Eduardo José Machado Correia Brioso Dias (A64478)
Análise Matemática II

2.1 Functions, Graphs, and Level Surfaces

Key Points in this Section.

1. A ***mapping*** or ***function*** $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ sends each point $\mathbf{x} \in A$ (the domain of f) to a specific point $f(\mathbf{x}) \in \mathbb{R}^m$. If $m = 1$, we call f a ***real valued function***.

DEFINITION: Graph of a Function Let $f: U \subset \mathbb{R}^n \to \mathbb{R}$. Define the ***graph*** of f to be the subset of $\mathbb{R}^{n+1}$ consisting of all the points

$$(x_1, \ldots, x_n, f(x_1, \ldots, x_n))$$

in $\mathbb{R}^{n+1}$ for $(x_1, \ldots, x_n)$ in U. In symbols,

$$\text{graph } f = \{(x_1, \ldots, x_n, f(x_1, \ldots, x_n)) \in \mathbb{R}^{n+1} \mid (x_1, \ldots, x_n) \in U\}.$$

DEFINITION: Level Curves and Surfaces Let $f: U \subset \mathbb{R}^n \to \mathbb{R}$ and let $c \in \mathbb{R}$. Then the ***level set of value*** c is defined to be the set of those points $\mathbf{x} \in U$ at which $f(\mathbf{x}) = c$. If $n = 2$, we speak of a ***level curve*** (of value c); and if $n = 3$, we speak of a ***level surface***. In symbols, the level set of value c is written

$$\{\mathbf{x} \in U \mid f(\mathbf{x}) = c\} \subset \mathbb{R}^n.$$

Note that the level set is always in the domain space.

4. A ***section*** of a graph is obtained by intersecting the graph with a vertical plane. For instance, for $z = f(x, y)$, setting $y = 0$ produces the section $z = f(x, 0)$ which is the graph of one function of one variable.

5. Level sets and sections are useful tools in constructing and visualizing graphs.

2.3 Differentiation

Key Points in this Section.

DEFINITION: Partial Derivatives Let $U \subset \mathbb{R}^n$ be an open set and suppose $f\colon U \subset \mathbb{R}^n \to \mathbb{R}$ is a real-valued function. Then $\partial f/\partial x_1, \ldots, \partial f/\partial x_n$, the ***partial derivatives*** of f with respect to the first, second, ... , nth variable, are the real-valued functions of n variables, which, at the point $(x_1, \ldots, x_n) = \mathbf{x}$, are defined by

$$\begin{aligned}\frac{\partial f}{\partial x_j}(x_1, \ldots, x_n) &= \lim_{h \to 0} \frac{f(x_1, x_2, \ldots, x_j + h, \ldots, x_n) - f(x_1, \ldots, x_n)}{h} \\ &= \lim_{h \to 0} \frac{f(\mathbf{x} + h\mathbf{e}_j) - f(\mathbf{x})}{h}\end{aligned}$$

if the limits exist, where $1 \le j \le n$ and $\mathbf{e}_j$ is the jth standard basis vector defined by $\mathbf{e}_j = (0, \ldots, 1, \ldots, 0)$, with 1 in the jth slot (see Section 1.5). The domain of the function $\partial f/\partial x_j$ is the set of $\mathbf{x} \in \mathbb{R}^n$ for which the limit exists.

DEFINITION: Differentiable: Two Variables Let $f\colon \mathbb{R}^2 \to \mathbb{R}$. We say f is ***differentiable*** at (x_0, y_0), if $\partial f/\partial x$ and $\partial f/\partial y$ exist at (x_0, y_0) and if

$$\frac{f(x, y) - f(x_0, y_0) - \left[\dfrac{\partial f}{\partial x}(x_0, y_0)\right](x - x_0) - \left[\dfrac{\partial f}{\partial y}(x_0, y_0)\right](y - y_0)}{\|(x, y) - (x_0, y_0)\|} \to 0 \tag{2}$$

as $(x, y) \to (x_0, y_0)$. This equation expresses what we mean by saying that

$$f(x_0, y_0) + \left[\frac{\partial f}{\partial x}(x_0, y_0)\right](x - x_0) + \left[\frac{\partial f}{\partial y}(x_0, y_0)\right](y - y_0)$$

is a ***good approximation*** to the function f.

