Appendix A Brief Review of Continuity and Differentiability

R^n will denote the set of n-tuples $(x_1, \ldots, x_n)$ of real numbers. Although we use only the cases $R^1 = R$, R^2, and R^3, the more general notion of R^n unifies the definitions and brings in no additional difficulties; the reader may think in R^2 or R^3, if he wishes so. In these particular cases, we shall use the following more traditional notation: x or t for R, (x, y) or (u, v) for R^2, and (x, y, z) for R^3.

A. Continuity in R^n

We start by making precise the notion of a point being ϵ-close to a given point $p_0 \in R^n$:

A *ball* (or *open ball*) in R^n with center $p_0 = (x_1^0, \ldots, x_n^0)$ and radius $\epsilon > 0$ is the set

$$B_\epsilon(p_0) = \{(x_1, \ldots, x_n) \in R^n; (x_1 - x_1^0)^2 + \cdots + (x_n - x_n^0) < \epsilon^2\}.$$

Thus, in R, $B_\epsilon(p_0)$ is an open interval with center p_0 and length 2ϵ; in R^2, $B_\epsilon(p_0)$ is the interior of a disk with center p_0 and radius ϵ; in R^3, $B_\epsilon(p_0)$ is the interior of a region bounded by a sphere of center p_0 and radius ϵ (see Fig. A2-1).

A set $U \subset R^n$ is an *open set* if for each $p \in U$ there is a ball $B_\epsilon(p) \subset U$; intuitively this means that points in U are entirely surrounded by points of U, or that points sufficiently close to points of U still belong to U.

For instance, the set

$$\{(x, y) \in R^2; a < x < b, c < y < d\}$$

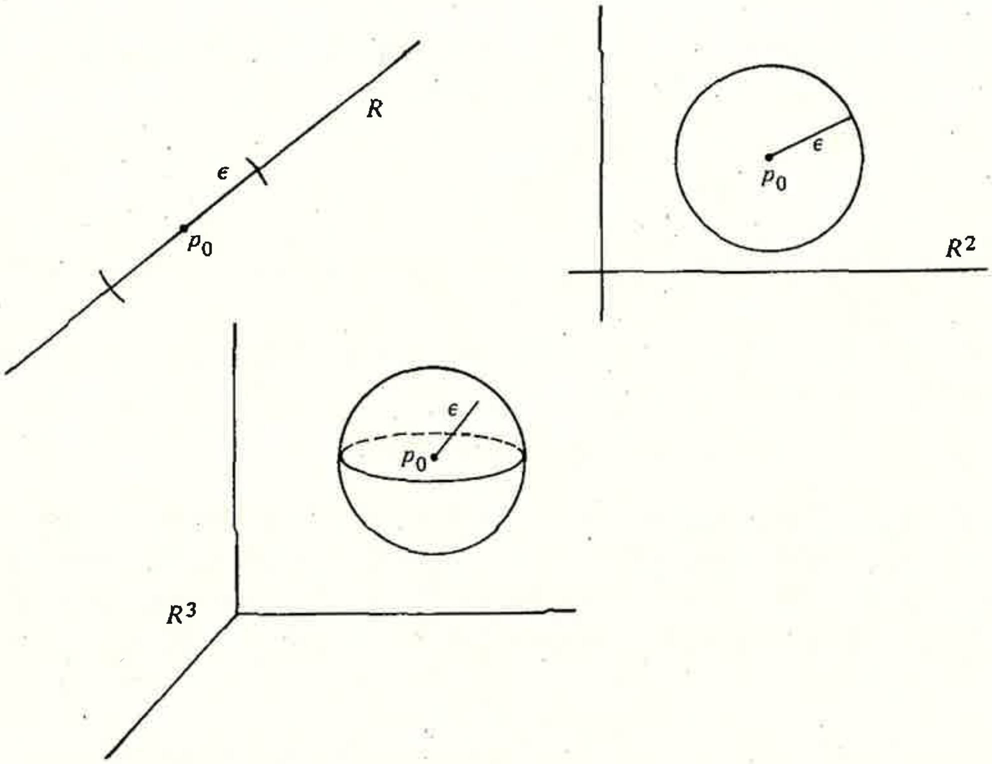

Figure A2-1

is easily seen to be open in R^2. However, if one of the strict inequalities, say $x < b$, is replaced by $x \leq b$, the set is no longer open; no ball with center at the point $(b, (d + c)/2)$, which belongs to the set, can be contained in the set (Fig. A2-2).

It is convenient to say that an open set in R^n containing a point $p \in R^n$ is a *neighborhood* of p.

From now on, $U \subset R^n$ will denote an open set in R^n.

