
- **Partial derivative** (of $w = f(x, y)$ with respect to x): The function

$$\frac{\partial w}{\partial x} = f_x(x, y) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}$$

  – To evaluate this, we apply the ordinary rules of differentiation, treating y as a constant.

- In either of the partial derivative definitions, Δx can be positive or negative. However, if we take the directional derivative in the positive x direction (for example), then Δx in the partial derivative definitions can only be positive.
  - Similarly, if f_x exists, it gives the directional derivative in the positive x-direction, whereas $-f_x$ is the directional derivative in the negative x-direction.

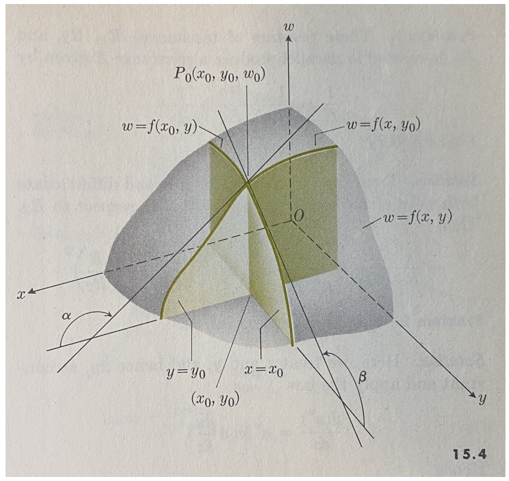


Figure 15.3: Geometric interpretation of the partial derivative.

- As in Figure 15.3, the geometric interpretation of the partial derivative (wrt. x) at a point $P(x_0, y_0, w_0)$ is as the slope of the curve $f(x, y_0)$, and symetrically wrt. y.
- We can define the partial derivative with respect to y similarly to how it is defined for x.

$$\frac{\partial w}{\partial y} = f_y(x, y) = \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}$$

- With higher order derivatives $\partial w / \partial z$, $\partial w / \partial u$, $\partial w / \partial v$, and more as in $w = f(x, y, z, u, v)$, we evaluate by holding all but the variable of interest constant.
- To denote the partial derivative at a point, we have two notations:

$$\left(\frac{\partial w}{\partial x}\right)_{(x_0, y_0)} \qquad\qquad f_x(x_0, y_0)$$

15.3 Tangent Plane and Normal Line

- **Tangent plane** (to $w = f(x, y)$ at $P_0(x_0, y_0, w_0)$): A plane T such that for any point P on the surface described by $f(x, y)$, as $P \to P_0$, the angle between T and $\overline{PP_0}$ approaches 0.
- **Normal line** (to $w = f(x, y)$ at $P_0(x_0, y_0, w_0)$): The line through P_0 which is normal to the tangent plane to $f(x, y)$ at P_0.
- The tangent plane is determined by the lines L_1 and L_2 tangent to the curves $C_1 : w = f(x_0, y)$ and $C_2 : w = f(x, y_0)$; the slopes of these lines are given by $\partial w / \partial y$ and $\partial w / \partial x$, respectively.
- Formulae for the tangent plane and normal line follow easily after finding a normal vector $\mathbf{N}$ to the plane of L_1 and L_2. To find $\mathbf{N}$, we can use the cross product of the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ lying along L_1 and L_2, respectively.

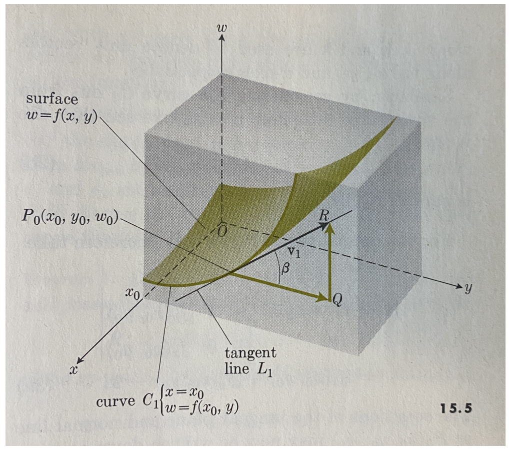


Figure 15.4: Deriving formulae for the tangent plane and normal line.

- From Figure 15.4, we can see that

$$\mathbf{v}_1 = \mathbf{j} + f_y(x_0, y_0)\mathbf{k} \qquad \mathbf{v}_2 = \mathbf{i} + f_x(x_0, y_0)\mathbf{k}$$

- Thus,

$$\mathbf{N} = \mathbf{i}f_x(x_0, y_0) + \mathbf{j}f_y(x_0, y_0) - \mathbf{k}$$

- Therefore, the formulae for the tangent plane and normal line, respectively, are

$$A(x - x_0) + B(y - y_0) + C(w - w_0) = 0 \qquad (x, y, w) = (x_0, y_0, w_0) + t(A, B, C)$$

where $A = f_x(x_0, y_0)$, $B = f_y(x_0, y_0)$, $C = -1$, and $t \in (-\infty, \infty)$.

- In vector form, if $\mathbf{R} = \mathbf{i}x + \mathbf{j}y + \mathbf{k}z$ and $\mathbf{R}_0 = \mathbf{i}x_0 + \mathbf{j}y_0 + \mathbf{k}z_0$, then

$$\mathbf{N} = \mathbf{i}f_x(x_0, y_0) + \mathbf{j}f_y(x_0, y_0) - \mathbf{k} \qquad \mathbf{N} \cdot (\mathbf{R} - \mathbf{R}_0) = 0 \qquad \mathbf{R} = \mathbf{R}_0 + t\mathbf{N}$$

15.4 Approximate Value of Δw

- **Linearization** (of f at P_0): The function (based off of the tangent plane)

$$w = f(x_0, y_0) + f_x(x_0, y_0) \cdot (x - x_0) + f_y(x_0, y_0) \cdot (y - y_0)$$

- Note that

$$\Delta w_{\text{tan}} = f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y$$

meaning that to calculate Δw_{tan}, we need only add the tangential components; no other interaction term is needed.

- Important results:

Theorem 15.1. Let the function $w = f(x, y)$ be continuous and possess partial derivatives f_x, f_y throughout a region $R : |x - x_0| < h,\ |y - y_0| < k$ of the xy-plane. Let f_x and f_y be continuous at (x_0, y_0). Let $\Delta w = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0)$. Then

$$\Delta w = f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y + \epsilon_1 \Delta x + \epsilon_2 \Delta y$$

where $\epsilon_1, \epsilon_2 \to 0$ when $\Delta x, \Delta y \to 0$.

Corollary 15.1. *Let $w = f(x, y)$ be continuous in a region $R : |x - x_0| < h,\ |y - y_0| < k$. Let f_x and f_y exist in R and be continuous at (x_0, y_0). Then the surface $w = f(x, y)$ has a tangent plane at $P_0(x_0, y_0, w_0)$, where $w_0 = f(x_0, y_0)$.*

- These results extend into finitely higher dimensions.

15.5 The Directional Derivative: General Case

12/17: • We first prove that the directional derivative can be expressed in terms of the partial derivatives.

Theorem 15.2. Let $w = f(x, y)$ be continuous and possess partial derivatives f_x, f_y throughout some neighborhood of the point $P_0(x_0, y_0)$. Let f_x and f_y be continuous at P_0. Then the directional derivative at P_0 exists for any direction angle ϕ and is given by

$$\frac{\mathrm{d}w}{\mathrm{d}s} = f_x(x_0, y_0)\cos\phi + f_y(x_0, y_0)\sin\phi$$

Proof. By Theorem 15.1, we know that $\Delta w = f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y + \epsilon_1\Delta x + \epsilon_2\Delta y$ where $\epsilon_1, \epsilon_2 \to 0$ as $\Delta x, \Delta y \to 0$. Thus,

$$\frac{\Delta w}{\Delta s} = f_x(x_0, y_0)\frac{\Delta x}{\Delta s} + f_y(x_0, y_0)\frac{\Delta y}{\Delta s} + \epsilon_1\frac{\Delta x}{\Delta s} + \epsilon_2\frac{\Delta y}{\Delta s}$$

where $\epsilon_1, \epsilon_2 \to 0$ as $\Delta x, \Delta y \to 0$. Hence, since $\lim \frac{\Delta w}{\Delta s} = \frac{\mathrm{d}w}{\mathrm{d}s}$, $\lim \frac{\Delta x}{\Delta s} = \frac{\mathrm{d}x}{\mathrm{d}s} = \cos\phi$, and $\lim \frac{\Delta y}{\Delta s} = \frac{\mathrm{d}y}{\mathrm{d}s} = \sin\phi$, we have

$$\frac{\mathrm{d}w}{\mathrm{d}s} = f_x(x_0, y_0)\cos\phi + f_y(x_0, y_0)\sin\phi$$

as desired. □

- Note that $\phi = 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$ lead to $\mathrm{d}w/\mathrm{d}s = \partial w/\partial x$, $\partial w/\partial y$, $-\partial w/\partial x$, $-\partial w/\partial y$, respectively.

• The directional derivative in three dimensions is given by

$$\frac{\mathrm{d}w}{\mathrm{d}s} = f_x(x_0, y_0, z_0)\cos\alpha + f_y(x_0, y_0, z_0)\cos\beta + f_z(x_0, y_0, z_0)\cos\gamma$$

- If we have a direction vector $\mathbf{u} = \mathbf{i}\cos\alpha + \mathbf{j}\cos\beta + \mathbf{k}\cos\gamma$ and another one that depends on the function and P_0 (the **gradient**) given by $\mathbf{v} = \mathbf{i}f_x(x_0, y_0, z_0) + \mathbf{j}f_y(x_0, y_0, z_0) + \mathbf{k}f_z(x_0, y_0, z_0)$, then

$$\frac{\mathrm{d}w}{\mathrm{d}s} = \mathbf{u} \cdot \mathbf{v}$$

15.6 The Gradient

• **Gradient** (of w): The vector function

$$\operatorname{grad} w = \nabla w = \mathbf{i}\frac{\partial w}{\partial x} + \mathbf{j}\frac{\partial w}{\partial y} + \mathbf{k}\frac{\partial w}{\partial z}$$

• The inverted capital delta is the **del operator**. In its own right, it is defined by

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}$$

• Recall that the directional derivative at $P_0(x_0, y_0, z_0)$ can be written as

$$\left(\frac{\mathrm{d}w}{\mathrm{d}s}\right)_0 = (\nabla w)_0 \cdot \mathbf{u}$$

- The significance of this is that it implies that

$$\left(\frac{\mathrm{d}w}{\mathrm{d}s}\right)_0 = |(\nabla w)_0|\,|u|\,\cos\theta = |(\nabla w)_0|\,\cos\theta$$

which means that the directional derivative is the "scalar projection of $\operatorname{grad} w$ at P_0, onto the direction $\mathbf{u}$" (Thomas, 1972, p. 511).

- Since $\mathrm{d}w/\mathrm{d}s$ is maximized when $\cos\theta = 1$, it must be that "the function $w = f(x, y, z)$ changes most rapidly in the direction given by the vector grad w itself. Moreover, the directional derivative in this direction is equal to the magnitude of the gradient" (Thomas, 1972, p. 511).

- The gradient vector at $P_0(x_0, y_0, z_0)$, where $w_0 = f(x_0, y_0, z_0)$ is also normal to the contour surface consisting of all points $P(x, y, z)$ for which $f(x, y, z) = w_0$.
  - We can prove this from the fact that the directional derivative in the direction of any line tangent to P_0 along the contour surface will be 0. But since $(\nabla w)_0 \neq \mathbf{0}$, we must have $\cos\theta = 0$, meaning that ∇w is perpendicular to any line tangent to P_0. It follows that it is normal to the contour surface, itself.

- Be careful with dimensions: The 3D vector grad w is points in the direction in 3-space of greatest change for a function with a 3D domain (a function best graphed in 4D), but is normal to a 2D surface that is a subset of this 3D domain.

15.7 The Chain Rule for Partial Derivatives

- Let $w = f(x, y, z)$ be a function with continuous partial derivatives f_x, f_y, f_z throughout some region R of xyz-space. If C is a curve lying in R defined by the parameterization $x = x(t)$, $y = y(t)$, and $z = z(t)$, then
$$\frac{\mathrm{d}w}{\mathrm{d}t} = \frac{\partial w}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial w}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} + \frac{\partial w}{\partial z}\frac{\mathrm{d}z}{\mathrm{d}t}$$
  - By generalizing Theorem 15.1 to three dimensions and dividing by Δt, we have
$$\frac{\Delta w}{\Delta t} = \left(\frac{\partial w}{\partial x}\right)_0 \frac{\Delta x}{\Delta t} + \left(\frac{\partial w}{\partial t}\right)_0 \frac{\Delta y}{\Delta t} + \left(\frac{\partial w}{\partial z}\right)_0 \frac{\Delta z}{\Delta t} + \epsilon_1 \frac{\Delta x}{\Delta t} + \epsilon_2 \frac{\Delta y}{\Delta t} + \epsilon_3 \frac{\Delta z}{\Delta t}$$
    where $\epsilon_1, \epsilon_2, \epsilon_3 \to 0$ as $\Delta x, \Delta y, \Delta z \to 0$.
  - Now if C is differentiable in R, too, (i.e., $\mathrm{d}x/\mathrm{d}t$, $\mathrm{d}y/\mathrm{d}t$, $\mathrm{d}z/\mathrm{d}t$ all exist) then $\Delta x, \Delta y, \Delta z \to 0$ as $\Delta t \to 0$.
  - Therefore, if we take the limit as $\Delta t \to 0$ of the above equation, we get the desired result.