DEFINITION: Tangent Plane Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be differentiable at $\mathbf{x}_0 = (x_0, y_0)$. The plane in $\mathbb{R}^3$ defined by the equation

$$z = f(x_0, y_0) + \left[\frac{\partial f}{\partial x}(x_0, y_0)\right](x - x_0) + \left[\frac{\partial f}{\partial y}(x_0, y_0)\right](y - y_0),$$

is called the ***tangent plane*** of the graph of f at the point (x_0, y_0).

DEFINITION: Differentiable, n Variables, m Functions Let U be an open set in $\mathbb{R}^n$ and let $f\colon U \subset \mathbb{R}^n \to \mathbb{R}^m$ be a given function. We say that f is ***differentiable*** at $\mathbf{x}_0 \in U$ if the partial derivatives of f exist at $\mathbf{x}_0$ and if

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{\|f(\mathbf{x}) - f(\mathbf{x}_0) - \mathbf{T}(\mathbf{x}-\mathbf{x}_0)\|}{\|\mathbf{x}-\mathbf{x}_0\|} = 0, \tag{4}$$

where $\mathbf{T} = \mathbf{D}f(\mathbf{x}_0)$ is the $m \times n$ matrix with matrix elements $\partial f_i/\partial x_j$ evaluated at $\mathbf{x}_0$ and $\mathbf{T}(\mathbf{x}-\mathbf{x}_0)$ means the product of $\mathbf{T}$ with $\mathbf{x}-\mathbf{x}_0$ (regarded as a column matrix). We call $\mathbf{T}$ the ***derivative*** of f at $\mathbf{x}_0$.

DEFINITION: Gradient Consider the special case $f\colon U \subset \mathbb{R}^n \to \mathbb{R}$. Here $\mathbf{D}f(\mathbf{x})$ is a $1 \times n$ matrix:

$$\mathbf{D}f(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} & \cdots & \dfrac{\partial f}{\partial x_n} \end{bmatrix}.$$

We can form the corresponding vector $(\partial f/\partial x_1, \ldots, \partial f/\partial x_n)$, called the ***gradient*** of f and denoted by ∇f or grad f.

THEOREM 8 Let $f\colon U \subset \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at $\mathbf{x}_0 \in U$. Then f is continuous at $\mathbf{x}_0$.

THEOREM 9 Let $f\colon U \subset \mathbb{R}^n \to \mathbb{R}^m$. Suppose the partial derivatives $\partial f_i/\partial x_j$ of f all exist and are continuous in a neighborhood of a point $\mathbf{x} \in U$. Then f is differentiable at $\mathbf{x}$.

2. The ***linear approximation*** to $f(x, y)$ at (x_0, y_0) is

$$\ell_{(x_0,y_0)}(x,y) = f(x_0,y_0) + \left[\frac{\partial f}{\partial x}(x_0,y_0)\right](x-x_0) + \left[\frac{\partial f}{\partial y}(x_0,y_0)\right](y-y_0)$$

6. If $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ has partial derivatives at $\mathbf{x}_0 \in U$, the ***derivative matrix*** is the $m \times n$ matrix $\mathbf{D}f(\mathbf{x}_0)$ given by

$$\mathbf{D}f(\mathbf{x}_0) = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \cdots & \dfrac{\partial f_2}{\partial x_n} \\ \vdots & & & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \dfrac{\partial f_m}{\partial x_2} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{bmatrix}$$

where the partials are all evaluated at $\mathbf{x}_0$.

9. If f is differentiable at $\mathbf{x}_0$, then it is continuous at $\mathbf{x}_0$. If the partials exist and are continuous in a neighborhood of $\mathbf{x}_0$ (that is, f is C^1), then f is differentiable at $\mathbf{x}_0$.

2.4 Introduction to Paths

Key Points in this Section.

