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

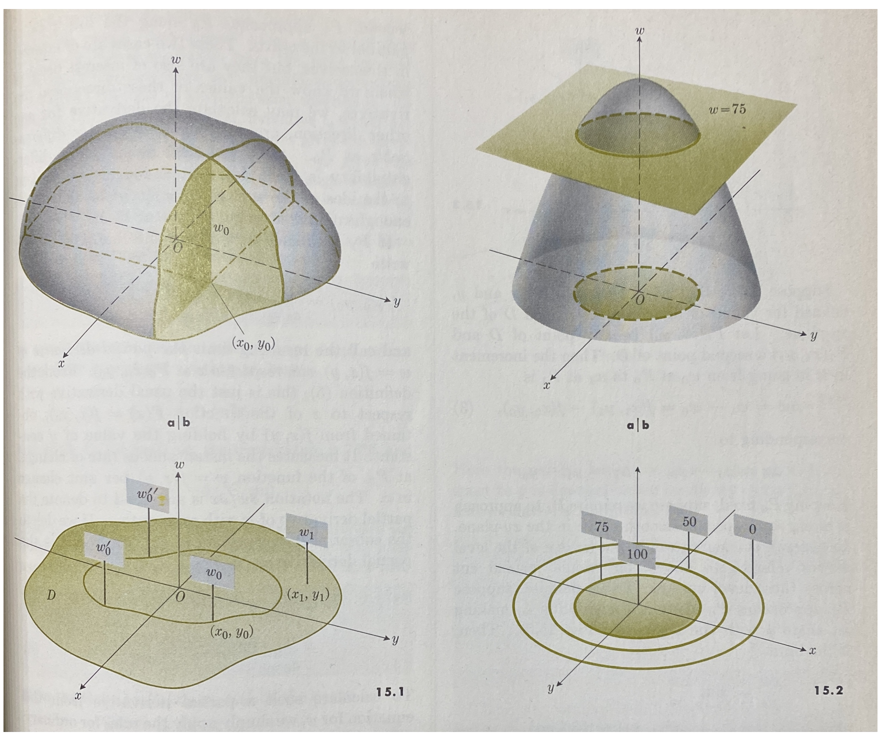


Figure 15.1: Surface plots and contour maps of 2D functions.

- 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

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

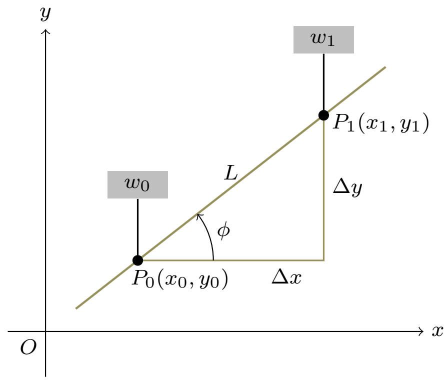


Figure 15.2: The directional derivative.

  - 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.

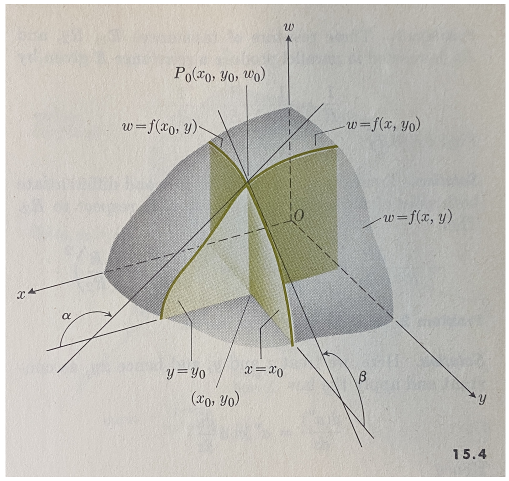


Figure 15.3: Geometric interpretation of the partial derivative.

- 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.