- Note that we can also write
$$\frac{\mathrm{d}w}{\mathrm{d}t} = \nabla w \cdot \mathbf{v}$$
  where $\mathbf{v}$ is the velocity vector along the curve C.

- We can also consider the behavior of a function along a surface S lying in R parameterized by $x = x(r, s)$, $y = y(r, s)$, and $z = z(r, s)$. In this case we can take
$$\frac{\partial w}{\partial r} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial r} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial r} \qquad \frac{\partial w}{\partial s} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial s}$$
  - We can also expand this to many more dimensions. This idea is best summarized in matrix form (for a function $w = f(x_1, \ldots, x_n)$ parameterized by $x_1 = x_1(y_1, \ldots, y_m)$, $\ldots$, $x_n = (y_1, \ldots, y_m)$):
$$\begin{bmatrix} \frac{\partial w}{\partial y_1} & \frac{\partial w}{\partial y_2} & \cdots & \frac{\partial w}{\partial y_m} \end{bmatrix} = \begin{bmatrix} \frac{\partial w}{\partial x_1} & \frac{\partial w}{\partial x_2} & \cdots & \frac{\partial w}{\partial x_n} \end{bmatrix} \begin{bmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \cdots & \frac{\partial x_1}{\partial y_m} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} & \cdots & \frac{\partial x_2}{\partial y_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \frac{\partial x_n}{\partial y_2} & \cdots & \frac{\partial x_n}{\partial y_m} \end{bmatrix}$$

- To clarify the above results, we investigate a few problems.

- "Suppose that $w = r^2 \cos 2\theta$, where $x = r\cos\theta$, $y = r\sin\theta$, $r = \sqrt{x^2+y^2}$, [and] $\theta = \tan^{-1}(y/x)$. Find $\partial w/\partial x$ and $\partial w/\partial y$" (Thomas, 1972, p. 516).
  - We use the matrix methods to get an equation for the desired results.

$$\begin{aligned}\begin{bmatrix}\frac{\partial w}{\partial x} & \frac{\partial w}{\partial y}\end{bmatrix} &= \begin{bmatrix}\frac{\partial w}{\partial r} & \frac{\partial w}{\partial \theta}\end{bmatrix}\begin{bmatrix}\frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y}\end{bmatrix} \\ &= \begin{bmatrix}2r\cos 2\theta & -2r^2\sin 2\theta\end{bmatrix}\begin{bmatrix}\frac{x}{r} & \frac{y}{r} \\ -\frac{y}{r^2} & \frac{x}{r^2}\end{bmatrix} \\ &= \begin{bmatrix}2x\cos 2\theta + 2y\sin 2\theta & 2y\cos 2\theta - 2x\sin 2\theta\end{bmatrix}\end{aligned}$$

  - Now, we can use some substitutions to put the above results in terms of the variable with respect to which we are differentiating.

$$\begin{aligned}\frac{\partial w}{\partial x} &= 2(x\cos 2\theta + y\sin 2\theta) \\ &= 2r(\cos\theta\cos 2\theta + \sin\theta\sin 2\theta) \\ &= 2r\cos(2\theta-\theta) \\ &= 2x\end{aligned} \qquad \begin{aligned}\frac{\partial w}{\partial y} &= 2(y\cos 2\theta - x\sin 2\theta) \\ &= 2r(\sin\theta\cos 2\theta - \cos\theta\sin 2\theta) \\ &= 2r\sin(\theta - 2\theta) \\ &= -2y\end{aligned}$$

- "Show that the change of variables from x and y to $r = y - ax$ [and] $s = y + ax$ transforms the differential equation

$$\frac{\partial w}{\partial x} - a\frac{\partial w}{\partial y} = 0$$

  into a form that is more easily solved, and solve it. (Here a is a constant.)" (Thomas, 1972, p. 516).
  - We divide into two cases ($a \neq 0$ and $a = 0$), beginning with the former.
  - Imagine that $w = f(x,y)$ is transformed into $w = \tilde{f}(r,s)$ via substitutions which can be derived by treating the definitions of r and s as a two-variable system of equations and solving for x and y.

$$x = \frac{1}{2a}(s-r) \qquad y = \frac{1}{2}(r+s)$$

  - Now to find $\partial w/\partial x$ and $\partial w/\partial y$, we use the chain rule.

$$\begin{aligned}\begin{bmatrix}\frac{\partial w}{\partial x} & \frac{\partial w}{\partial y}\end{bmatrix} &= \begin{bmatrix}\frac{\partial w}{\partial r} & \frac{\partial w}{\partial s}\end{bmatrix}\begin{bmatrix}\frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial s}{\partial x} & \frac{\partial s}{\partial y}\end{bmatrix} \\ &= \begin{bmatrix}\frac{\partial w}{\partial r} & \frac{\partial w}{\partial s}\end{bmatrix}\begin{bmatrix}-a & 1 \\ a & 1\end{bmatrix} \\ &= \begin{bmatrix}-a\frac{\partial w}{\partial r} + a\frac{\partial w}{\partial s} & \frac{\partial w}{\partial r} + \frac{\partial w}{\partial s}\end{bmatrix}\end{aligned}$$

  - By substituting into the original differential equation, we get $-2a\,\partial w/\partial r = 0$, or $\frac{\partial w}{\partial r} = 0$.
  - Thus, this differential equation is easy to solve — we need only require that w is constant in the r-direction. Indeed, w can be any function of s. Therefore, the solution is

$$w = \phi(s) = \phi(y + ax)$$

    where $\phi(s)$ is any differentiable function of s, whatsoever.

  - If $a = 0$, then we have the similar case $\partial w/\partial x = 0$.
  - Note that in solving an ordinary differential equation, we often get constants of integration. In solving a partial differentiable equation, arbitrary functions (such as $\phi(s)$) are analogous to these constants of integration. Extending the analogy, they can sometimes be solved for with "initial conditions," as in the next problem.

- Find an explicit formula for w in the above problem if its values along the x-axis are given by $w = \sin x$ and if $a \neq 0$.
  - The general solution is $w = f(x, y) = \phi(y + ax)$.
  - We are given $f(x, 0) = \phi(0 + ax) = \sin x$.
  - Thus, if we let $u = ax$, we have $\phi(u) = \sin \frac{u}{a}$, meaning that $w = \phi(y + ax) = \sin \frac{y+ax}{a}$.

15.8 The Total Differential

- **Partial differential** (of $w = f(x, y, z)$ with respect to x): The infinitesimal value

$$\frac{\partial w}{\partial x}\,\mathrm{d}x$$

  - There also exist symmetric partial differentials with respect to y and z.

- **Total differential** (of $w = f(x, y, z)$): The infinitesimal value

$$\mathrm{d}w = \frac{\partial w}{\partial x}\,\mathrm{d}x + \frac{\partial w}{\partial y}\,\mathrm{d}y + \frac{\partial w}{\partial z}\,\mathrm{d}z$$

that is the sum of the partial differentials of w.

- If x, y, z are given as functions of a single variable t, we have

$$\mathrm{d}x = x'(t)\,\mathrm{d}t \qquad \mathrm{d}y = y'(t)\,\mathrm{d}t \qquad \mathrm{d}z = z'(t)\,\mathrm{d}t$$

- If x, y, z are given as functions of two variables r, s, their total differentials are

$$\mathrm{d}x = \frac{\partial x}{\partial r}\,\mathrm{d}r + \frac{\partial x}{\partial s}\,\mathrm{d}s \qquad \mathrm{d}y = \frac{\partial y}{\partial r}\,\mathrm{d}r + \frac{\partial y}{\partial s}\,\mathrm{d}s \qquad \mathrm{d}z = \frac{\partial z}{\partial r}\,\mathrm{d}r + \frac{\partial z}{\partial s}\,\mathrm{d}s$$

  - If we take the perspective that $w = f[x(r, s), y(r, s), z(r, s)] = \tilde{f}(r, s)$ is a function of two variables, then we should have

$$\mathrm{d}w = \frac{\partial w}{\partial r}\,\mathrm{d}r + \frac{\partial w}{\partial s}\,\mathrm{d}s$$

  - Now this $\mathrm{d}w$ is the same as the one given earlier with respect to x, y, z, as a consequence of the chain rule. Indeed,

$$\begin{aligned}
\mathrm{d}w &= \frac{\partial w}{\partial x}\left(\frac{\partial x}{\partial r}\,\mathrm{d}r + \frac{\partial x}{\partial s}\,\mathrm{d}s\right) + \frac{\partial w}{\partial y}\left(\frac{\partial y}{\partial r}\,\mathrm{d}r + \frac{\partial y}{\partial s}\,\mathrm{d}s\right) + \frac{\partial w}{\partial z}\left(\frac{\partial z}{\partial r}\,\mathrm{d}r + \frac{\partial z}{\partial s}\,\mathrm{d}s\right) \\
&= \left(\frac{\partial w}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial r} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial r}\right)\mathrm{d}r + \left(\frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial s}\right)\mathrm{d}s \\
&= \frac{\partial w}{\partial r}\,\mathrm{d}r + \frac{\partial w}{\partial s}\,\mathrm{d}s
\end{aligned}$$

  - Note that r and s are the independent variables here, so while we can approximate $\mathrm{d}r$ and $\mathrm{d}s$ with Δr and Δs, respectively, we should not approximate $\mathrm{d}x$, $\mathrm{d}y$, $\mathrm{d}z$ with $\Delta x, \Delta y, \Delta z$. Indeed, we should use the differentials $\mathrm{d}x$, $\mathrm{d}y$, $\mathrm{d}z$ as defined in terms of $\mathrm{d}r$, $\mathrm{d}s$ to approximate $\Delta x, \Delta y, \Delta z$.

- These results can be generalized to higher dimensions.
- The following example should make clear the power of these differential definitions: "Consider the function $w = x^2 + y^2 + z^2$ with $x = r\cos s$, $y = r\sin s$, [and] $z = r$" (Thomas, 1972, p. 519).
  - The total differential is $\mathrm{d}w = 2(x\,\mathrm{d}x + y\,\mathrm{d}y + z\,\mathrm{d}z)$.
  - The differentials in terms of r and s are $\mathrm{d}x = \cos s\,\mathrm{d}r - r\sin s\,\mathrm{d}s$, $\mathrm{d}y = \sin s\,\mathrm{d}r + r\cos s\,\mathrm{d}s$, and $\mathrm{d}z = \mathrm{d}r$.
  - Hence, the total differential can also be written as

$$\begin{aligned}\mathrm{d}w &= 2(x\cos s + y\sin s + z)\,\mathrm{d}r + 2(-xr\sin s + yr\cos s)\,\mathrm{d}s\\ &= 2(r\cos^2 s + r\sin^2 s + r)\,\mathrm{d}r + 2(-r^2\cos s\sin s + r^2\sin s\cos s)\,\mathrm{d}s\\ &= 4r\,\mathrm{d}r\end{aligned}$$

  - Integrating, the above yields $w = 2r^2$.
  - Critically, we can also derive this formula for w from the original function and parameterization since $w = r^2\cos^2 s + r^2\sin^2 s + r^2 = 2r^2$.
- If $F(x,y) = 0$, then for the plane curve generated, $\mathrm{d}y/\mathrm{d}x = -F_x(x,y)/F_y(x,y)$ if $F_y(x,y) \neq 0$.
  - We can also derive this result by implicitly differentiating $F(x,y) = 0$, as we are allowed to by the **implicit function theorem**.

15.9 Maxima and Minima of Functions of Two Independent Variables

12/18:

- **Relative minimum** (of $f(x,y)$): A point (a,b), in a region R where f is defined, continuous, and has continuous partial derivatives with respect to x and y, such that $f(x,y) \geq f(a,b)$ for all points (x,y) sufficiently close to (a,b). *Also known as* **local minimum**.
- **Relative maximum** (of $f(x,y)$): A point (a,b), in a region R where f is defined, continuous, and has continuous partial derivatives with respect to x and y, such that $f(x,y) \leq f(a,b)$ for all points (x,y) sufficiently close to (a,b). *Also known as* **local maximum**.
- **Absolute minimum** (of $f(x,y)$): A relative minimum (a,b) of $f(x,y)$ such that $f(x,y) \geq f(a,b)$ for all $(x,y) \in R$.
- **Absolute maximum** (of $f(x,y)$): A relative maximum (a,b) of $f(x,y)$ such that $f(x,y) \leq f(a,b)$ for all $(x,y) \in R$.
- If (a,b) is a relative maximum[1], then $f_x(a,b) = 0$ and $f_y(a,b) = 0$.
  - We can see from Figure 15.5 that the curve lying in the plane $y = b$ given by $z = f(x,b)$ has a high turning point at $x = a$.
  - We can similarly see that $z = f(a,y)$ has a high turning point at $y = b$.
  - Thus,

$$\left(\frac{\partial z}{\partial x}\right)_{(a,b)} = 0 \qquad\qquad \left(\frac{\partial z}{\partial y}\right)_{(a,b)} = 0$$

  - This implies the desired result.

[1]Or minimum. We choose arbitrarily to work with maxima from here on out, but every statement is symmetric for minima.