2. A particle on the rim of a rolling circle of radius 1 traces out a path called a ***cycloid***:

$$\mathbf{c}(t) = (t - \sin t, 1 - \cos t).$$

Paths and Curves A ***path*** in $\mathbb{R}^n$ is a map $\mathbf{c}\colon [a, b] \to \mathbb{R}^n$; it is a ***path in the plane*** if $n = 2$ and a ***path in space*** if $n = 3$. The collection C of points $\mathbf{c}(t)$ as t varies in $[a, b]$ is called a ***curve***, and $\mathbf{c}(a)$ and $\mathbf{c}(b)$ are its ***endpoints***. The path $\mathbf{c}$ is said to ***parametrize*** the curve C. We also say $\mathbf{c}(t)$ ***traces out*** C as t varies.

If $\mathbf{c}$ is a path in $\mathbb{R}^3$, we can write $\mathbf{c}(t) = (x(t), y(t), z(t))$ and we call $x(t)$, $y(t)$, and $z(t)$ the ***component functions*** of $\mathbf{c}$. We form component functions similarly in $\mathbb{R}^2$ or, generally, in $\mathbb{R}^n$.

DEFINITION: Velocity Vector If $\mathbf{c}$ is a path and it is differentiable, we say $\mathbf{c}$ is a ***differentiable path***. The ***velocity*** of $\mathbf{c}$ at time t is defined by[3]

$$\mathbf{c}'(t) = \lim_{h\to 0} \frac{\mathbf{c}(t+h) - \mathbf{c}(t)}{h}.$$

We normally draw the vector $\mathbf{c}'(t)$ with its tail at the point $\mathbf{c}(t)$. The ***speed*** of the path $\mathbf{c}(t)$ is $s = \|\mathbf{c}'(t)\|$, the length of the velocity vector. If $\mathbf{c}(t) = (x(t), y(t))$ in $\mathbb{R}^2$, then

$$\mathbf{c}'(t) = (x'(t), y'(t)) = x'(t)\mathbf{i} + y'(t)\mathbf{j}$$

and if $\mathbf{c}(t) = (x(t), y(t), z(t))$ in $\mathbb{R}^3$, then

$$\mathbf{c}'(t) = (x'(t), y'(t), z'(t)) = x'(t)\mathbf{i} + y'(t)\mathbf{j} + z'(t)\mathbf{k}.$$

Tangent Vector The velocity $\mathbf{c}'(t)$ is a vector ***tangent*** to the path $\mathbf{c}(t)$ at time t. If C is a curve traced out by $\mathbf{c}$ and if $\mathbf{c}'(t)$ is not equal to $\mathbf{0}$, then $\mathbf{c}'(t)$ is a vector tangent to the curve C at the point $\mathbf{c}(t)$.

Tangent Line to a Path If $\mathbf{c}(t)$ is a path, and if $\mathbf{c}'(t_0) \neq \mathbf{0}$, the equation of its ***tangent line*** at the point $\mathbf{c}(t_0)$ is

$$\boldsymbol{l}(t) = \mathbf{c}(t_0) + (t - t_0)\mathbf{c}'(t_0).$$

If C is the curve traced out by $\mathbf{c}$, then the line traced out by $\boldsymbol{l}$ is the tangent line to the curve C at $\mathbf{c}(t_0)$.

2.5 Properties of the Derivative

Key Points in this Section.

THEOREM 10: Sums, Products, Quotients

(i) **Constant Multiple Rule**. Let $f\colon U \subset \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at $\mathbf{x}_0$ and let c be a real number. Then $h(\mathbf{x}) = cf(\mathbf{x})$ is differentiable at $\mathbf{x}_0$ and

$$\mathbf{D}h(\mathbf{x}_0) = c\mathbf{D}f(\mathbf{x}_0) \qquad \text{(equality of matrices)}.$$

(ii) **Sum Rule**. Let $f\colon U \subset \mathbb{R}^n \to \mathbb{R}^m$ and $g\colon U \subset \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at $\mathbf{x}_0$. Then $h(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$ is differentiable at $\mathbf{x}_0$ and

$$\mathbf{D}h(\mathbf{x}_0) = \mathbf{D}f(\mathbf{x}_0) + \mathbf{D}g(\mathbf{x}_0) \qquad \text{(sum of matrices)}.$$

(iii) **Product Rule**. Let $f\colon U \subset \mathbb{R}^n \to \mathbb{R}$ and $g\colon U \subset \mathbb{R}^n \to \mathbb{R}$ be differentiable at $\mathbf{x}_0$ and let $h(\mathbf{x}) = g(\mathbf{x})f(\mathbf{x})$. Then $h\colon U \subset \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\mathbf{x}_0$ and

$$\mathbf{D}h(\mathbf{x}_0) = g(\mathbf{x}_0)\mathbf{D}f(\mathbf{x}_0) + f(\mathbf{x}_0)\mathbf{D}g(\mathbf{x}_0).$$

(Note that each side of this equation is a $1 \times n$ matrix; a more general product rule is presented in Exercise 29 at the end of this section.)