We recall that a real function $f: U \subset R \to R$ of a real variable is continuous at $x_0 \in U$ if given an $\epsilon > 0$ there exists a $\delta > 0$ such that if $|x - x_0| < \delta$, then

$$|f(x) - f(x_0)| < \epsilon.$$

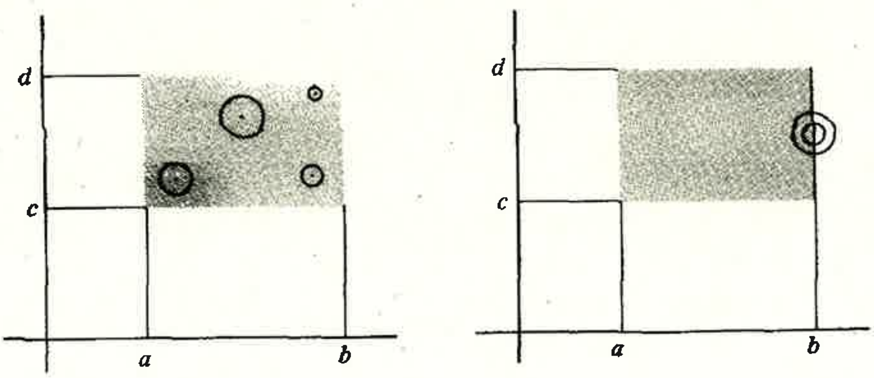

Figure A2-2

Similarly, a real function $f: U \subset R^2 \longrightarrow R$ of two real variables is continuous at $(x_0, y_0) \in U$ if given an $\epsilon > 0$ there exists a $\delta > 0$ such that if $(x - x_0)^2 + (y - y_0)^2 < \delta^2$, then

$$|f(x, y) - f(x_0, y_0)| < \epsilon.$$

The notion of ball unifies these definitions as particular cases of the following general concept:

A map $F: U \subset R^n \longrightarrow R^m$ is *continuous* at $p \in U$ if given $\epsilon > 0$, there exists a $\delta > 0$ such that

$$F(B_\delta(p)) \subset B_\epsilon(F(p)).$$

In other words, F is continuous at p if points arbitrarily close to $F(p)$ are images of points sufficiently close to p. It is easily seen that in the particular cases of $n = 1, 2$ and $m = 1$, this agrees with the previous definitions. We say that F is *continuous* in U if F is *continuous* for all $p \in U$ (Fig. A2-3).

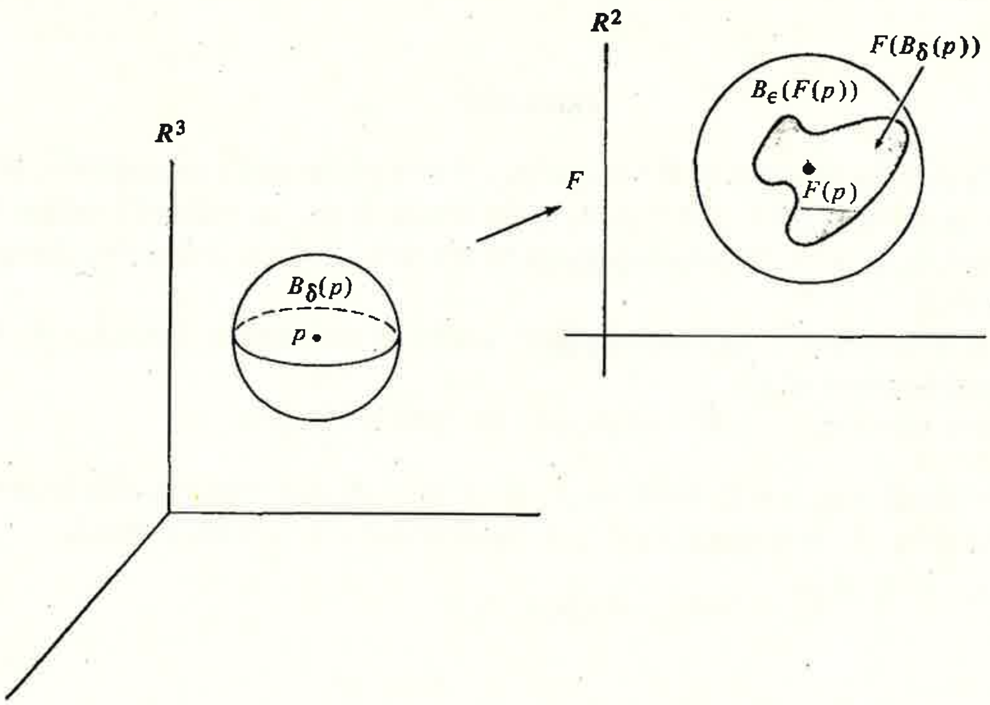

Figure A2-3

Given a map $F: U \subset R^n \longrightarrow R^m$, we can determine m functions of n variables as follows. Let $p = (x_1, \ldots, x_n) \in U$ and $f(p) = (y_1, \ldots, y_m)$. Then we can write