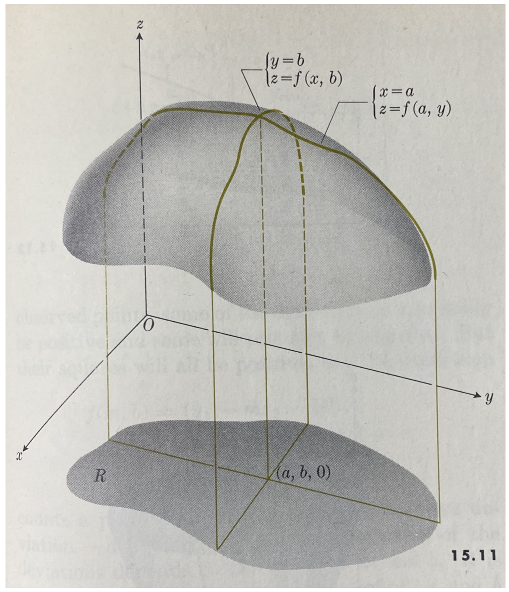


Figure 15.5: Relative maximum of $f(x, y)$.

- As to the second derivative test, Thomas (1972) does not go into it deeply, but mentions that the key is that $D = f(x, y) - f(a, b)$ is nonnegative (positive or zero) for a minimum and nonpositive (negative or zero) for a maximum for all (x, y) sufficiently close to (a, b).
  - He recommends checking $x = a + h$ and $y = b + k$ for small values of h and k to confirm.
  - This may seem to not be rigorous, but it can actually work quite well, as we will see in the following problem.
- Find the minima and maxima on the surface

$$z = f(x, y) = x^2 - xy + y^2 + 2x + 2y - 4$$

  - Since the domain of this functions is not restricted, there are no boundary points to check. Thus, we need only apply the necessary conditions

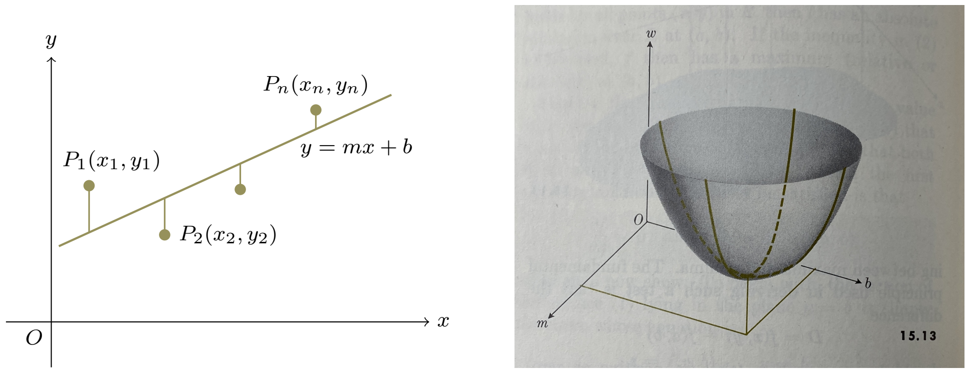
$$\begin{aligned} 0 &= \frac{\partial z}{\partial x} \\ &= 2x - y + 2 \end{aligned} \qquad \begin{aligned} 0 &= \frac{\partial z}{\partial y} \\ &= 2y - x + 2 \end{aligned}$$

  - If we solve the above as a two-variable system of equations, we find that $(-2, -2)$ is the solution.
  - Thus, the critical point is $(-2, -2, f(-2, -2)) = (-2, -2, -8)$.
  - Applying the pseudo-second-derivative test, we have

$$\begin{aligned} D &= f(-2 + h, -2 + k) - f(-2, -2) \\ &= h^2 - hk + k^2 \\ &= \left(h - \frac{k}{2}\right)^2 + \frac{3k^2}{4} \end{aligned}$$

    which is clearly positive unless $h = k = 0$.
  - Therefore, $(-2, -2, -8)$ is an absolute minimum, and there are no other high or low points.

15.10 The Method of Least Squares



(a) Best fit line. (b) Minimization.

Figure 15.6: Method of least squares.

- **Method of least squares**: A technique for fitting a straight line $y = mx+b$ to a set of experimentally observed points $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$.
- **Deviation**: The difference between the observed y-value and the y-value predicted by the straight line. *Also known as* **dev**, $\boldsymbol{d_n}$.

$$\text{dev} = y_{\text{obs}} - (mx_{\text{obs}} + b)$$

- "For a straight line which comes close to fitting all of the observed points, some of the deviations will probably be positive and some will probably be negative. But their squares will all be positive, and the expression

$$f(m, b) = (y_1 - mx_1 - b)^2 + (y_2 - mx_2 - b)^2 + \cdots + (y_n - mx_n - b)^2$$

counts a positive deviation $+d$ and a negative deviation $-d$ equally. This sum of squares of the deviations depends on the choice of m and b. It is never negative, and it can be zero only if m and b have values that produce a straight line that is a perfect fit" (Thomas, 1972, pp. 524–25).
  - Essentially, the method of least squares says "Take as the line $y = mx + b$ of best fit that one for which the sum of squares of the deviations

$$f(m, b) = d_1^2 + d_2^2 + \cdots + d_n^2$$

is a minimum" (Thomas, 1972, p. 525).
- Note that we can use the pseudo-second-derivative test to show that the point on $f(m, b)$ found by the method of least squares is a minimum. However, it is customary to omit this step since it can be shown that for the general case of fitting a straight line, the answer is always a minimum.

15.11 Maxima and Minima of Functions of Several Independent Variables

- This is necessary in certain statistical applications.
- As we would expect, minima and maxima of functions of the form $w = f(x_1, \dots, x_n)$ can lie at boundary points, or at points where $0 = \partial f/\partial x_1, \dots, \partial f/\partial x_n$.

- Sometimes a function $w = f(x_1, \dots, x_n)$ is given with certain constraints of the form $g(x_1, \dots, x_n) = 0$. In these cases, use the constraints to express some of the variables in terms of the remaining ones (so that the remaining ones are independent) before taking partial derivatives.
- The minimization process may lead to an answer that lies outside the region where f is defined. In these cases, it can help to choose a different set of independent variables. Let's look at one example of this phenomenon.
- "Find the minimum distance from the origin to the surface $x^2 - z^2 = 1$" (Thomas, 1972, p. 528).

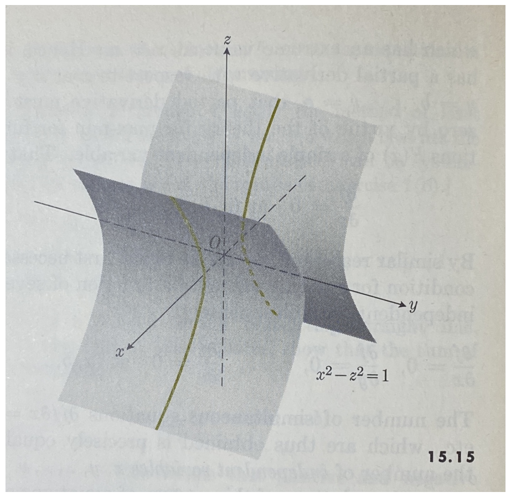


Figure 15.7: Minimizing the distance from the origin across a hyperbolic cylinder.

  - We want to minimize $\sqrt{(x-0)^2 + (y-0)^2 + (z-0)^2}$, the formula for the distance between a point (x, y, z) and the origin, over all points (x, y, z) on the surface defined by $x^2 - z^2 = 1$.
  - Thus, we choose to minimize $w = x^2 + y^2 + z^2$ (it's easier to work without the radical, and the latter formula minimizes at the same points as the former). To limit the set of points in the domain of w, we arbitrarily choose x and y to be our independent variables and use the constraint to substitute out z, yielding $w = 2x^2 + y^2 - 1$.
  - However, $0 = \partial w/\partial x = 4x$ and $0 = \partial w/\partial y = 2y$ lead to $(0, 0, i)$ as our answer. The problem here is that $x \in (-1, 1)$ is not in the domain of $x^2 - z^2 = 1$, yet it is in the domain of $w = 2x^2 + y^2 - 1$.
  - Thus, we need a different substitution. If we eliminate x instead, then we have $w = 1 + y^2 + 2z^2$ in terms of variables that have meaning for all values in the set $(-\infty, \infty)$ (importantly, no new elements are added to the domain). Indeed, minimizing this, we get $(\pm 1, 0, 0)$ as our answers, and we can see from Figure 15.7 that these are correct.

- **Method of Lagrange multipliers**: "To minimize (or maximize) a function $f(x, y, z)$, subject to the constraint $g(x, y, z) = 0$, construct the auxiliary function

$$H(x, y, z, \lambda) = f(x, y, z) - \lambda g(x, y, z)$$

  and find values of x, y, z, λ for which the partial derivatives of H are all 0: $H_x = 0$, $H_y = 0$, $H_z = 0$, [and] $H_\lambda = 0$" (Thomas, 1972, p. 528).

- "Find the point on the plane

$$2x - 3y + 5z = 19$$

  that is nearest the origin, using the method of Lagrange multipliers" (Thomas, 1972, p. 528).

  - Find H.

$$H(x, y, z, \lambda) = x^2 + y^2 + z^2 - \lambda(2x - 3y + 5z - 19)$$

- Take partial derivatives.

$$0 = H_x = 2x - 2\lambda \qquad 0 = H_y = 2y + 3\lambda \qquad 0 = H_z = 2z - 5\lambda \qquad 0 = H_\lambda = -g(x, y, z)$$

- Solve H_x, H_y, H_z for x, y, z.

$$x = \lambda \qquad\qquad y = -\frac{3}{2}\lambda \qquad\qquad z = \frac{5}{2}\lambda$$

- Plug into H_λ and solve for λ.

$$\begin{aligned} 2\lambda + \frac{9}{2}\lambda + \frac{25}{2}\lambda &= 19 \\ \lambda &= 1 \end{aligned}$$

- Return the substitutions.

$$(x, y, z) = \left(1, -\frac{3}{2}, \frac{5}{2}\right)$$

- Thomas (1972) derives the method of Lagrange multiplies, finding in the process the important equation $\nabla f = \lambda \nabla g$ where $\lambda = f_z/g_z$[2].
- If there exist two constraints $g(x, y, z)$ and $h(x, y, z)$, then we work with

$$H(x, y, z, \lambda, \mu) = f(x, y, z) - \lambda g(x, y, z) - \mu h(x, y, z)$$

and set $H_x, H_y, H_z, H_\lambda, H_\mu = 0$ in the process.

  - We also have in this case $\nabla f = \lambda \nabla g + \mu \nabla h$.

15.12 Higher-Order Derivatives

12/19: • To denote higher-order partial derivatives, use symbols such as

$$\frac{\partial^2 f}{\partial x^2},\ \frac{\partial^2 f}{\partial y^2},\ \frac{\partial^2 f}{\partial x \partial y},\ \frac{\partial^2 f}{\partial y \partial x}$$

or, respectively,

$$f_{xx},\ f_{yy},\ f_{yx},\ f_{xy}$$

- The order of differentiation is "conserved" left to right, in a sense. For example,

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$$

- However, in general, $\partial^2 f/\partial x \partial y = \partial^2 f/\partial y \partial x$:

Theorem 15.3. Let the function $w = f(x, y)$, together with the partial derivatives f_x, f_y, f_{xy}, and f_{yx} be continuous in some neighborhood of a point $P(a, b)$. Then, at that point,

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$$

[2]There is some important but rather indecipherable geometric meaning of this equation. Additionally, the proof overall would be good to understand. This is something to come back to at a later date, though.

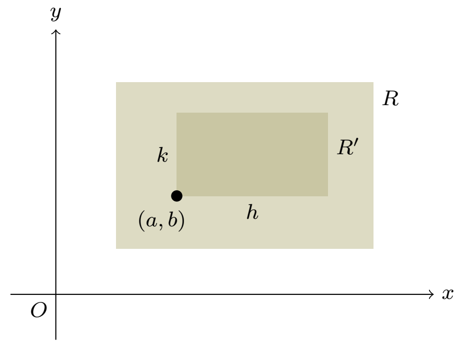


Figure 15.8: Equivalence of higher-order partial derivative ordering.

Proof. Let h and k be numbers such that $(a+h, b+k)$ lies in the rectangle R, and let $F(x)$ be defined by $F(x) = f(x, b+k) - f(x, b)$. Consider the difference

$$\Delta = F(a+h) - F(a)$$

If we apply the Mean Value Theorem to F, we get

$$\Delta = hF'(c_1)$$

for some $c_1 \in [a, a+h]$. Substituting, we have

$$\Delta = h[f_x(c_1, b+k) - f_x(c_1, b)]$$

If we apply the Mean Value Theorem to the difference $f_x(c_1, b+k) - f_x(c_1, b)$, we get $f_x(c_1, b+k) - f_x(c_1, b) = kf_{xy}(c_1, d_1)$ where $d_1 \in [b, b+k]$. Therefore,

$$\Delta = hkf_{xy}(c_1, d_1)$$

where $(c_1, d_1) \in R'$ in Figure 15.8. Now, if we let $G(y) = f(a+h, y) - f(a, y)$, we can arrive at the similar result

$$\Delta = hkf_{yx}(c_2, d_2)$$

where $(c_2, d_2) \in R'$ in Figure 15.8. Thus, by transitivity, we have

$$f_{xy}(c_1, d_1) = f_{yx}(c_2, d_2)$$

That is not what we wanted to prove. However, since h and k can be made arbitrarily small, and since the continuity of f_{xy} and f_{yx} implies that

$$f_{xy}(c_1, d_1) = f_{xy}(a, b) + \epsilon_1 \qquad\qquad f_{yx}(c_2, d_2) = f_{yx}(a, b) + \epsilon_2$$

where $\epsilon_1, \epsilon_2 \to 0$ as $h, k \to 0$, we let $h, k \to 0$ and get

$$f_{xy}(a, b) = f_{yx}(a, b)$$

as desired. □

- Note that the quantity Δ in the above proof of Theorem 15.3 is known as the **second difference**.
  - Although it is an advanced skill to know introduce such a tool in a proof, it makes sense to use the *second* difference in a proof about *second* derivatives.
  - Note that for sufficiently small h, k, $\Delta \approx hkf_{xy}(a, b)$.