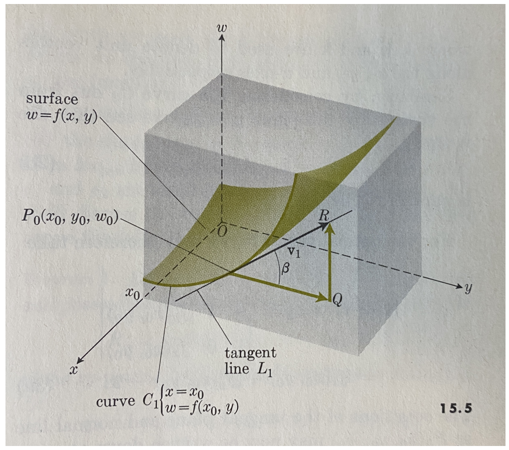


Figure 15.4: Deriving formulae for the tangent plane and normal line.

- 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 $w = f(x_1,\dots,x_n)$ parameterized by $x_1 = x_1(y_1,\dots,y_m)$, …, $x_n = (y_1,\dots,y_m)$):
$$\begin{bmatrix} \frac{\partial w}{\partial y_1} & \frac{\partial w}{\partial y_2} & \cdots & \frac{\partial w}{\partial y_m} \end{bmatrix} = \begin{bmatrix} \frac{\partial w}{\partial x_1} & \frac{\partial w}{\partial x_2} & \cdots & \frac{\partial w}{\partial x_n} \end{bmatrix} \begin{bmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \cdots & \frac{\partial x_1}{\partial y_m} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} & \cdots & \frac{\partial x_2}{\partial y_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \frac{\partial x_n}{\partial y_2} & \cdots & \frac{\partial x_n}{\partial y_m} \end{bmatrix}$$
- 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.

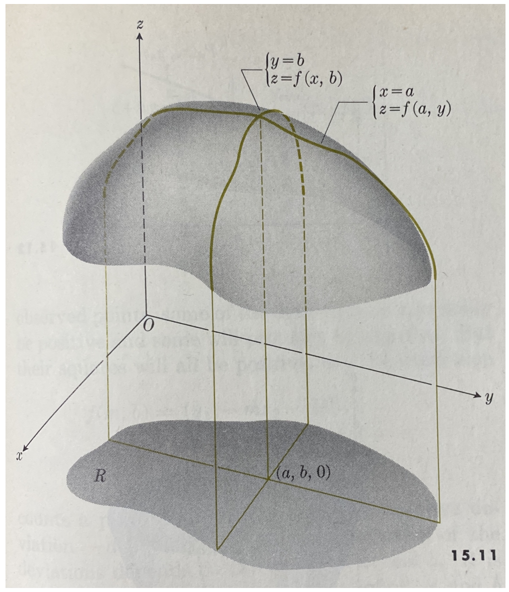


Figure 15.5: Relative maximum of $f(x,y)$.

- 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

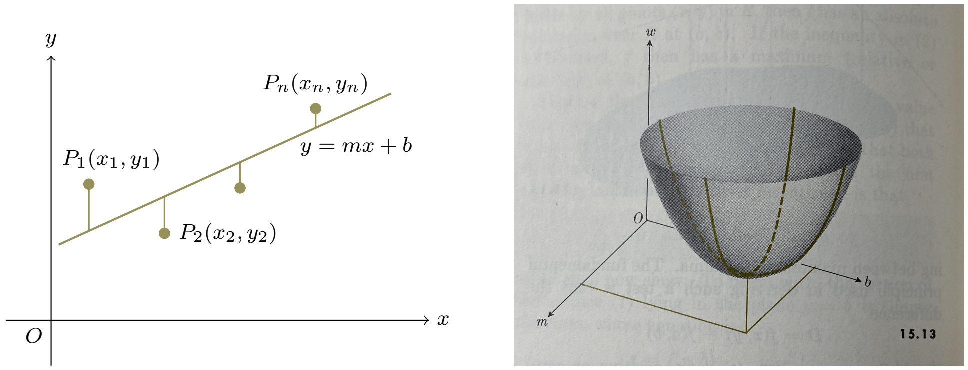
$$\begin{aligned} 0 &= \frac{\partial z}{\partial x} \\ &= 2x - y + 2 \end{aligned} \qquad\qquad \begin{aligned} 0 &= \frac{\partial z}{\partial y} \\ &= 2y - x + 2 \end{aligned}$$

  - 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).