$$y_1 = f_1(x_1, \ldots, x_n), \ldots, y_m = f_m(x_1, \ldots, x_n).$$

The functions $f_i: U \longrightarrow R$, $i = 1, \ldots, m$, are the *component functions* of F.

Example 1 (*Symmetry*). Let $F: R^3 \longrightarrow R^3$ be the map which assigns to each $p \in R^3$ the point which is symmetric to p with respect to the origin $O \in R^3$. Then $F(p) = -p$, or

$$F(x, y, z) = (-x, -y, -z),$$

and the component functions of F are

$$f_1(x, y, z) = -x, \qquad f_2(x, y, z) = -y, \qquad f_3(x, y, z) = -z.$$

Example 2 (*Inversion*). Let $F: R^2 - \{(0, 0)\} \longrightarrow R^2$ be defined as follows. Denote by $|p|$ the distance to the origin $(0, 0) = O$ of a point $p \in R^2$. By definition, $F(p), p \neq 0$, belongs to the half-line Op and is such that $|F(p)|\cdot|p| = 1$. Thus, $F(p) = p/|p|^2$, or

$$F(x, y) = \left(\frac{x}{x^2 + y^2}, \frac{y}{x^2 + y^2}\right), \qquad (x, y) \neq (0, 0),$$

and the component functions of F are

$$f_1(x, y) = \frac{x}{x^2 + y^2}, \qquad f_2(x, y) = \frac{y}{x^2 + y^2}.$$

Example 3 (*Projection*). Let $\pi: R^3 \longrightarrow R^2$ be the projection $\pi(x, y, z) = (x, y)$. Then $f_1(x, y, z) = x, f_2(x, y, z) = y$.

The following proposition shows that the continuity of the map F is equivalent to the continuity of its component functions.

PROPOSITION 1. $\mathrm{F}: \mathrm{U} \subset \mathrm{R}^n \longrightarrow \mathrm{R}^m$ *is continuous if and only if each component function* $\mathrm{f}_i: \mathrm{U} \subset \mathrm{R}^n \longrightarrow \mathrm{R}$, $\mathrm{i} = 1, \ldots, \mathrm{m}$, *is continuous.*

Proof. Assume that F is continuous at $p \in U$. Then given $\epsilon > 0$, there exists $\delta > 0$ such that $F(B_\delta(p)) \subset B_\epsilon(F(p))$. Thus, if $q \in B_\delta(p)$, then

$$F(q) \in B_\epsilon(F(p)),$$

that is,

$$(f_1(q) - f_1(p))^2 + \cdots + (f_m(q) - f_m(p))^2 < \epsilon^2,$$

which implies that, for each $i = 1, \ldots, m$, $|f_i(q) - f_i(p)| < \epsilon$. Therefore, given $\epsilon > 0$ there exists $\delta > 0$ such that if $q \in S_\delta(p)$, then $|f_i(q) - f_i(p)| < \epsilon$. Hence, each f_i is continuous at p.

Conversely, let f_i, $i = 1, \ldots, m$, be continuous at p. Then given $\epsilon > 0$ there exists $\delta_i > 0$ such that if $q \in S_{\delta_i}(p)$, then $|f_i(q) - f_i(p)| < \epsilon/\sqrt{m}$. Set

$\delta < \min \delta_i$ and let $q \in S_\delta(p)$. Then

$$(f_1(q) - f_1(p))^2 + \cdots + (f_m(q) - f_m(p))^2 < \epsilon^2,$$

and hence, the continuity of F at p. **Q.E.D.**

It follows that the maps in Examples 1, 2, and 3 are continuous.

Example 4. Let $F: U \subset R \longrightarrow R^m$. Then

$$F(t) = (x_1(t), \ldots, x_m(t)), \qquad t \in U.$$

This is usually called a *vector-valued function*, and the component functions of F are the components of the vector $F(t) \in R^m$. When F is continuous, or, equivalently, the functions $x_i(t)$, $i = 1, \ldots, m$, are continuous, we say that F is a *continuous curve* in R^n.

In most applications, it is convenient to express the continuity in terms of neighborhoods instead of balls.