- From Theorem 15.3, we can prove higher-order partial derivative equalities such as

$$\frac{\partial^3 f}{\partial x^2 \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial x \partial y}\right) = \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial y \partial x}\right) = \frac{\partial}{\partial y}\left(\frac{\partial^2 f}{\partial x \partial x}\right) = \frac{\partial^3 f}{\partial y \partial x^2}$$

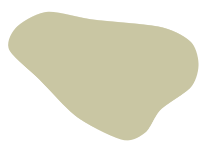
- "In fact, if all the partial derivatives that appear are continuous, the notation

$$\frac{\partial^{m+n} f}{\partial x^m \partial y^n}$$

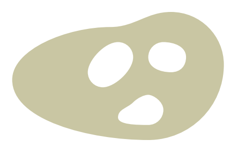
may be used to denote the result of differentiating the function $f(x, y)$ m times with respect to x and n times with respect to y, the order in which these differentiations are performed being entirely arbitrary" (Thomas, 1972, p. 535).

15.13 Exact Differentials

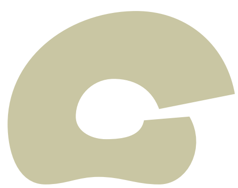
- **Exact differential**: The expression $\mathrm{d}f(x, y) = M(x, y)\,\mathrm{d}x + N(x, y)\,\mathrm{d}y$ corresponding to a function $w = f(x, y)$.



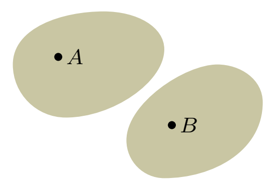
(a) A simply connected region.



(b) A (not simply) connected region.



(c) A simply connected region.



(d) A disconnected region (no path AB lies entirely within the set, for example).

Figure 15.9: Types of regions in the plane.

- **Simply connected region** (of the plane): A set of points that is **open**, **connected**, and satisfies the property that "if C is any simple closed curve, all of whose points are in the set, then all points in the interior of C are also in the set" (Thomas, 1972, p. 537).
- **Open** (set in the plane): A set such that "each point of the set is an **interior** point of the set" (Thomas, 1972, p. 537).
- **Connected** (set): A set such that "any two points of the set can be joined by a polygonal path, all of whose points are in the set" (Thomas, 1972, p. 537).
- **Interior** (point of a set in the plane): A point that "can be the center of a small circle whose entire interior is in the set" (Thomas, 1972, p. 537).
- We now prove a theorem that answers two questions with respect to exact differentials: (1) "How can we tell whether a given expression is or is not an exact differential?" and (2) "If the expression is exact, how do we find the function $f(x, y)$ of which it is the differential?" (Thomas, 1972, p. 536).

  Theorem 15.4. Let the functions $M(x, y)$ and $N(x, y)$ be continuous, and let them possess continuous partial derivatives M_x, M_y, N_x, N_y for all real values of x and y in some simply connected region G.

Then a necessary and sufficient condition for $M(x,y)\,\mathrm{d}x + N(x,y)\,\mathrm{d}y$ to be an exact differential in G is that

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$$

Proof. We first demonstrate that $\partial M/\partial y = \partial N/\partial x$ is a necessary condition. Suppose that there exists a function $f(x,y)$ with an exact differential at all points in G. We also know that $\mathrm{d}f = \partial f/\partial x\,\mathrm{d}x + \partial f/\partial y\,\mathrm{d}y$ by the definition of the total differential. Thus,

$$M(x,y)\,\mathrm{d}x + N(x,y)\,\mathrm{d}y = \frac{\partial f}{\partial x}\,\mathrm{d}x + \frac{\partial f}{\partial y}\,\mathrm{d}y$$

But since $\mathrm{d}x$ and $\mathrm{d}y$ are independent variables in the above equation, we can set either to 0 and the equality must be maintained. Consequently, we must have

$$\frac{\partial f}{\partial x} = M(x,y) \qquad\qquad \frac{\partial f}{\partial y} = N(x,y)$$

Now by Theorem 15.3, we have

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$$

and it follows from this that

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$$

as desired.

To demonstrate that $\partial M/\partial y = \partial N/\partial x$ is a sufficient condition, we will show how to find the function $f(x,y)$ from the exact differential. To begin, let $\partial f/\partial x = M(x,y)$. From this definition, the condition that $\partial M/\partial y = \partial N/\partial x$, and Theorem 15.3, we can immediately learn another interesting fact about f:

$$\begin{aligned}\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) &= \frac{\partial M}{\partial y}\\ \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) &= \frac{\partial N}{\partial x}\\ \frac{\partial f}{\partial y} &= N(x,y)\end{aligned}$$

With this established, we can now solve for $f(x,y)$. Start with

$$\frac{\partial f}{\partial x} = M(x,y)$$

From here, we integrate both sides with respect to x, holding y constant. Notice that we introduce $g(y)$, a function of just y, as a kind of constant of integration[3].

$$f(x,y) = \int_x M(x,y)\,\mathrm{d}x + g(y)$$

We now take the partial derivative with respect to y.

$$\frac{\partial f}{\partial y} = \frac{\partial}{\partial y}\int_x M(x,y)\,\mathrm{d}x + \frac{\partial g}{\partial y}$$

Since g is purely a function of y, the partial derivative of g with respect to y is simply the derivative of g with respect to y.

$$= \frac{\partial}{\partial y}\int_x M(x,y)\,\mathrm{d}x + \frac{\mathrm{d}g}{\mathrm{d}y}$$

We now make use of the fact that we proved in the beginning, and substitute out $\frac{\partial f}{\partial y}$ for $N(x,y)$. We then rearrange the terms to create a differential equation for g.

$$\begin{aligned}N(x,y) &= \frac{\partial}{\partial y}\int_x M(x,y)\,\mathrm{d}x + \frac{\mathrm{d}g}{\mathrm{d}y}\\ \frac{\mathrm{d}g}{\mathrm{d}y} &= N(x,y) - \frac{\partial}{\partial y}\int_x M(x,y)\,\mathrm{d}x\end{aligned}$$

This differential equation can be solved for g by integrating the right-hand side with respect to y. After that has been accomplished, we can substitute the definition of g back into $f(x,y) = \int_x M(x,y)\,\mathrm{d}x + g(y)$ to generate the final formula for $f(x,y)$:

$$f(x,y) = \int_x M(x,y)\,\mathrm{d}x + \int_y \left(N(x,y) - \frac{\partial}{\partial y}\int_x M(x,y)\,\mathrm{d}x\right)\mathrm{d}y$$

□

- We now use an example to tangibly demonstrate both how to determine that a differential is a exact, and how to solve for $f(x,y)$.
- Consider the differential $(x^2+y^2)\,\mathrm{d}x + 2xy\,\mathrm{d}y$. Use the condition of Theorem 15.4 to show that it is exact, and then solve for the function $w = f(x,y)$ for which $\mathrm{d}w = (x^2+y^2)\,\mathrm{d}x + 2xy\,\mathrm{d}y$.
  - Let $M(x,y) = x^2+y^2$ and let $N(x,y) = 2xy$. Then

    $$\frac{\partial M}{\partial y} = 2y = \frac{\partial N}{\partial x}$$

    so the given differential is exact.
  - Now let $\partial f/\partial x = x^2+y^2$, and recall that we can prove from this that $\partial f/\partial y = 2xy$. It follows that

    $$\begin{aligned}f(x,y) &= \int_x (x^2+y^2)\,\mathrm{d}x + g(y)\\ &= \frac{x^3}{3} + xy^2 + g(y)\end{aligned}$$

    The above equation is important, and we'll end up substituting $g(y)$ into it once we find it as follows.

    $$\begin{aligned}\frac{\partial f}{\partial y} &= \frac{\partial}{\partial y}\left(\frac{x^3}{3} + xy^2\right) + \frac{\partial g}{\partial y}\\ 2xy &= 0 + 2xy + \frac{\mathrm{d}g}{\mathrm{d}y}\\ \frac{\mathrm{d}g}{\mathrm{d}y} &= 0\\ g(y) &= C \qquad C \in \mathbb{R}\end{aligned}$$

  - Therefore,

    $$f(x,y) = \frac{x^3}{3} + xy^2 + C$$

[3] In the sense that if we took the partial derivative of both sides of the following with respect to x, we would get the above.

15.14 Derivatives of Integrals

- From the Fundamental Theorem of Integral Calculus, we know that if f is a continuous function of $a \leq t \leq b$, then

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_a^x f(t)\,\mathrm{d}t = f(x) \qquad\qquad \frac{\mathrm{d}}{\mathrm{d}x}\int_x^b f(t)\,\mathrm{d}t = -f(x)$$

- These results allow us to prove the following.

Theorem 15.5. Let f be continuous on $a \leq t \leq b$. Let u and v be differentiable functions of x such that $u(x)$ and $v(x)$ lie between a and b. Then

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_{u(x)}^{v(x)} f(t)\,\mathrm{d}t = f[v(x)]\frac{\mathrm{d}v}{\mathrm{d}x} - f[u(x)]\frac{\mathrm{d}u}{\mathrm{d}x}$$

Proof. Let $F[u(x), v(x)] = \int_{u(x)}^{v(x)} f(t)\,\mathrm{d}t$. Then by the above results from the Fundamental Theorem of Calculus, we respectively have

$$\frac{\partial F}{\partial v} = f[v(x)] \qquad\qquad \frac{\partial F}{\partial u} = -f[u(x)]$$

Since $\mathrm{d}F/\mathrm{d}x = \partial F/\partial u\ \mathrm{d}u/\mathrm{d}x + \partial F/\partial v\ \mathrm{d}v/\mathrm{d}x$ (the chain rule for partial derivatives), we therefore have

$$\frac{\mathrm{d}F}{\mathrm{d}x} = -f[u(x)]\frac{\mathrm{d}u}{\mathrm{d}x} + f[v(x)]\frac{\mathrm{d}v}{\mathrm{d}x}$$
$$\frac{\mathrm{d}}{\mathrm{d}x}\int_{u(x)}^{v(x)} f(t)\,\mathrm{d}t = f[v(x)]\frac{\mathrm{d}v}{\mathrm{d}x} - f[u(x)]\frac{\mathrm{d}u}{\mathrm{d}x}$$

as desired. $\square$

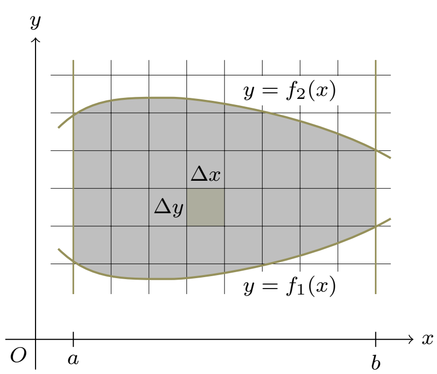
Chapter 16

Multiple Integrals

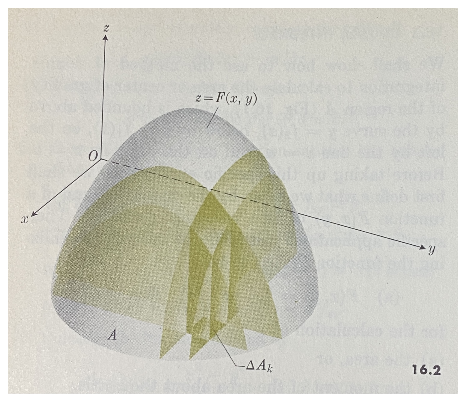
16.1 Double Integrals

12/19: - **Double integral** (of $F(x, y)$ over the region A): The limit as $\Delta A \to 0$ of the sum of every ΔA_k (composing A) times $F(x, y)$ for some $(x, y) \in \Delta A_k$. Mathematically,

$$\int_A \int F(x, y)\, \mathrm{d}A = \lim_{\Delta A \to 0} \sum_{k=1}^{n} F(x_k, y_k)\, \Delta A_k$$



(a) Subdividing the region A.



(b) Parts of $F(x, y)$ over a ΔA_k.

Figure 16.1: The double integral.

  - To conceptualize the double integral, first imagine that a region A of the plane is bounded above by the curve $y = f_2(x)$, below by the curve $y = f_1(x)$, on the left by the line $x = a$, and on the right by the line $x = b$ (see Figure 16.1a).
  - Now imagine that A is subdivided by a grid into n pieces $\Delta A = \Delta x \Delta y = \Delta y \Delta x$. We disregard the pieces that lie partially or entirely outside of the bounds.
  - As discussed in Chapter 15, $F(x, y)$ can be thought of as a surface in three-space. For the sake of simplicity, we will imagine for right now that $F(x, y)$ is positive for all $(x, y) \in A$, i.e., that it lies above A (see Figure 16.1b).
  - With this picture, we can imagine summing the partial volumes $F(x, y) \cdot \Delta A_k$ for each ΔA_k where (x, y) is some point in ΔA_k to approximate the total volume under the surface (analogous to the area under the curve).

  - All that the double integral does at this point is find the exact volume under the surface by taking the limit of this summation as we consider increasingly more increasingly small slivers of volume.