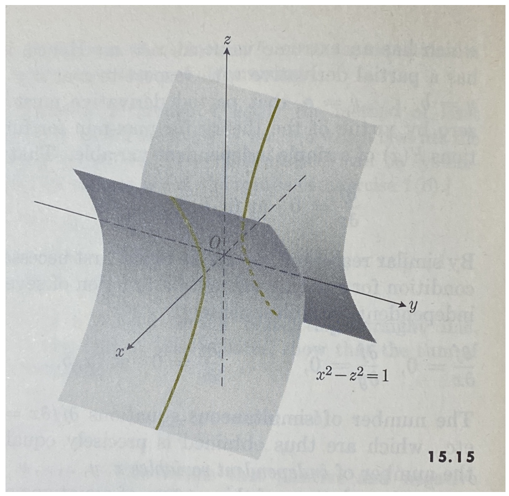


Figure 15.7: Minimizing the distance from the origin across a hyperbolic cylinder.

  - 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.

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

– 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 multipliers, 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.

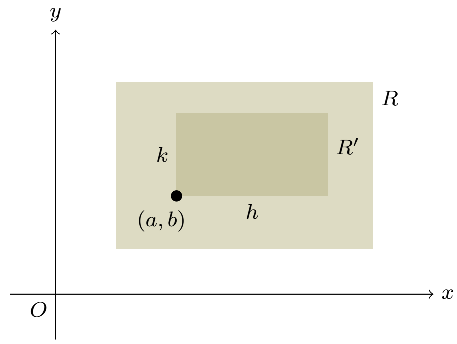


Figure 15.8: Equivalence of higher-order partial derivative ordering.

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

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

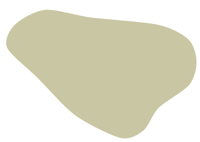
- "In fact, if all the partial derivatives that appear are continuous, the notation

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

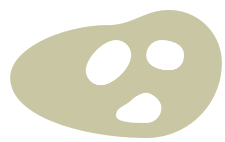
may be used to denote the result of differentiating the function $f(x, y)$ m times with respect to x and n times with respect to y, the order in which these differentiations are performed being entirely arbitrary" (Thomas, 1972, p. 535).

15.13 Exact Differentials

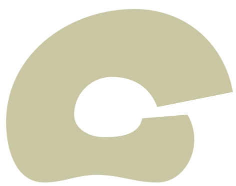
- **Exact differential**: The expression $\mathrm{d}f(x, y) = M(x, y)\,\mathrm{d}x + N(x, y)\,\mathrm{d}y$ corresponding to a function $w = f(x, y)$.



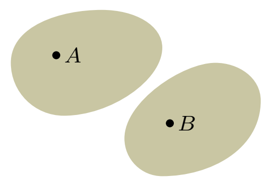
(a) A simply connected region.



(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.

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

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 $f(x, y)$:

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

$\square$

- 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
  $$f(x,y) = \int_x (x^2+y^2)\,\mathrm{d}x + g(y)$$
  $$= \frac{x^3}{3} + xy^2 + g(y)$$
  The above equation is important, and we'll end up substituting $g(y)$ into it once we find it as follows.
  $$\frac{\partial f}{\partial y} = \frac{\partial}{\partial y}\left(\frac{x^3}{3} + xy^2\right) + \frac{\partial g}{\partial y}$$
  $$2xy = 0 + 2xy + \frac{\mathrm{d}g}{\mathrm{d}y}$$
  $$\frac{\mathrm{d}g}{\mathrm{d}y} = 0$$
  $$g(y) = C \qquad C \in \mathbb{R}$$
  - 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. $\square$