PROPOSITION 2. *A map* $F: U \subset R^n \longrightarrow R^m$ *is continuous at* $p \in U$ *if and only if, given a neighborhood* V *of* $F(p)$ *in* R^m *there exists a neighborhood* W *of* p *in* R^n *such that* $F(W) \subset V$.

Proof. Assume that F is continuous at p. Since V is an open set containing $F(p)$, it contains a ball $B_\epsilon(F(p))$ for some $\epsilon > 0$. By continuity, there exists a ball $B_\delta(p) = W$ such that

$$F(W) = F(B_\delta(p)) \subset B_\epsilon(F(p)) \subset V,$$

and this proves that the condition is necessary.

Conversely, assume that the condition holds. Let $\varepsilon > 0$ be given and set $V = B_\epsilon(F(p))$. By hypothesis, there exists a neighborhood W of p in R^n such that $F(W) \subset V$. Since W is open, there exists a ball $B_\delta(p) \subset W$. Thus,

$$F(B_\delta(p)) \subset F(W) \subset V = B_\epsilon(F(p)),$$

and hence the continuity of F at p. **Q.E.D.**

The composition of continuous maps yields a continuous map. More precisely, we have the following proposition.

PROPOSITION 3. *Let* $F: U \subset R^n \longrightarrow R^m$ *and* $G: V \subset R^m \longrightarrow R^k$ *be continuous maps, where* U *and* V *are open sets such that* $F(U) \subset V$. *Then* $G \circ F: U \subset R^n \longrightarrow R^k$ *is a continuous map.*

Proof. Let $p \in U$ and let V be a neighborhood of $G \circ F(p)$ in R^k. By continuity of G, there is a neighborhood Q of $F(p)$ in R^m with $G(Q) \subset V$. By continuity of F, there is a neighborhood W of p in R^n with $F(W) \subset Q$. Thus,

$$G \circ F(W) \subset G(Q) \subset V,$$

and hence the continuity of $G \circ F$. **Q.E.D.**

It is often necessary to deal with maps defined on arbitrary (not necessarily open) sets of R^n. To extend the previous ideas to this situation, we shall proceed as follows.

Let $F: A \subset R^n \longrightarrow R^m$ be a map, where A is an arbitrary set in R^n. We say that F is *continuous in A* if there exists an open set $U \subset R^n$, $U \supset A$, and a continuous map $\bar{F}: U \longrightarrow R^m$ such that the restriction $\bar{F}|A = F$. In other words, F is continuous in A if it is the restriction of a continuous map defined in an open set containing A.

It is clear that if $F: A \subset R^n \longrightarrow R^m$ is continuous, given a neighborhood V of $F(p)$ in R^m, $p \in A$, there exists a neighborhood W of p in R^n such that $F(W \cap A) \subset V$. For this reason, it is convenient to call the set $W \cap A$ a *neighborhood of* p *in* A (Fig. A2-4).

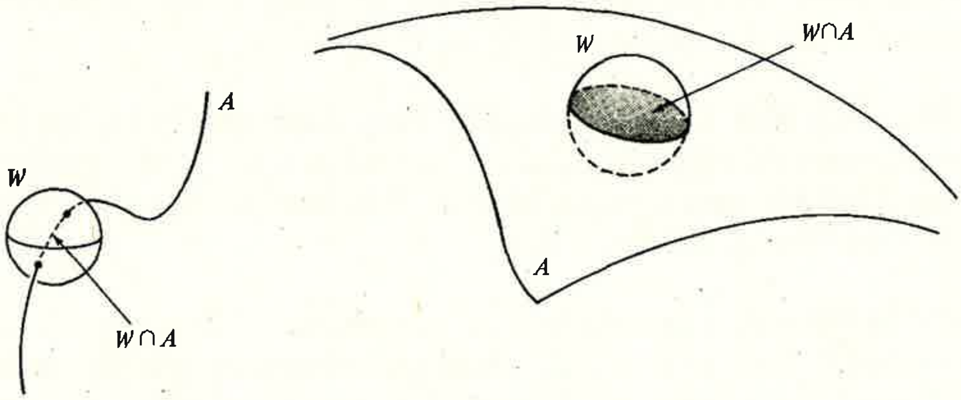

Figure A2-4

Example 5. Let

$$E = \left\{(x, y, z) \in R^3; \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1\right\}$$

be an ellipsoid, and let $\pi: R^3 \longrightarrow R^2$ be the projection of Example 3. Then the restriction of π to E is a continuous map from E to R^2.