- We can evaluate the double integral of $F(x, y)$ over A if $F(x, y)$ is continuous throughout A, if the boundary curves are continuous and have finite total length, and if we let $\Delta y = 2\Delta x$ (or some other similar function) and let $\Delta x \to 0$.
- To evaluate double integrals, we calculate one or the other of the **iterated** integrals[1]

$$\int_A \int F(x,y)\,\mathrm{d}x\,\mathrm{d}y \qquad\qquad \int_A \int F(x,y)\,\mathrm{d}y\,\mathrm{d}x$$

- To evaluate the latter integral above, for example, we integrate "$\int F(x,y)\,\mathrm{d}y$ with respect to y (with x held fixed) and [evaluate] the resulting integral between the limits $y = f_1(x)$ and $y = f_2(x)$, and then [integrate] the result... with respect to x between the limits $x = a$ and $x = b$" (Thomas, 1972, p. 549). Mathematically,

$$\int_A \int F(x,y)\,\mathrm{d}y\,\mathrm{d}x = \int_a^b \left(\int_{f_1(x)}^{f_2(x)} F(x,y)\,\mathrm{d}y \right) \mathrm{d}x$$

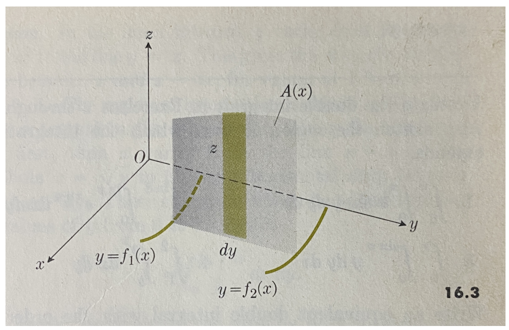


Figure 16.2: Visualizing iterated integration.

  - To visualize iterated integration, imagine determining the volume under the surface by summing infinitely many infinitely thin cross sections of the solid parallel to the y-axis.
  - The cross section at $x = x_0$ would have area $A(x_0) = \int_{f_1(x_0)}^{f_2(x_0)} F(x_0, y)\,\mathrm{d}y$.
  - The sum of all such cross sections' contributions to the volume would be the integral $\int_a^b A(x)\,\mathrm{d}x = \int_a^b \left(\int_{f_1(x)}^{f_2(x)} F(x,y)\,\mathrm{d}y \right) \mathrm{d}x$.

16.2 Area by Double Integration

- The area of the region of the xy-plane is given by either of the following integrals (with proper limits of integration).

$$A = \iint \mathrm{d}x\,\mathrm{d}y = \iint \mathrm{d}y\,\mathrm{d}x$$

- In cases such as that of Figure 16.1a, it makes sense to integrate with respect to y first and x second. However, if we have a region bounded by $y = c$, $y = d$, $x = g_1(y)$, and $x = g_2(y)$, then it would make more sense to do the opposite.
- Conceptualize the iterated integration here as summing the infinitesimal areas of strips parallel to the x- or y-axis, the lengths of which are given by $f_2(x) - f_1(x) = \int_{f_1(x)}^{f_2(x)} \mathrm{d}y$ or $g_2(y) - g_1(y) = \int_{g_1(y)}^{g_2(y)} \mathrm{d}x$.

[1] Proving that evaluating these integrals is equivalent to evaluating the double integral is a more advanced theorem of analysis.

16.3 Physical Applications

12/20:

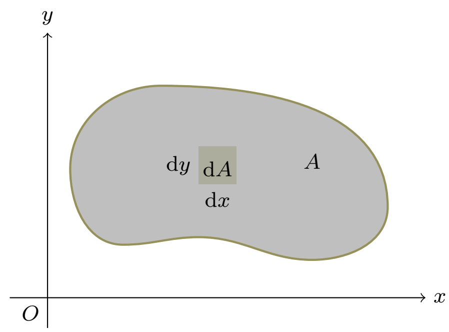


Figure 16.3: A planar mass.

- Imagine that there is a mass continuously distributed over a region A in the xy-plane, where the 2D density at every point $(x,y) \in A$ is given by $\delta(x,y)$.
- It follows that $\mathrm{d}m = \delta(x,y)\,\mathrm{d}A$, so the mass, the first moment of the mass with respect to the x-axis, and the first moment of the mass with respect to the y-axis are, respectively,

$$M = \iint \delta(x,y)\,\mathrm{d}A \qquad M_x = \iint y\,\delta(x,y)\,\mathrm{d}A \qquad M_y = \iint x\,\delta(x,y)\,\mathrm{d}A$$

- From these equations, we can also get the coordinates of the center of mass:

$$(x_{\text{cm}}, y_{\text{cm}}) = \left(\frac{M_y}{M}, \frac{M_x}{M} \right)$$

- The moments of inertia (second moments of the mass) with respect to the x- and y-axes, and the polar moment of inertia about the origin are, respectively,

$$I_x = \iint y^2\,\delta(x,y)\,\mathrm{d}A \qquad I_y = \iint x^2\,\delta(x,y)\,\mathrm{d}A \qquad I_0 = \iint r^2\,\delta(x,y)\,\mathrm{d}A$$

- Thomas (1972) defines the moment of inertia from a physics perspective, including the formula

$$I = \lim_{\Delta m \to 0} \sum r^2\,\Delta m = \int r^2\,\mathrm{d}m$$

  and then talks a bit about stiffness and why I-beams are strong.
- Statistical importance of moments:
  - The first moment helps compute the mean $\bar{r}$:

$$\bar{r} = \frac{\sum m_k r_k}{\sum m_k} = \frac{M_1}{M}$$

  - The second moment helps compute the variance σ^2 and standard deviation σ:

$$\sigma^2 = \frac{\sum (r_k - \bar{r})^2 m_k}{\sum m_k} = \frac{M_2}{M} - \bar{r}^2$$

    - Both σ^2 and σ are "measures of the way in which the r-values tend to bunch up close to $\bar{r}$ (small values of σ) or to be spread out (large values of σ)" (Thomas, 1972, p. 554).

- The third moment helps compute the **skewness**.
- The fourth moment helps compute the **kurtosis**.
- The t^{th} moment is defined as

$$M_t = \sum_{k=1}^{n} m_k r_k^t$$

  - r_k ranges over all values of the statistic under consideration, and m_k is the number of times it occurs (e.g., if 5 students score a 75% on a quiz, then $r_k = 75$ and $m_k = 5$ where a 75% is the k^{th} score earned).

- **Frequency distribution**: "A table of values of m_k versus r_k" (Thomas, 1972, p. 553).
  - M_t is "the t^{th} moment of this frequency distribution" (Thomas, 1972, p. 553).
- Note that area integrals such as $A = \int_a^b y \, \mathrm{d}x$ require substituting for y whereas double integrals do not, as we are integrating with respect to y in part, and the boundary curves are only applied as limits of integration. In double integrals, the point (x, y) is an element of $\mathrm{d}A$, and x, y are both independent variables.
- **Radius of gyration** (of an object of mass M about an axis where the object has moment of inertia I): The number $\sqrt{I/M}$ representing the radius a radial point particle (a loop) with equal mass would have to have in order to have the same moment of inertia as the object at hand.

16.4 Polar Coordinates

- **Double integral** (of $F(r, \theta)$ over the region A): The limit as $\Delta A \to 0$ of the sum of every ΔA_k (composing A) times $F(r_k, \theta_k)$ where (r_k, θ_k) is the point in the center[2] of ΔA_k. Mathematically,

$$\int_A \int F(r, \theta) \, \mathrm{d}A = \lim_{\Delta A \to 0} \sum_{k=1}^{n} F(r_k, \theta_k) \, \Delta A_k$$

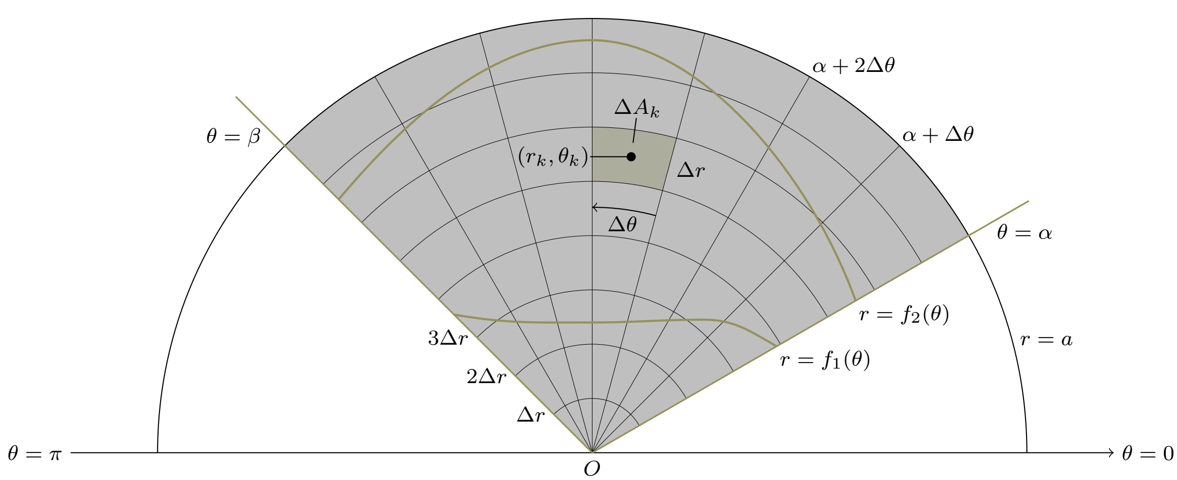


Figure 16.4: The double integral in polar coordinates.

[2]The intersection of the arc halfway between r and $r + \Delta r$ and the ray halfway between θ and $\theta + \Delta\theta$.

- If we subdivide A as in Figure 16.4, we find that

$$\begin{aligned}\Delta A_k &= \frac{1}{2}\left(r_k + \frac{1}{2}\Delta r\right)^2 \Delta\theta - \frac{1}{2}\left(r_k - \frac{1}{2}\Delta r\right)^2 \Delta\theta \\ &= r_k\,\Delta\theta\,\Delta r\end{aligned}$$

- Therefore, to evaluate the double integral of $F(r,\theta)$ over the region A, we calculate the iterated integral on the right below.

$$\int_A\int F(r,\theta)\,\mathrm{d}A = \int_{\theta=\alpha}^{\beta}\int_{r=f_1(\theta)}^{f_2(\theta)} F(r,\theta)\,r\,\mathrm{d}r\,\mathrm{d}\theta$$

- To change from xy- to uv- coordinates in a double integral, where we integrate over a region A in the xy-plane that is equivalent to the region G in the uv-plane and have $x = f(u,v)$, $y = g(u,v)$, use the formula

$$\int_A\int \phi(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_G\int \phi[f(u,v),g(u,v)]\frac{\partial(x,y)}{\partial(u,v)}\,\mathrm{d}u\,\mathrm{d}v$$

where

$$\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix}\frac{\partial x}{\partial u} & \frac{\partial x}{\partial v}\\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v}\end{vmatrix}$$

is the **Jacobian** of the transformation.

- For example, to switch from xy- to polar coordinates, we have $x = r\cos\theta$, $y = r\sin\theta$ and

$$\frac{\partial(x,y)}{\partial(r,\theta)} = \begin{vmatrix}\frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta}\\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta}\end{vmatrix} = \begin{vmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{vmatrix} = r(\cos^2\theta + \sin^2\theta) = r$$

so we get

$$\iint \phi(x,y)\,\mathrm{d}x\,\mathrm{d}y = \iint \phi(r\cos\theta, r\sin\theta)\,r\,\mathrm{d}r\,\mathrm{d}\theta$$

16.5 Triple Integrals: Volume

- **Riemann triple integral** (of $F(x,y,z)$ over V): The limit as $\Delta V \to 0$ of the sum of every ΔV_k (composing V) times $F(x,y,z)$ for some point $(x,y,z) \in \Delta V_k$. Mathematically,

$$\iiint_V F(x,y,z)\,\mathrm{d}V = \lim_{\Delta V\to 0}\sum_{k=1}^{n} F(x_k,y_k,z_k)\,\Delta V_k$$

- Some interpretations of the triple integral:
  - $F(x,y,z) = 1$: The volume of V.
  - $F(x,y,z) = x$: The first moment of the volume V with respect to the yz-plane.
  - $F(x,y,z)$ is the density at (x,y,z): The mass in V.
  - $F(x,y,z)$ is "the product of the density at (x,y,z) and the square of the distance from (x,y,z) to an axis L" (Thomas, 1972, p. 558): The moment of inertia of the mass with respect to L.
- To evaluate triple integrals, we calculate an iterated integral of the following form, where $f_1(x,y)$ and $f_2(x,y)$ are surfaces that enclose V on the top and bottom, and A is the region of the xy-plane enclosed by the cylinder C that bounds V laterally with elements parallel to the z-axis.

$$\int_A\iint_{f_1(x,y)}^{f_2(x,y)} F(x,y,z)\,\mathrm{d}z\,\mathrm{d}y\,\mathrm{d}x$$

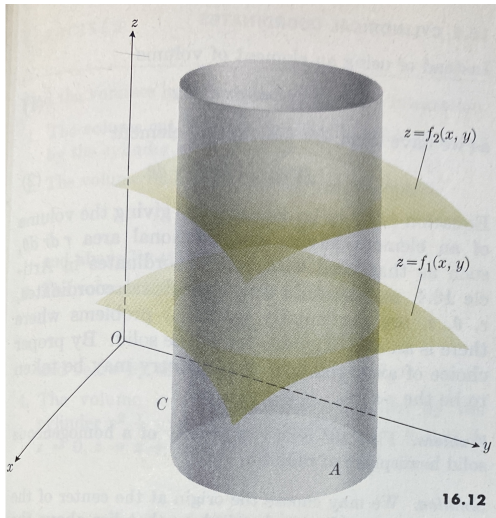


Figure 16.5: The triple integral.