(iv) **Quotient Rule**. With the same hypotheses as in rule (iii), let $h(\mathbf{x}) = f(\mathbf{x})/g(\mathbf{x})$ and suppose g is never zero on U. Then h is differentiable at $\mathbf{x}_0$ and

$$\mathbf{D}h(\mathbf{x}_0) = \frac{g(\mathbf{x}_0)\mathbf{D}f(\mathbf{x}_0) - f(\mathbf{x}_0)\mathbf{D}g(\mathbf{x}_0)}{[g(\mathbf{x}_0)]^2}.$$

THEOREM 11: Chain Rule Let $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ be open sets. Let $g\colon U \subset \mathbb{R}^n \to \mathbb{R}^m$ and $f\colon V \subset \mathbb{R}^m \to \mathbb{R}^p$ be given functions such that g maps U into V, so that $f \circ g$ is defined. Suppose g is differentiable at $\mathbf{x}_0$ and f is differentiable at $\mathbf{y}_0 = g(\mathbf{x}_0)$. Then $f \circ g$ is differentiable at $\mathbf{x}_0$ and

$$\mathbf{D}(f \circ g)(\mathbf{x}_0) = \mathbf{D}f(\mathbf{y}_0)\mathbf{D}g(\mathbf{x}_0). \tag{1}$$

The right-hand side is the matrix product of $\mathbf{D}f(\mathbf{y}_0)$ with $\mathbf{D}g(\mathbf{x}_0)$.

2.6 Gradients and Directional Derivatives

Key Points in this Section.

3. The direction in which f is *increasing the fastest* at $\mathbf{x}$ is the direction parallel to $\nabla f(\mathbf{x})$. The direction of fastest *decrease* is parallel to $-\nabla f(\mathbf{x})$.

5. The gravitational force field

$$\mathbf{F} = -\frac{GMm}{r^3}\mathbf{r} = -\frac{GMm}{r^2}\mathbf{n}$$

(the inverse square law), where $\mathbf{n} = \mathbf{r}/r$, $\mathbf{r} = x\mathbf{i}+y\mathbf{j}+z\mathbf{k}$ and $r = \|\mathbf{r}\|$, is a gradient. Namely,

$$\mathbf{F} = -\nabla V,$$

where

$$V = -\frac{GMm}{r}.$$

DEFINITION: The Gradient If $f\colon U \subset \mathbb{R}^3 \to \mathbb{R}$ is differentiable, the ***gradient*** of f at (x, y, z) is the vector in space given by

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right).$$

This vector is also denoted $\nabla f(x, y, z)$. Thus, ∇f is just the matrix of the derivative $\mathbf{D}f$, written as a vector.

DEFINITION: Directional Derivatives If $f\colon \mathbb{R}^3 \to \mathbb{R}$, the ***directional derivative*** of f at $\mathbf{x}$ along the vector $\mathbf{v}$ is given by

$$\left.\frac{d}{dt} f(\mathbf{x} + t\mathbf{v})\right|_{t=0}$$

if this exists.

In the definition of a directional derivative, we normally choose $\mathbf{v}$ to be a *unit* vector. In this case we are moving in the direction $\mathbf{v}$ with unit speed and we refer to $\nabla f(\mathbf{x}) \cdot \mathbf{v}$ as the ***directional derivative of f in the direction $\mathbf{v}$***.