We say that a continuous map $F: A \subset R^n \longrightarrow R^n$ is a *homeomorphism* onto $F(A)$ if F is one-to-one and the inverse $F^{-1}: F(A) \subset R^n \longrightarrow R^n$ is continuous. In this case A and $F(A)$ are *homeomorphic sets*.

Example 6. Let $F: R^3 \longrightarrow R^3$ be given by

$$F(x, y, z) = (xa, yb, zc).$$

F is clearly continuous, and the restriction of F to the sphere

$$S^2 = \{(x, y, z) \in R^3;\ x^2 + y^2 + z^2 = 1\}$$

is a continuous map $\tilde{F}: S^2 \longrightarrow R^3$. Observe that $\tilde{F}(S^2) = E$, where E is the ellipsoid of Example 5. It is also clear that F is one-to-one and that

$$F^{-1}(x, y, z) = \left(\frac{x}{a}, \frac{y}{b}, \frac{z}{c}\right).$$

Thus, $\tilde{F}^{-1} = F^{-1} \mid E$ is continuous. Therefore, $\tilde{F}$ is a homeomorphism of the sphere S^2 onto the ellipsoid E.

Finally, we want to describe two properties of real continuous functions defined on a closed interval $[a, b]$,

$$[a, b] = \{x \in R;\ a \leq x \leq b\}$$

(Props. 4 and 5 below), and an important property of the closed interval $[a, b]$ itself. They will be used repeatedly in this book.

PROPOSITION 4 (The Intermediate Value Theorem). *Let* $f: [a, b] \longrightarrow R$ *be a continuous function defined on the closed interval* $[a, b]$. *Assume that* $f(a)$ *and* $f(b)$ *have opposite signs; that is,* $f(a)f(b) < 0$. *Then there exists a point* $c \in (a, b)$ *such that* $f(c) = 0$.

PROPOSITION 5. *Let* $f: [a, b]$ *be a continuous function defined in the closed interval* $[a, b]$. *Then* f *reaches its maximum and its minimum in* $[a, b]$; *that is, there exist points* $x_1, x_2 \in [a, b]$ *such that* $f(x_1) \leq f(x) \leq f(x_2)$ *for all* $x \in [a, b]$.

PROPOSITION 6 (Heine-Borel). *Let* $[a, b]$ *be a closed interval and let* I_α, $\alpha \in A$, *be a collection of open intervals in* $[a, b]$ *such that* $\bigcup_\alpha I_\alpha = [a, b]$. *Then it is possible to choose a finite number* $I_{k_1}, I_{k_2}, \ldots, I_{k_n}$ *of* I_α *such that* $\bigcup I_{k_i} = I$, $i = 1, \ldots, n$.

These propositions are standard theorems in courses on advanced calculus, and we shall not prove them here. However, proofs are provided in the appendix to Chap. 5 (Props. 6, 13, and 11, respectively).

B. Differentiability in R^n

Let $f: U \subset R \longrightarrow R$. The *derivative* $f'(x_0)$ of f at $x_0 \in U$ is the limit (when it exists)

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}, \qquad x_0 + h \in U.$$

When f has derivatives at all points of a neighborhood V of x_0, we can consider the derivative of $f': V \longrightarrow R$ at x_0, which is called the *second derivative* $f''(x_0)$ of f at x_0, and so forth. f is *differentiable* at x_0 if it has continuous derivatives of all orders at x_0. f is *differentiable* in U if it is differentiable at all points in U.

Remark. We use the word differentiable for what is sometimes called infinitely differentiable (or of class C^∞). Our usage should not be confused with the usage of elementary calculus, where a function is called differentiable if its first derivative exists.

Let $F: U \subset R^2 \longrightarrow R$. The *partial derivative of* f *with respect to* x *at* $(x_0, y_0) \in U$, denoted by $(\partial f/\partial x)(x_0, y_0)$, is (when it exists) the derivative at x_0 of the function of one variable: $x \longrightarrow f(x, y_0)$. Similarly, the partial derivative with respect to y at (x_0, y_0), $(\partial f/\partial y)(x_0, y_0)$, is defined as the derivative at y_0 of $y \longrightarrow f(x_0, y)$. When f has partial derivatives at all points of a neighborhood V of (x_0, y_0), we can consider the *second partial derivatives at* (x_0, y_0):

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

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

and so forth. f is *differentiable* at (x_0, y_0) if it has continuous partial derivatives of all orders at (x_0, y_0). f is *differentiable* in U if it is differentiable at all points of U. We sometimes denote partial derivatives by

$$\frac{\partial f}{\partial x} = f_x, \qquad \frac{\partial f}{\partial y} = f_y, \qquad \frac{\partial^2 f}{\partial x^2} = f_{xx}, \qquad \frac{\partial^2 f}{\partial x\, \partial y} = f_{xy}, \qquad \frac{\partial^2 f}{\partial y^2} = f_{yy}.$$