- The first integral is a normal integral with x, y held constant, and the second one is a double integral over A.

- "Find the volume enclosed between the two surfaces $z = 8 - x^2 - y^2$ and $z = x^2 + 3y^2$" (Thomas, 1972, p. 559).

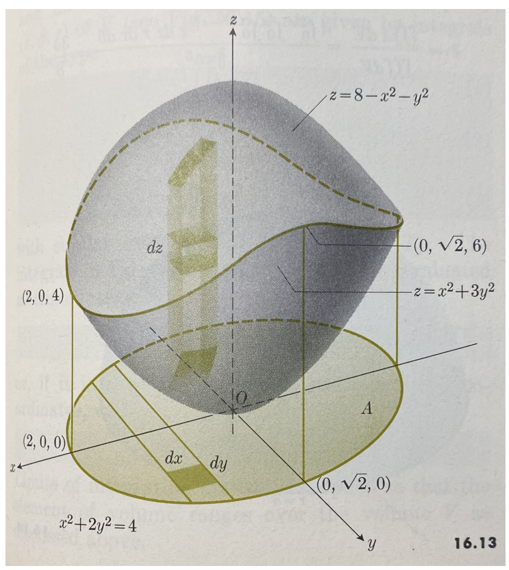


Figure 16.6: Volume by triple integration.

- By eliminating z in the two given equations, we can find the elliptic cylinder enclosing V laterally.

$$8 - x^2 - y^2 = x^2 + 3y^2$$
$$x^2 + 2y^2 = 4$$

– Thus, to find volume, we evaluate the integral[3]

$$V = \int_{-2}^{2} \int_{-\sqrt{(4-x^2)/2}}^{\sqrt{(4-x^2)/2}} \int_{x^2+3y^2}^{8-x^2-y^2} \mathrm{d}z\,\mathrm{d}y\,\mathrm{d}x = 8\pi\sqrt{2}$$

16.6 Cylindrical Coordinates

- "Instead of using an element of volume $\mathrm{d}V_{xyz} = \mathrm{d}z\,\mathrm{d}y\,\mathrm{d}x$ as we have done, we may use an element $\mathrm{d}V_{r\theta z} = \mathrm{d}z\ r\,\mathrm{d}r\,\mathrm{d}\theta$" (Thomas, 1972, p. 560).
- Cylindrical coordinates r, θ, z are particularly useful in problems with an axis of radial symmetry, which we can make to be the z-axis.
- "Find the center of gravity of a homogenous solid hemisphere of radius a" (Thomas, 1972, p. 560).

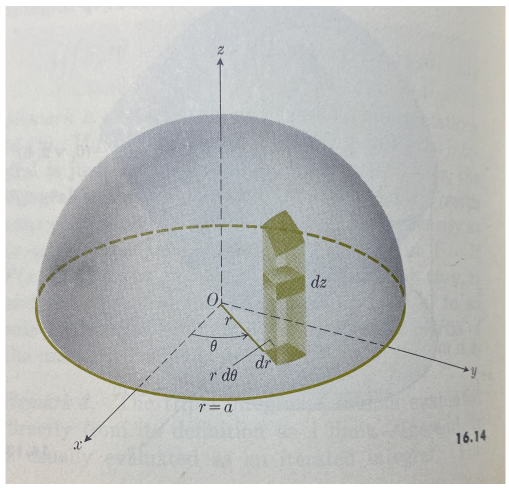


Figure 16.7: Center of mass of a hemisphere by triple integration.

– We let the base of the hemisphere rest on the xy-plane so that it exhibits radial symmetry about the z-axis, as in Figure 16.7.

– By symmetry, we have $x_{\text{cm}} = y_{\text{cm}} = 0$.

– As to z_{cm}, in cylindrical coordinates, the height of the hemisphere above the xy-plane is given by $\sqrt{a^2 - r^2}$. Thus, we evaluate

$$z_{\text{cm}} = \frac{\iiint z\,\mathrm{d}V}{\iiint \mathrm{d}V} = \frac{\int_0^{2\pi} \int_0^a \int_0^{\sqrt{a^2-r^2}} z\,\mathrm{d}z\ r\,\mathrm{d}r\,\mathrm{d}\theta}{\frac{2}{3}\pi a^3} = \frac{3a}{8}$$

16.7 Physical Applications of Triple Integration

- The mass, center of gravity, and moment of inertia of a mass M distributed over a region V of xyz-space with density $\delta(x, y, z)$ at the point (x, y, z) are, respectively,

$$M = \iiint \delta\,\mathrm{d}V \qquad x_{\text{cm}} = \frac{\iiint x\,\delta\,\mathrm{d}V}{\iiint \delta\,\mathrm{d}V} \qquad I_z = \iiint (x^2 + y^2)\,\delta\,\mathrm{d}V$$

[3] One good way to think about this integral is as dividing V up into cross sections parallel to the yz-plane at each $x_0 \in [-2, 2]$, taking the double integral of $F(x_0, y, z)$ over this cross section, and then summing (integrating) all of these cross sections.

16.8 Spherical Coordinates

- Spherical coordinates ρ, ϕ, θ are particularly useful in problems with symmetry about a point, which we can make the origin.

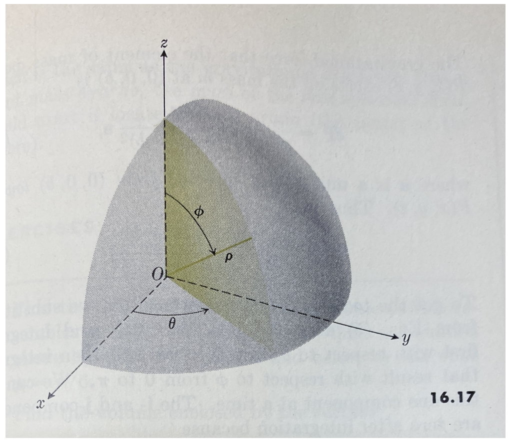


Figure 16.8: Spherical coordinates.

- We interconvert between spherical and Cartesian coordinates with the equations

$$x = \rho \sin\phi \cos\theta \qquad y = \rho \sin\phi \sin\theta \qquad z = \rho \cos\phi$$

- Here, we substitute a volume element $\mathrm{d}V_{xyz}$ for a volume element

$$\begin{aligned} \mathrm{d}V_{\rho\phi\theta} &= \mathrm{d}\rho \cdot \rho\,\mathrm{d}\phi \cdot \rho \sin\phi\,\mathrm{d}\theta \\ &= \rho^2 \sin\phi\,\mathrm{d}\rho\,\mathrm{d}\phi\,\mathrm{d}\theta \end{aligned}$$

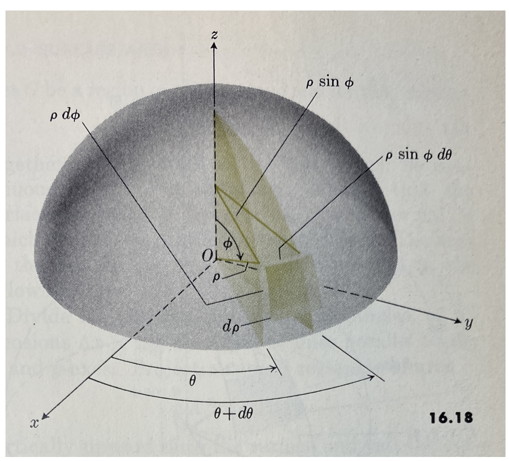


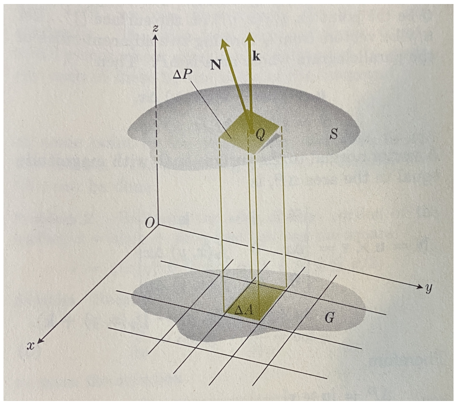
Figure 16.9: Infinitesimal volume element of spherical coordinates.

  - Figure 16.9 visually justifies the choice of these lengths.

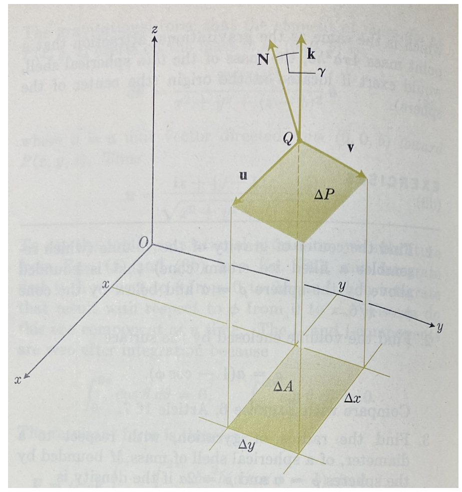
- Thomas (1972) proves that the gravitational attraction of a planet is the same as a point particle of identical mass located at the planet's center.

16.9 Surface Area

- If G is a region of the xy-plane in which $z = f(x, y)$ and its first partial derivatives are continuous, then we can find the surface area of the surface S defined by $f(x, y)$ lying over G.



(a) Analyzing the surface.



(b) Analyzing the approximation.

Figure 16.10: Surface area of a surface defined by a function of two variables.

  - We begin by dividing G into rectangles parallel to the xy-axes as in Figure 16.10a.
  - Consider one such rectangle ΔA whose four vertices are $(x, y, 0)$, $(x + \Delta x, y, 0)$, $(x, y + \Delta y, 0)$, and $(x + \Delta x, y + \Delta y, 0)$, and let $Q(x, y, f(x, y))$ be the point above $(x, y, 0)$ on S. We are going to approximate the area ΔS of the part of S above ΔA with the area ΔP of the tangent plane to Q above ΔA.
  - We can define two vectors $\mathbf{u}$ and $\mathbf{v}$ that lie along two of the edges of ΔP, as in Figure 16.10b, as follows.

$$\mathbf{u} = \mathbf{i}\Delta x + \mathbf{k} f_x(x, y)\Delta x \qquad\qquad \mathbf{v} = \mathbf{j}\Delta y + \mathbf{k} f_y(x, y)\Delta y$$

  - The cross product of these vectors gives both a normal vector $\mathbf{N}$ to P (and to S at Q), and the area of ΔP (via its magnitude).

$$\begin{aligned} \mathbf{N} &= \mathbf{u} \times \mathbf{v} \\ &= \Delta x\, \Delta y(-\mathbf{i} f_x(x, y) - \mathbf{j} f_y(x, y) + \mathbf{k}) \end{aligned} \qquad\qquad \begin{aligned} \Delta P &= |\mathbf{u} \times \mathbf{v}| \\ &= \Delta x\, \Delta y \sqrt{f_x^2(x, y) + f_y^2(x, y) + 1} \end{aligned}$$

  - Now that we have an analytical formula for ΔP, we know that the surface area of S is given by[4]

$$SA = \lim_{\Delta x, \Delta y \to 0} \sum_G \Delta P = \int_G \int \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + 1}\, \mathrm{d}x\, \mathrm{d}y$$

- Since $(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{k}$ is equal to the area of the projection of ΔP in the xy-plane (this can be confirmed by the fact that $[\Delta x\, \Delta y(-\mathbf{i} f_x(x, y) - \mathbf{j} f_y(x, y) + \mathbf{k})] \cdot \mathbf{k} = \Delta x\, \Delta y = \Delta A$), and since $(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{k} = |\mathbf{u} \times \mathbf{v}|\, |\mathbf{k}| \cos\gamma = \Delta P \cos\gamma$, we know that $\Delta P = \Delta A / \cos\gamma$, meaning that

$$SA = \lim_{\Delta x, \Delta y \to 0} \sum_G \Delta P = \int_G \int \frac{\mathrm{d}A}{\cos\gamma}$$

[4]It makes sense that this formula is similar to the single-variable calculus arc length formula since surface area is analogous to arc length, just one dimension higher.

  - If the equation of S is given in the form $F(x, y, z) = 0$, then $\mathbf{N} = \nabla F$ and $\cos\gamma = \mathbf{N} \cdot \mathbf{k}/(|\mathbf{N}| \cdot |\mathbf{k}|)$

- Discusses numerical approximations of the surface area integral.

Chapter 17

Vector Analysis

17.1 Introduction: Vector Fields

12/29:

- In this chapter, we will consider vector functions of several variables, such as the function giving the velocity $\mathbf{v} = \mathbf{F}(x, y, z, t)$ of a particle in a fluid located at position (x, y, z) at time t.
- **Steady-state flow**: A flow for which the velocity function does not depend on the time t.
- **Vector field**: The collection of all vectors $\mathbf{F}(P)$ assigned to each point P in a region G.
- **Gradient field**: The vector field defined for points in the domain G of a scalar function T such that $\mathbf{F}(P) = \nabla T(P)$.

17.2 Surface Integrals

12/30:

- Just like we have $\mathrm{d}s = \sqrt{1 + f_x^2}\,\mathrm{d}x$, we have

$$\mathrm{d}\sigma = g(x, y)\,\mathrm{d}A$$

  where $\mathrm{d}\sigma$ is "an element of surface area in the tangent plane that approximates the corresponding portion $\Delta\sigma$ of the surface itself" (Thomas, 1972, p. 581) and $g(x, y) = \sqrt{1 + f_x^2 + f_y^2}$.