THEOREM 12 If $f\colon \mathbb{R}^3 \to \mathbb{R}$ is differentiable, then all directional derivatives exist. The directional derivative at $\mathbf{x}$ in the direction $\mathbf{v}$ is given by

$$\mathbf{D}f(\mathbf{x})\mathbf{v} = \operatorname{grad} f(\mathbf{x}) \cdot \mathbf{v} = \nabla f(\mathbf{x}) \cdot \mathbf{v} = \left[\frac{\partial f}{\partial x}(\mathbf{x})\right]v_1 + \left[\frac{\partial f}{\partial y}(\mathbf{x})\right]v_2 + \left[\frac{\partial f}{\partial z}(\mathbf{x})\right]v_3,$$

where $\mathbf{v} = (v_1, v_2, v_3)$.

THEOREM 13 Assume $\nabla f(\mathbf{x}) \neq \mathbf{0}$. Then $\nabla f(\mathbf{x})$ points in the direction along which f is increasing the fastest.

THEOREM 14: The Gradient is Normal to Level Surfaces Let $f: \mathbb{R}^3 \to \mathbb{R}$ be a C^1 map and let (x_0, y_0, z_0) lie on the level surface S defined by $f(x, y, z) = k$, for k a constant. Then $\nabla f(x_0, y_0, z_0)$ is normal to the level surface in the following sense: If $\mathbf{v}$ is the tangent vector at $t = 0$ of a path $\mathbf{c}(t)$ in S with $\mathbf{c}(0) = (x_0, y_0, z_0)$, then $\nabla f(x_0, y_0, z_0) \cdot \mathbf{v} = 0$ (see Figure 2.6.2).

DEFINITION: Tangent Planes to Level Surfaces Let S be the surface consisting of those (x, y, z) such that $f(x, y, z) = k$, for k a constant. The ***tangent plane*** of S at a point (x_0, y_0, z_0) of S is defined by the equation

$$\nabla f(x_0, y_0, z_0) \cdot (x - x_0, y - y_0, z - z_0) = 0 \tag{1}$$

if $\nabla f(x_0, y_0, z_0) \neq \mathbf{0}$. That is, the tangent plane is the set of points (x, y, z) that satisfy equation (1).

3.1 Iterated Partial Derivatives

Key Points in this Section.

1. **Equality of Mixed Partials.** If $f(x, y)$ is C^2 (has continuous 2nd partial derivatives), then

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}.$$

3. Higher order partials are also symmetric; for example, for $f(x, y, z)$,

$$\frac{\partial^4 f}{\partial x \partial^2 z \partial y} = \frac{\partial^4 f}{\partial x \partial y \partial^2 z}$$

3.3 Extrema of Real Valued Functions

Key Points in this Section.

DEFINITION If $f\colon U \subset \mathbb{R}^n \to \mathbb{R}$ is a given scalar function, a point $\mathbf{x}_0 \in U$ is called a ***local minimum*** of f if there is a neighborhood V of $\mathbf{x}_0$ such that for all points $\mathbf{x}$ in V, $f(\mathbf{x}) \geq f(\mathbf{x}_0)$. (See Figure 3.3.3.) Similarly, $\mathbf{x}_0 \in U$ is a ***local maximum*** if there is a neighborhood V of $\mathbf{x}_0$ such that $f(\mathbf{x}) \leq f(\mathbf{x}_0)$ for all $\mathbf{x} \in V$. The point $\boldsymbol{x_0} \in U$ is said to be a ***local*** or ***relative***, ***extremum*** if i tis either a local minimum or a local maximum. A point x_0 is a ***critical point*** of f if either f is not *differenciable* at $\boldsymbol{x_0}$,or if it is, $\mathbf{D}\boldsymbol{f}(\boldsymbol{x_0}) = \mathbf{0}$. A critical point that is not a local extremum is called a ***saddle point***.

2. **First Derivative Test.** If $U \subset \mathbb{R}^n$ is open, $f : U \subset \mathbb{R}^n \to \mathbb{R}$ is differentiable and $\mathbf{x}_0$ is a local extremum, then $\mathbf{x}_0$ is a ***critical point***; that is, all the partials of f vanish at $\mathbf{x}_0$:

$$\frac{\partial f}{\partial x_1}(\mathbf{x}_0) = 0, \cdots, \frac{\partial f}{\partial x_0}(\mathbf{x}_0) = 0.$$

The idea of the proof is to apply the one-variable first derivative test to f restricted to lines through $\mathbf{x}_0$.