It is an important fact that when f is differentiable the partial derivatives of f are independent of the order in which they are performed; that is,

$$\frac{\partial^2 f}{\partial x\, \partial y} = \frac{\partial^2 f}{\partial y\, \partial x}, \qquad \frac{\partial^3 f}{\partial^2 x\, \partial y} = \frac{\partial^3 f}{\partial x\, \partial y\, \partial x}, \qquad \text{etc.}$$

The definitions of partial derivatives and differentiability are easily extended to functions $f\colon U \subset R^n \longrightarrow R$. For instance, $(\partial f/\partial x_3)(x_1^0, x_2^0, \ldots, x_n^0)$ is the derivative of the function of one variable

$$x_3 \longrightarrow f(x_1^0, x_2^0, x_3, x_4^0, \ldots, x_n^0).$$

A further important fact is that partial derivatives obey the so-called *chain rule*. For instance, if $x = x(u, v)$, $y = y(u, v)$, $z = z(u, v)$ are real differentiable functions in $U \subset R^2$ and $f(x, y, z)$ is a real differentiable function in R^3, then the composition $f(x(u, v), y(u, v), z(u, v))$ is a differentiable function in U, and the partial derivative of f with respect to, say, u is given by

$$\frac{\partial f}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial u}.$$

We are now interested in extending the notion of differentiability to maps $F\colon U \subset R^n \longrightarrow R^m$. We say that F is *differentiable* at $p \in U$ if its component functions are differentiable at p; that is, by writing

$$F(x_1, \ldots, x_n) = (f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n)),$$

the functions f_i, $i = 1, \ldots, m$, have continuous partial derivatives of all orders at p. F is *differentiable* in U if it is differentiable at all points in U.

For the case $m = 1$, this repeats the previous definition. For the case $n = 1$, we obtain the notion of a (parametrized) *differentiable curve* in R^m. In Chap. 1, we have already seen such an object in R^3. For our purposes, we need to extend the definition of tangent vector of Chap. 1 to the present situation. A *tangent vector* to a map $\alpha\colon U \subset R \longrightarrow R^m$ at $t_0 \in U$ is the vector in R^m

$$\alpha'(t_0) = (x_1'(t_0), \ldots, x_m'(t_0)).$$

Example 7. Let $F\colon U \subset R^2 \longrightarrow R^3$ be given by

$$F(u, v) = (\cos u \cos v, \cos u \sin v, \cos^2 v), \qquad (u, v) \in U.$$

The component functions of F, namely,

$$f_1(u, v) = \cos u \cos v, \qquad f_2(u, v) = \cos u \sin v, \qquad f_3(u, v) = \cos^2 v$$

have continuous partial derivatives of all orders in U. Thus, F is differentiable in U.

Example 8. Let $\alpha\colon U \subset R \longrightarrow R^4$ be given by

$$\alpha(t) = (t^4, t^3, t^2, t), \qquad t \in U.$$

Then α is a differentiable curve in R^4, and the tangent vector to α at t is $\alpha'(t) = (4t^3, 3t^2, 2t, 1)$.

Example 9. Given a vector $w \in R^m$ and a point $p_0 \in U \subset R^m$, we can always find a differentiable curve $\alpha: (-\epsilon, \epsilon) \longrightarrow U$ with $\alpha(0) = p_0$ and $\alpha'(0) = w$. Simply define $\alpha(t) = p_0 + tw$, $t \in (-\epsilon, \epsilon)$. By writing $p_0 = (x_1^0, \ldots, x_m^0)$ and $w = (w_1, \ldots, w_m)$, the component functions of α are $x_i(t) = x_i^0 + tw_i$, $i = 1, \ldots, m$. Thus, α is differentiable, $\alpha(0) = p_0$ and

$$\alpha'(0) = (x_1'(0), \ldots, x_m'(0)) = (w_1, \ldots, w_m) = w.$$

We shall now introduce the concept of differential of a differentiable map. It will play an important role in this book.