- Thus, we can think of surface area as either the lefthand or righthand side of the below equation.

$$\iint\limits_{\Sigma} \mathrm{d}\sigma = \iint\limits_{R} g(x, y)\,\mathrm{d}A$$

  - The lefthand interpretation sums infinitely many, infinitely small pieces $\mathrm{d}\sigma$ of the surface Σ.
  - The righthand interpretation sums infinitely many, infinitely small pieces $\mathrm{d}A$ of the shadow R of the surface Σ on the xy-plane, adjusted by the factor $g(x, y)$.

- These formulations are important because sometimes we want to conceive and evaluate an integral of the form $\iint_\Sigma h(x, y, z)\,\mathrm{d}\sigma$.
- **Surface integral** (of $h(x, y, z)$ over the surface Σ): The limit as $\Delta\sigma \to 0$ of the sum of every $\Delta\sigma_k$ (composing Σ) times $h(x, y, z)$ for some $(x, y, z) \in \Delta\sigma_k$. Mathematically,

$$\iint\limits_{\Sigma} h(x, y, z)\,\mathrm{d}\sigma = \lim_{\Delta\sigma \to 0} \sum_{k=1}^{n} h(x_k, y_k, z_k)\,\Delta\sigma_k$$

- Consider a surface Σ consisting of all points $P(x,y,z)$ satisfying $z = f(x,y)$ for $(x,y) \in R$, where R is a closed, bounded region of the xy-plane and f, f_x, f_y are continuous throughout R and its boundary.
- Approximate R by dividing it into n rectangles using lines parallel to the y-axis spaced Δx apart and lines parallel to the x-axis spaced Δy apart.
- Let the part of Σ above each rectangle be denoted by $\Delta\sigma_k$ for some $1 \leq k \leq n$.
- Now if $P_k(x_k, y_k, z_k)$ is a point in $\Delta\sigma_k$, we can consider the above sum and take its limit.

12/31: • To evaluate the surface integral, we substitute $\Delta\sigma_k = g(x_k, y_k)\,\Delta x\,\Delta y$ and $z_k = f(x_k, y_k)$ in the sum, and take iterated integrals over R (the shadow of Σ on the xy-plane) instead of Σ.

$$\iint\limits_{\Sigma} h(x,y,z)\,\mathrm{d}\sigma = \iint\limits_{R} h[x,y,f(x,y)]g(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

- We now explore a useful surface integration technique through a problem.
- Evaluate $\iint (x^2+y^2)\,\mathrm{d}\sigma$ over the hemisphere Σ described by $z = \sqrt{a^2-x^2-y^2}$.
  - Because of a *sphere* 2Σ of radius a's high degree of symmetry,

$$\iint\limits_{2\Sigma} x^2\,\mathrm{d}\sigma = \iint\limits_{2\Sigma} y^2\,\mathrm{d}\sigma = \iint\limits_{2\Sigma} z^2\,\mathrm{d}\sigma = \frac{1}{3}\iint\limits_{2\Sigma} (x^2+y^2+z^2)\,\mathrm{d}\sigma = \frac{1}{3}\iint\limits_{2\Sigma} a^2\,\mathrm{d}\sigma$$

Thus, for the *hemisphere* Σ,

$$\begin{aligned}
\iint\limits_{\Sigma} (x^2+y^2)\,\mathrm{d}\sigma &= \frac{1}{2}\iint\limits_{2\Sigma} (x^2+y^2)\,\mathrm{d}\sigma \\
&= \frac{1}{2}\left(\iint\limits_{2\Sigma} x^2\,\mathrm{d}\sigma + \iint\limits_{2\Sigma} y^2\,\mathrm{d}\sigma\right) \\
&= \frac{1}{2}\left(\frac{1}{3}\iint\limits_{2\Sigma} a^2\,\mathrm{d}\sigma + \frac{1}{3}\iint\limits_{2\Sigma} a^2\,\mathrm{d}\sigma\right) \\
&= \frac{a^2}{3}\iint\limits_{2\Sigma} \mathrm{d}\sigma \\
&= \frac{a^2}{3}\cdot 4\pi a^2 \\
&= \frac{4}{3}\pi a^4
\end{aligned}$$

- Alternate formulations of $\mathrm{d}\sigma$.
  - Let the surface Σ be defined by the equation $F(x,y,z) = 0$.
  - For the same reasons discussed in Chapter 17,

$$\mathrm{d}\sigma = \frac{\mathrm{d}A}{\cos\phi}$$

where ϕ is the angle between $\mathbf{N} = \nabla F$ and the unit vector normal to the plane onto which Σ is projected, which we will take to be the xy-plane at first (this means that this normal vector is $\mathbf{k}$).

- Since

$$\cos\phi = \frac{\mathbf{N}\cdot\mathbf{k}}{|\mathbf{N}|\,|\mathbf{k}|} = \frac{|F_z|}{\sqrt{F_x^2+F_y^2+F_z^2}}$$

we thus have that

$$\mathrm{d}\sigma = \frac{\sqrt{F_x^2+F_y^2+F_z^2}}{|F_z|}\,\mathrm{d}x\,\mathrm{d}y$$

- Note that if we project Σ onto a different plane, an analog to the above can easily be derived.

17.3 Line Integrals

- **Line integral** (of $w(x,y,z)$ along the curve C from A to B): The limit as $\Delta s \to 0$ of the sum of every Δs_k (composing the section of C between points A and B along C) times $w(x,y,z)$ for some $(x,y,z) \in \Delta s_k$. Mathematically,

$$\int_C w\,\mathrm{d}s = \lim_{\Delta s\to 0}\sum_{k=1}^{n} w(x_k,y_k,z_k)\,\Delta s_k$$

  - Suppose that C is a directed curve in three-space from A to B. Let $w(x,y,z)$ be a scalar function of position that is continuous in a region D containing C.
  - Divide C into n segments, and let $P_k(x_k,y_k,z_k)$ be an arbitrary point on the kth subarc.
  - If the above sum has a limit as $n\to\infty$ and the largest $\Delta s_k \to 0$, and if this limit is the same for all ways of subdividing C and all choices of the points P_k, then we call this limit the line integral.
- If C is parameterized by the functions $x=f(t)$, $y=g(t)$, and $z=h(t)$ for $t_A \leq t \leq t_B$, where f,g,h are continuous and have bounded and piecewise-continuous first derivatives on $[t_A,t_B]$, then we may evaluate the line integral of $w(x,y,z)$ along C from A to B with the following formula.

$$\int_C w\,\mathrm{d}s = \int_{t_A}^{t_B} w[f(t),g(t),h(t)]\sqrt{\left(\frac{\mathrm{d}f}{\mathrm{d}t}\right)^2+\left(\frac{\mathrm{d}g}{\mathrm{d}t}\right)^2+\left(\frac{\mathrm{d}h}{\mathrm{d}t}\right)^2}\,\mathrm{d}t$$

  - Note that the line integral is the same for any appropriate parameterization of C, or no parameterization.

9/2:

- The line integral can be geometrically interpreted as the area of the region R that lies above the curve C, offset by distance w.
- If C is a straight line, we can take the line integral over it directly wrt. s by expressing f in terms of s and rewriting the limits:
- For example, "let C be the line segment from $A(0,0)$ to $B(1,1)$ and let $w=x+y^2$. Evaluate $\int_C w\,\mathrm{d}s$" (Thomas, 1972, p. 585).
  - Let $x=t$ and $y=t$ for $0\leq t\leq 1$. Then

$$\begin{aligned}\int_C w\,\mathrm{d}s &= \int_0^1 (t+t^2)\sqrt{1+1}\,\mathrm{d}t\\ &= \sqrt{2}\left[\frac{t^2}{2}+\frac{t^3}{3}\right]_0^1\\ &= \frac{5\sqrt{2}}{6}\end{aligned}$$

- By the Pythagorean theorem, $s = \sqrt{x^2 + y^2} = \sqrt{2x^2} = x\sqrt{2}$. Thus, $w = s/\sqrt{2} + s^2/2$. Additionally, as $0 \leq x \leq 1$, $0 \leq s \leq \sqrt{2}$. Therefore,

$$\begin{aligned}\int_C w \, \mathrm{d}s &= \int_0^{\sqrt{2}} \left(\frac{s}{\sqrt{2}} + \frac{s^2}{2} \right) \mathrm{d}s \\ &= \left[\frac{s^2}{2\sqrt{2}} + \frac{s^3}{6} \right]_0^{\sqrt{2}} \\ &= \frac{5\sqrt{2}}{6}\end{aligned}$$

- To generalize the above notion, we can always think of w as a function $\phi(s)$, where s is arc length.
- "If the point of application of a force $\mathbf{F} = \mathbf{i}M(x,y,z) + \mathbf{j}N(x,y,z) + \mathbf{k}P(x,y,z)$ moves along a curve C from a point $A(a_1, a_2, a_3)$ to a point $B(b_1, b_2, b_3)$, then the work done by the force is

$$W = \int_C \mathbf{F} \cdot \mathrm{d}\mathbf{R}$$

where $\mathbf{R}$ [is the position vector]" (Thomas, 1972, p. 586).

  - Since $\mathrm{d}\mathbf{R} = \frac{\mathrm{d}\mathbf{R}}{\mathrm{d}s} \mathrm{d}s$ and $\mathrm{d}\mathbf{R}/\mathrm{d}s = \mathbf{T}$, the work can also be thought of as "the value of the line integral along C of the tangential component of the force field $\mathbf{F}$" (Thomas, 1972, p. 587).

- The line integral between two points A and B is independent of the path C joining them if and only if the force field $\mathbf{F}$ is a **gradient field**, that is, if

$$\mathbf{F}(x,y,z) = \nabla f$$

for some differentiable function f.

  - Thomas (1972) proves this.
  - If $\mathbf{F}$ is a gradient field, then

$$\int_A^B \mathbf{F} \cdot \mathrm{d}\mathbf{R} = \int_A^B \nabla f \cdot \mathrm{d}\mathbf{R} = f(B) - f(A)$$

  - Furthermore, from $\mathbf{F}$, we define f by

$$f(x', y', z') = \int_A^{(x',y',z')} \mathbf{F} \cdot \mathrm{d}\mathbf{R}$$

- "Find a function f such that if $\mathbf{F} = 2x\mathbf{i} + 2y\mathbf{j} + 2z\mathbf{k}$, then $\mathbf{F} = \nabla f$" (Thomas, 1972, p. 589).
  - Choose $A = (0,0,0)$ to simplify calculations.
  - Assume that $\mathbf{F}$ is a gradient field, i.e., that evaluating $\int_A^{(x',y',z')} \mathbf{F} \cdot \mathrm{d}\mathbf{R}$ along any path will yield the same result.
  - Thus, choose to evaluate the line integral along the line segment from A to (x', y', z'), which we may define by the parameterization $x = x't$, $y = y't$, $z = z't$ for $0 \leq t \leq 1$.
  - Therefore,

$$\begin{aligned}f(x', y', z') &= \int_{(0,0,0)}^{(x',y',z')} \mathbf{F} \cdot \mathrm{d}\mathbf{R} \\ &= \int_{(0,0,0)}^{(x',y',z')} (2x\mathbf{i} + 2y\mathbf{j} + 2z\mathbf{k}) \cdot (\mathbf{i} \, \mathrm{d}x + \mathbf{j} \, \mathrm{d}y + \mathbf{k} \, \mathrm{d}z)\end{aligned}$$

$$\begin{aligned}
&= \int_0^1 (2x't\mathbf{i} + 2y't\mathbf{j} + 2z't\mathbf{k}) \cdot (\mathbf{i}x'\,\mathrm{d}t + \mathbf{j}y'\,\mathrm{d}t + \mathbf{k}z'\,\mathrm{d}t) \\
&= \int_0^1 (2x'^2t\,\mathrm{d}t + 2y'^2t\,\mathrm{d}t + 2z'^2t\,\mathrm{d}t) \\
&= (x'^2 + y'^2 + z'^2)\int_0^1 2t\,\mathrm{d}t \\
&= x'^2 + y'^2 + z'^2
\end{aligned}$$

- **Conservative** (force field): A force field $\mathbf{F}$ such that the work integral from A to B is the same for all paths joining them.
- Another criterion besides $\mathbf{F} = \nabla f$ for some differentiable f is that $\mathrm{d}f = \mathbf{F}\cdot\mathrm{d}\mathbf{R} = M\,\mathrm{d}x + N\,\mathrm{d}y + P\,\mathrm{d}z$ is an exact differential.
  - By an extension of Theorem 15.4, we know that $M\,\mathrm{d}x + N\,\mathrm{d}y + P\,\mathrm{d}z$ is an exact differential if and only if

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x} \qquad \frac{\partial M}{\partial z} = \frac{\partial P}{\partial x} \qquad \frac{\partial N}{\partial z} = \frac{\partial P}{\partial y}$$

- "Suppose $\mathbf{F} = \mathbf{i}(\mathrm{e}^x\cos y + yz) + \mathbf{j}(xz - \mathrm{e}^x\sin y) + \mathbf{k}(xy + z)$. Is $\mathbf{F}$ conservative? If so, find f such that $\mathbf{F} = \nabla f$" (Thomas, 1972, p. 591).
  - Apply the exact differential test:

$$\frac{\partial M}{\partial y} = -\mathrm{e}^x\sin y + z = \frac{\partial N}{\partial x} \qquad \frac{\partial M}{\partial z} = y = \frac{\partial P}{\partial x} \qquad \frac{\partial N}{\partial z} = x = \frac{\partial P}{\partial y}$$

    - Therefore, $\mathbf{F}$ is conservative.
  - To calculate f, we integrate the system of equations

$$\frac{\partial f}{\partial x} = \mathrm{e}^x\cos y + yz \qquad \frac{\partial f}{\partial y} = xz - \mathrm{e}^x\sin y \qquad \frac{\partial f}{\partial z} = xy + z$$

    - Starting with the first one, we obtain

$$f = \mathrm{e}^x\cos y + xyz + g(y,z)$$

      where $g(y,z)$ is a function of integration.
    - Differentiating wrt. y, we get

$$\begin{aligned}
xz - \mathrm{e}^x\sin y = \frac{\partial f}{\partial y} &= -e[x]\sin y + xz + \frac{\partial g}{\partial y} \\
\frac{\partial g}{\partial y} &= 0 \\
g(y,z) &= h(z)
\end{aligned}$$

    - Differentiating wrt. z, we get

$$\begin{aligned}
xy + z = \frac{\partial f}{\partial z} &= xy + \frac{\partial h}{\partial z} \\
\frac{\partial h}{\partial z} &= z \\
h(z) &= \frac{1}{2}z^2 + C
\end{aligned}$$

    - Therefore,

$$f(x,y,z) = \mathrm{e}^x\cos y + xyz + \frac{1}{2}z^2 + C$$

- **Potential function**: A function $f(x,y,z)$ which has the property that its gradient gives the force vector $\mathbf{F}$.