DEFINITION Suppose that $f\colon U \subset \mathbb{R}^n \to \mathbb{R}$ has second-order continuous derivatives $(\partial^2 f/\partial x_i\, \partial x_j)(\mathbf{x}_0)$, for $i, j = 1, \ldots, n$, at a point $\mathbf{x}_0 \in U$. The ***Hessian of f at*** $\mathbf{x}_0$ is the quadratic function defined by

$$Hf(\mathbf{x}_0)(\mathbf{h}) = \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i\, \partial x_j}(\mathbf{x}_0) h_i h_j$$

$$= \frac{1}{2}(h_1, \ldots, h_n)\begin{pmatrix} \frac{\partial^2 f}{\partial x_1\, \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_1\, \partial x_n} \\ \vdots & & \\ \frac{\partial^2 f}{\partial x_n\, \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n\, \partial x_n} \end{pmatrix}\begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}.$$

4. **Second Derivative Test—n Variables.** If $f : U \subset \mathbb{R}^n \to \mathbb{R}$ is C^3 (and again U is open), $\mathbf{x}_0$ is a critical point, and if $Hf(\mathbf{x}_0)(\mathbf{h}) > 0$ for all $\mathbf{h} \neq \mathbf{0}$ (that is, $Hf(\mathbf{x}_0)$ is ***positive definite***), then $\mathbf{x}_0$ is a local minimum. Likewise, if $Hf(\mathbf{x}_0)(\mathbf{h}) < 0$ for all $\mathbf{h} \neq \mathbf{0}$, (that is, $Hf(\mathbf{x}_0)$ is ***negative definite***), then $\mathbf{x}_0$ is a local maximum.

5. The idea of the proof of the second derivative test is to apply the second order Taylor theorem and show that the remainder term can be ignored.

6. **Second Derivative Test—Two Variables.** Let $f : U \subset \mathbb{R}^2 \to \mathbb{R}$ (again with U open) be of class C^3. A point $(x_0, y_0) \in U$ is a local minimum if the following conditions are satisfied:

   (i) $\dfrac{\partial f}{\partial x}(x_0, y_0) = \dfrac{\partial f}{\partial y}(x_0, y_0) = 0$ (that is, (x_0, y_0) is a critical point)

   (ii) $\dfrac{\partial^2 f}{\partial x^2}(x_0, y_0) > 0$

   (iii) $D = \begin{vmatrix} \dfrac{\partial^2 f}{\partial x^2}(x_0, y_0) & \dfrac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \\ \dfrac{\partial^2 f}{\partial x \partial y}(x_0, y_0) & \dfrac{\partial^2 f}{\partial y^2}(x_0, y_0) \end{vmatrix} > 0.$

   If (i) and (iii) hold, but $\partial^2 f/\partial x^2$ at (x_0, y_0) is negative, then (x_0, y_0) is a local maximum. If the ***discriminant*** D is negative, then (x_0, y_0) is a ***saddle point*** (that is, (x_0, y_0) is neither a local maximum nor a local minimum).

DEFINITION Suppose $f\colon A \to \mathbb{R}$ is a function defined on a set A in $\mathbb{R}^2$ or $\mathbb{R}^3$. A point $\mathbf{x}_0 \in A$ is said to be an ***absolute maximum*** (or ***absolute minimum***) point of f if $f(\mathbf{x}) \le f(\mathbf{x}_0)$ [or $f(\mathbf{x}) \ge f(\mathbf{x}_0)$] for all $\mathbf{x} \in A$.

DEFINITION A set $D \in \mathbb{R}^n$ is said to be ***bounded*** if there is a number $M > 0$ such that $\|\mathbf{x}\| < M$ for all $\mathbf{x} \in D$. A set is ***closed*** if it contains all its boundary points.

THEOREM 7: Global Existence Theorem for Maxima and Minima Let D be closed and bounded in $\mathbb{R}^n$ and let $f\colon D \to \mathbb{R}$ be continuous. Then f assumes its absolute maximum and minimum values at some points $\mathbf{x}_0$ and $\mathbf{x}_1$ of D.

Strategy for Finding the Absolute Maxima and Minima on a Region with Boundary Let f be a continuous function of two variables defined on a closed and bounded region D in $\mathbb{R}^2$, which is bounded by a smooth closed curve. To find the absolute maximum and minimum of f on D:

(i) Locate all critical points for f in U.