DEFINITION 1. *Let* $F: U \subset R^n \longrightarrow R^m$ *be a differentiable map. To each* $p \in U$ *we associate a linear map* $dF_p: R^n \longrightarrow R^m$ *which is called the* differential *of* F *at* p *and is defined as follows. Let* $w \in R^n$ *and let* $\alpha: (-\epsilon, \epsilon) \longrightarrow U$ *be a differentiable curve such that* $\alpha(0) = p$, $\alpha'(0) = w$. *By the chain rule, the curve* $\beta = F \circ \alpha: (-\epsilon, \epsilon) \longrightarrow R^m$ *is also differentiable.* Then (Fig. A2-5)

$$dF_p(w) = \beta'(0).$$

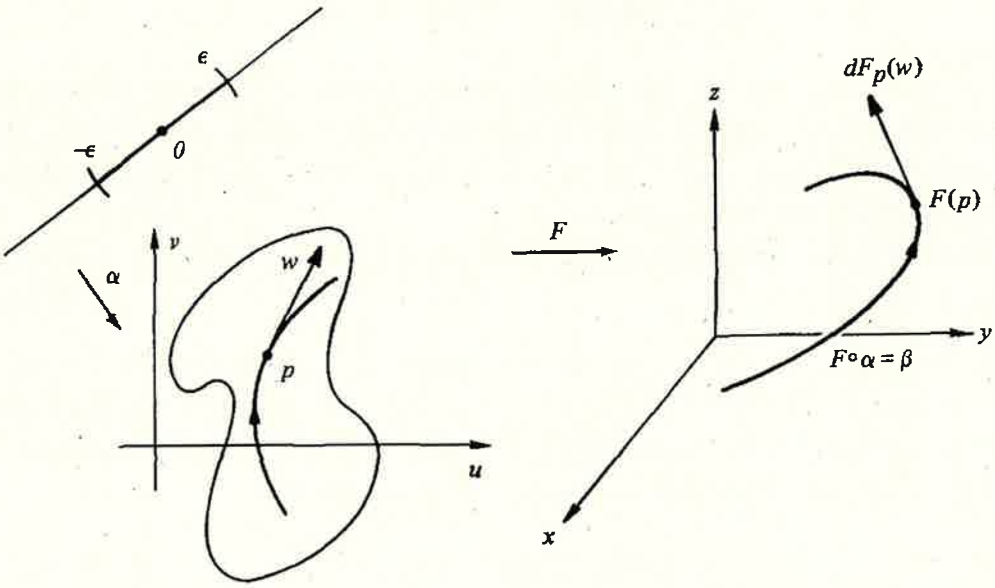

Figure A2-5

PROPOSITION 7. *The above definition of* dF_p *does not depend on the choice of the curve which passes through* p *with tangent vector* w, *and* dF_p *is, in fact, a linear map.*

Proof. To simplify notation, we work with the case $F: U \subset R^2 \longrightarrow R^3$. Let (u, v) be coordinates in R^2 and (x, y, z) be coordinates in R^3. Let

$e_1 = (1, 0)$, $e_2 = (0, 1)$ be the canonical basis in R^2 and $f_1 = (1, 0, 0)$, $f_2 = (0, 1, 0)$, $f_3 = (0, 0, 1)$ be the canonical basis in R^3. Then we can write $\alpha(t) = (u(t), v(t))$, $t \in (-\epsilon, \epsilon)$,

$$\alpha'(0) = w = u'(0)e_1 + v'(0)e_2,$$

$F(u, v) = (x(u, v), y(u, v), z(u, v))$, and

$$\beta(t) = F \circ \alpha(t) = (x(u(t), v(t)), y(u(t), v(t)), z(u(t), v(t))).$$

Thus, using the chain rule and taking the derivatives at $t = 0$, we obtain

$$\beta'(0) = \left(\frac{\partial x}{\partial u}\frac{du}{dt} + \frac{\partial x}{\partial v}\frac{dv}{dt}\right)f_1 + \left(\frac{\partial y}{\partial u}\frac{du}{dt} + \frac{\partial y}{\partial v}\frac{dv}{dt}\right)f_2$$
$$+ \left(\frac{\partial z}{\partial u}\frac{du}{dt} + \frac{\partial z}{\partial v}\frac{dv}{dt}\right)f_3$$
$$= \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} \end{pmatrix} \begin{pmatrix} \frac{du}{dt} \\ \\ \frac{dv}{dt} \end{pmatrix} = dF_p(w).$$

This shows that dF_p is represented, in the canonical bases of R^2 and R^3, by a matrix which depends only on the partial derivatives at p of the component functions x, y, z of F. Thus, dF_p is a linear map, and clearly $dF_p(w)$ does not depend on the choice of α.

The reader will have no trouble in extending this argument to the more general situation. **Q.E.D.**

The matrix of dF_p: $R^n \longrightarrow R^m$ in the canonical bases of R^n and R^m, that is, the matrix $(\partial f_i/\partial x_j)$, $i = 1, \ldots, m$, $j = 1, \ldots, n$, is called the *Jacobian matrix* of F at p. When $n = m$, this is a square matrix and its determinant is called the *Jacobian determinant;* it is usual to denote it by

$$\det\left(\frac{\partial f_i}{\partial x_j}\right) = \frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)}.$$

Remark. There is no agreement in the literature regarding the notation for the differential. It is also of common usage to call dF_p the derivative of F at p and to denote it by $F'(p)$.