17.4 Two-Dimensional Fields: Line Integrals in the Plane and Their Relation to Surface Integrals on Cylinders

- Features of a two-dimensional field $\mathbf{F}$:
  1. The vectors in $\mathbf{F}$ are all parallel to one plane, which we have taken to be the xy-plane.
     - Mathematically, the vectors have no $\mathbf{k}$ component.
  2. In every plane parallel to the xy-plane, the field is the same as it is in that plane.
     - Mathematically, the vectors do not depend on z.
- Imagine fluid of planar mass density δ flowing out from the origin with velocity defined by a vector velocity function $\mathbf{v}$.

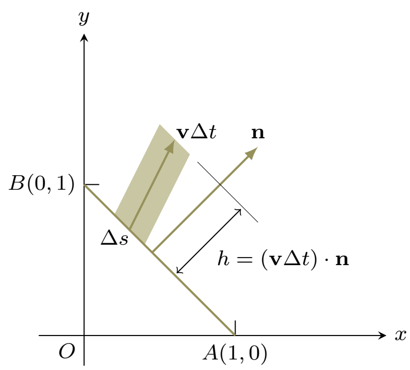


Figure 17.1: Fluid flowing over a line segment.

  - Then the flow rate $\mathrm{d}M/\mathrm{d}t$ over a curve C in the plane is given by

$$\frac{\mathrm{d}M}{\mathrm{d}t} = \int_C \delta(\mathbf{v}\cdot\mathbf{n})\,\mathrm{d}s$$

- **Flux** (of $\mathbf{F} = \delta\mathbf{v}$ across C): The quantity

$$\int_C \mathbf{F}\cdot\mathbf{n}\,\mathrm{d}s$$

- If C is a closed curve, we canonically choose $\mathbf{n}$ to point outwards and the orientation to be in the counterclockwise direction.
  - We also choose $\mathbf{n} = \mathbf{T}\times\mathbf{k}$.
- With these conventions, if we let $\mathbf{F}(x,y) = \mathbf{i}M(x,y) + \mathbf{j}N(x,y)$, then

$$\begin{aligned}
\text{flux} &= \int_C \mathbf{F}\cdot\mathbf{n}\,\mathrm{d}s \\
&= \int_C \mathbf{F}\cdot(\mathbf{T}\times\mathbf{k})\,\mathrm{d}s \\
&= \int_C \mathbf{F}\cdot\left(\left(\frac{\mathrm{d}x}{\mathrm{d}s}\mathbf{i} + \frac{\mathrm{d}y}{\mathrm{d}s}\mathbf{j}\right)\times\mathbf{k}\right)\mathrm{d}s \\
&= \int_C \mathbf{F}\cdot\left(\frac{\mathrm{d}y}{\mathrm{d}s}\mathbf{i} - \frac{\mathrm{d}x}{\mathrm{d}s}\mathbf{j}\right)\mathrm{d}s \\
&= \int_C \left(M\frac{\mathrm{d}y}{\mathrm{d}s} - N\frac{\mathrm{d}x}{\mathrm{d}s}\right)\mathrm{d}s \\
&= \int_C (M\,\mathrm{d}y - N\,\mathrm{d}x)
\end{aligned}$$

- Any flux integral can be reinterpreted as a work integral on a related field and vice versa: If $\mathbf{F}(x, y) = \mathbf{i}M + \mathbf{j}N$ and $\mathbf{G}(x, y) = -\mathbf{i}N + \mathbf{j}M$, then

$$\int_C \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}s = \int_C \mathbf{G} \cdot \mathbf{T} \, \mathrm{d}s$$

17.5 Green's Theorem

- The following is a formal statement of **Green's theorem**.

Theorem 17.1 (Green's theorem)**.** Let C be a simple closed curve in the xy-plane such that a line parallel to either axis cuts C in at most two points. Let M, N, $\partial N/\partial x$, and $\partial M/\partial y$ be continuous functions of (x, y) inside and on C. Let R be the region inside C. Then

$$\oint_C M \, \mathrm{d}x + N \, \mathrm{d}y = \iint_R \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} \right) \mathrm{d}x \, \mathrm{d}y \text{ [1]}$$

Proof. We will prove that $\iint_R (-\,\partial M/\partial y)\, \mathrm{d}x \, \mathrm{d}y = \oint_C M \, \mathrm{d}x$. It will follow by a symmetric argument that $\iint_R (\,\partial N/\partial x)\, \mathrm{d}x \, \mathrm{d}y = \oint_C N \, \mathrm{d}y$. The sum of these two qualities will yield Green's theorem. Let's begin.

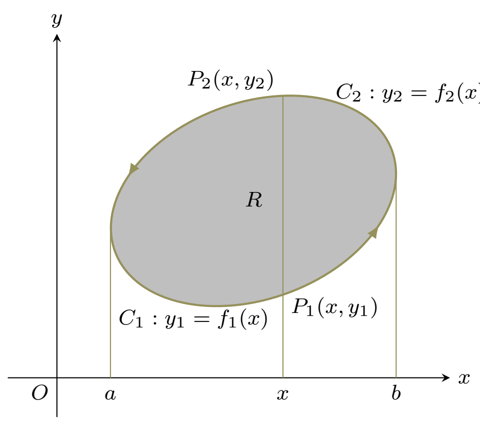


Figure 17.2: Proving Green's theorem.

Consider the curve C enclosing a region R. We divide C into a lower boundary curve C_1 and an upper boundary curve C_2, both of which are functions of x (the constraint that a line parallel to either axis cuts C in at most two points allows us to do this).

Since $\partial M/\partial y$ is continuous, it is integrable, meaning that at any $x \in [a, b]$, we can determine that

$$\begin{aligned} \int_{y_1}^{y_2} \frac{\partial M}{\partial y} \, \mathrm{d}y &= [M(x, y)]_{y=f_1(x)}^{y=f_2(x)} \\ &= M(x, f_2(x)) - M(x, f_1(x)) \end{aligned}$$

[1] The symbol $\oint$ denotes a line integral over a closed curve C.

It follows since M is continuous, and therefore integrable, that

$$\begin{aligned}
\iint_R -\frac{\partial M}{\partial y}\,\mathrm{d}x\,\mathrm{d}y &= \int_a^b -\int_{f_1(x)}^{f_2(x)} \frac{\partial M}{\partial y}\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_a^b [M(x, f_1(x)) - M(x, f_2(x))]\,\mathrm{d}x \\
&= \int_a^b [M(x, f_1(x))\,\mathrm{d}x + \int_b^a M(x, f_2(x))]\,\mathrm{d}x \\
&= \int_{C_1} M\,\mathrm{d}x + \int_{C_2} M\,\mathrm{d}x \\
&= \oint_C M\,\mathrm{d}x
\end{aligned}$$

as desired.

It follows by a symmetric argument that

$$\iint_R \frac{\partial N}{\partial x}\,\mathrm{d}x\,\mathrm{d}y = \oint_C N\,\mathrm{d}y$$

Therefore, we have by addition that

$$\oint_C M\,\mathrm{d}x + N\,\mathrm{d}y = \iint_R \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right)\mathrm{d}x\,\mathrm{d}y$$ [2]

as desired. ☐

- Green's theorem provides an easy to calculate the area enclosed by a curve.

Corollary 17.1. *If C is a simple closed curve such that a line parallel to either axis cuts it in at most two points, them the area enclosed by C is equal to*

$$\frac{1}{2}\oint_C (x\,\mathrm{d}y - y\,\mathrm{d}x)$$

Proof. From Section 16.2, we have that

$$\begin{aligned}
A &= \iint_R 1\,\mathrm{d}x\,\mathrm{d}y \\
&= \iint_R \left(\frac{1}{2} - \left(-\frac{1}{2}\right)\right)\mathrm{d}x\,\mathrm{d}y \\
&= \iint_R \left(\frac{\partial}{\partial x}\left(\frac{x}{2}\right) - \frac{\partial}{\partial y}\left(-\frac{y}{2}\right)\right)\mathrm{d}x\,\mathrm{d}y \\
&= \oint_C \left(-\frac{1}{2}y\,\mathrm{d}x + \frac{1}{2}x\,\mathrm{d}y\right) && \text{Theorem 17.1} \\
&= \frac{1}{2}\oint_C (x\,\mathrm{d}y - y\,\mathrm{d}x)
\end{aligned}$$

as desired. ☐

- Note that Green's theorem also applies to a number of shapes that don't fit the theorem statement's direct criteria.
  - For instance, we can prove that it holds for a rectangle in the xy-plane with sides parallel to the x- or y-axes.

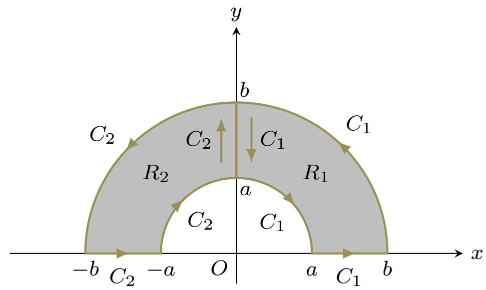


Figure 17.3: Creating composite regions that satisfy Green's theorem.

- Additionally, we can add together regions that satisfy Green's theorem individually to form bigger regions that satisfy it (the line integrals in the overlapping part of Figure 17.3 cancel).
- In fact, we can add together any finite number of subregions that satisfy Green's theorem.

- **Curl** (of a vector $\mathbf{F} = M\mathbf{i} + N\mathbf{j} + P\mathbf{k}$): The cross product of the del operator and $\mathbf{F}$. *Given by*

$$\begin{aligned}
\operatorname{curl}\mathbf{F} &= \nabla \times \mathbf{F} \\
&= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ M & N & P \end{vmatrix} \\
&= \mathbf{i}\left(\frac{\partial P}{\partial y} - \frac{\partial N}{\partial z}\right) + \mathbf{j}\left(\frac{\partial M}{\partial z} - \frac{\partial P}{\partial x}\right) + \mathbf{k}\left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right)
\end{aligned}$$

- It follows that Green's theorem in vector form is

$$\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{R} = \iint_R (\nabla \times \mathbf{F}) \cdot \mathrm{d}\mathbf{A}$$

where $\mathbf{F} = M\mathbf{i} + N\mathbf{j} + P\mathbf{k}$, $\mathrm{d}\mathbf{R} = \mathbf{i}\,\mathrm{d}x + \mathbf{j}\,\mathrm{d}y$, and $\mathrm{d}\mathbf{A} = \mathbf{k}\,\mathrm{d}x\,\mathrm{d}y$.

  - "In words, Green's theorem states that the integral around C of the tangential component of $\mathbf{F}$ is equal to the integral, over the region R bounded by C, of the component of $\operatorname{curl}\mathbf{F}$ that is normal to R; this integral, specifically, is the flux through R of $\operatorname{curl}\mathbf{F}$" (Thomas, 1972, p. 604).

- **Divergence** (of a vector $\mathbf{F} = M\mathbf{i} + N\mathbf{j} + P\mathbf{k}$): The dot product of the del operator and $\mathbf{F}$. *Given by*

$$\begin{aligned}
\operatorname{div}\mathbf{G} &= \nabla \cdot \mathbf{G} \\
&= \frac{\partial M}{\partial x} + \frac{\partial N}{\partial x} + \frac{\partial P}{\partial x}
\end{aligned}$$

- If $\mathbf{F} = \mathbf{i}M(x,y) + \mathbf{j}N(x,y)$ is a field and $\mathbf{G} = \mathbf{i}N(x,y) - \mathbf{j}M(x,y)$ is the orthogonal field, then an alternate vector formulation of Green's theorem is

$$\int_C \mathbf{G} \cdot \mathbf{n}\,\mathrm{d}s = \iint_R \nabla \cdot \mathbf{G}\,\mathrm{d}x\,\mathrm{d}y$$

  - "In words, [this] says that the line integral of the normal compoment of any vector field $\mathbf{G}$ around the boundary of a region R in which $\mathbf{G}$ is continuous and has continuous partial derivatives is equal to the double integral of the divergence of $\mathbf{G}$ over R" (Thomas, 1972, p. 604).

References

Olmsted, J. M. H. (1956). *Intermediate analysis: An introduction to the theory of functions of one real variable* [Retrieved from https://hdl.handle.net/2027/mdp.39015000977804.].

Sears, F. W. (1949). *Optics* [Retrieved from https://hdl.handle.net/2027/wu.89074764515.].

Thomas, G. B., Jr. (1972). *Calculus and analytic geometry* (fourth). Addison-Wesley.