(ii) Find the maximum and minimum of f viewed as a function only on ∂U.

(iii) Compute the value of f at all of these critical points.

(iv) Compare all these values and select the largest and the smallest.

3.4 Constrained Extrema and Lagrange Multipliers

THEOREM 8: The Method of Lagrange Multipliers Suppose that $f\colon U \subset \mathbb{R}^n \to \mathbb{R}$ and $g\colon U \subset \mathbb{R}^n \to \mathbb{R}$ are given C^1 real-valued functions. Let $\mathbf{x}_0 \in U$ and $g(\mathbf{x}_0) = c$, and let S be the level set for g with value c (recall that this is the set of points $\mathbf{x} \in \mathbb{R}^n$ satisfying $g(\mathbf{x}) = c$). Assume $\nabla g(\mathbf{x}_0) \neq \mathbf{0}$.

If $f|S$, which denotes "f restricted to S," has a local maximum or minimum on S at $\mathbf{x}_0$, then there is a real number λ such that

$$\nabla f(\mathbf{x}_0) = \lambda \nabla g(\mathbf{x}_0). \tag{1}$$

2. The idea of the proof is to use the fact that f has a critical point along any curve in the level set through $\mathbf{x}_0$, which shows, via the chain rule, that $\nabla f(\mathbf{x}_0)$ is perpendicular to that level set; but $\nabla g(\mathbf{x}_0)$ is also perpendicular, so these two vectors are parallel.

3. The Lagrange multiplier method produces *candidates* for extrema; one must make sure there is an extremum and then f can be evaluated at the candidates to choose the maximum or minimum as desired.

4. If there are k constraints

$$g_1 = c_1, \cdots, g_k = c_k,$$

for C^1 functions $g(x_1, \ldots, x_n), \ldots g_k(x_1, \ldots, x_n)$ and constants $c_1, \ldots, c_k$, then the Lagrange multiplier equations become

$$\nabla f(\mathbf{x}_0) = \lambda_1 \nabla g(\mathbf{x}_0) + \cdots + \lambda_k \nabla g(\mathbf{x}_0).$$

5. The Lagrange multiplier method is an effective tool for finding the extrema of $f|\partial U$ in the strategy for finding global extrema described in the last section.

THEOREM 9 If f, when constrained to a surface S, has a maximum or minimum at $\mathbf{x}_0$, then $\nabla f(\mathbf{x}_0)$ is perpendicular to S at $\mathbf{x}_0$ (see Figure 3.4.2).

THEOREM 10 Let $f\colon U \subset \mathbb{R}^2 \to \mathbb{R}$ and $g\colon U \subset \mathbb{R}^2 \to \mathbb{R}$ be smooth (at least C^2) functions. Let $\mathbf{v}_0 \in U$, $g(\mathbf{v}_0) = c$, and S be the level curve for g with value c. Assume that $\nabla g(\mathbf{v}_0) \neq \mathbf{0}$ and that there is a real number λ such that $\nabla f(\mathbf{v}_0) = \lambda \nabla g(\mathbf{v}_0)$. Form the auxiliary function $h = f - \lambda g$ and the ***bordered Hessian*** determinant

$$|\bar{H}| = \begin{vmatrix} 0 & -\dfrac{\partial g}{\partial x} & -\dfrac{\partial g}{\partial y} \\ -\dfrac{\partial g}{\partial x} & \dfrac{\partial^2 h}{\partial x^2} & \dfrac{\partial^2 h}{\partial x\,\partial y} \\ -\dfrac{\partial g}{\partial y} & \dfrac{\partial^2 h}{\partial x\,\partial y} & \dfrac{\partial^2 h}{\partial y^2} \end{vmatrix} \quad \text{evaluated at } \mathbf{v}_0.$$

(i) If $|\bar{H}| > 0$, then $\mathbf{v}_0$ is a local maximum point for $f|S$.

(ii) If $|\bar{H}| < 0$, then $\mathbf{v}_0$ is a local minimum point for $f|S$.

(iii) If $|\bar{H}| = 0$, the test is inconclusive and $\mathbf{v}_0$ may be a minimum, a maximum, or neither.