Example 10. Let F: $R^2 \longrightarrow R^2$ be given by

$$F(x, y) = (x^2 - y^2, 2xy), \qquad (x, y) \in R^2.$$

F is easily seen to be differentiable, and its differential dF_p at $p = (x, y)$ is

$$dF_p = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}.$$

For instance, $dF_{(1,1)}(2, 3) = (-2, 10)$.

One of the advantages of the notion of differential of a map is that it allows us to express many facts of calculus in a geometric language. Consider, for instance, the following situation: Let $F: U \subset R^2 \to R^3$; $G: V \subset R^3 \to R^2$ be differentiable maps, where U and V are open sets such that $F(U) \subset V$. Let us agree on the following set of coordinates,

$$\begin{array}{ccccc} U \subset R^2 & \xrightarrow{F} & V \subset R^3 & \xrightarrow{G} & R^2 \\ (u, v) & & (x, y, z) & & (\xi, \eta) \end{array}$$

and let us write

$$F(u, v) = (x(u, v), y(u, v), z(u, v)),$$
$$G(x, y, z) = (\xi(x, y, z), \eta(x, y, z)).$$

Then

$$G \circ F(u, v) = (\xi(x(u, v), y(u, v), z(u, v)), \eta(x(u, v), y(u, v), z(u, v))),$$

and, by the chain rule, we can say that $G \circ F$ is differentiable and compute the partial derivatives of its component functions. For instance,

$$\frac{\partial \xi}{\partial u} = \frac{\partial \xi}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial \xi}{\partial y}\frac{\partial y}{\partial u} + \frac{\partial \xi}{\partial z}\frac{\partial z}{\partial u}.$$

Now, a simple way of expressing the above situation is by using the following general fact.

PROPOSITION 8 (The Chain Rule for Maps). *Let* $F: U \subset R^n \to R^m$ *and* $G: V \subset R^m \to R^k$ *be differentiable maps, where* U *and* V *are open sets such that* $F(U) \subset V$. *Then* $G \circ F: U \to R^k$ *is a differentiable map, and*

$$d(G \circ F)_p = dG_{F(p)} \circ dF_p, \qquad p \in U.$$

Proof. The fact that $G \circ F$ is differentiable is a consequence of the chain rule for functions. Now, let $w_1 \in R^n$ be given and let us consider a curve $\alpha: (-\epsilon_2, \epsilon_2) \to U$, with $\alpha(0) = p$, $\alpha'(0) = w_1$. Set $dF_p(w_1) = w_2$ and observe that $dG_{F(p)}(w_2) = (d/dt)(G \circ F \circ \alpha)|_{t=0}$. Then

$$d(G \circ F)_p(w_1) = \frac{d}{dt}(G \circ F \circ \alpha)_{t=0} = dG_{F(p)}(w_2) = dG_{F(p)} \circ dF_p(w_1).$$

Q.E.D.

Notice that, for the particular situation we were considering before, the relation $d(G \circ F)_p = dG_{F(p)} \circ dF_p$ is equivalent to the following product of Jacobian matrices,

$$\begin{pmatrix} \frac{\partial \xi}{\partial u} & \frac{\partial \xi}{\partial v} \\ \\ \frac{\partial \eta}{\partial u} & \frac{\partial \eta}{\partial v} \end{pmatrix} = \begin{pmatrix} \frac{\partial \xi}{\partial x} & \frac{\partial \xi}{\partial y} & \frac{\partial \xi}{\partial z} \\ \\ \frac{\partial \eta}{\partial x} & \frac{\partial \eta}{\partial y} & \frac{\partial \eta}{\partial z} \end{pmatrix} \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} \end{pmatrix},$$

which contains the expressions of all partial derivatives $\partial\xi/\partial u$, $\partial\xi/\partial v$, $\partial\eta/\partial u$, $\partial\eta/\partial v$. Thus, the simple expression of the chain rule for maps embodies a great deal of information on the partial derivatives of their component functions.

An important property of a differentiable function $f: (a, b) \subset R \longrightarrow R$ defined in an open interval (a, b) is that if $f'(x) \equiv 0$ on (a, b), then f is constant on (a, b). This generalizes for differentiable functions of several variables as follows.

We say that an open set $U \subset R^n$ is *connected* if given two points $p, q \in U$ there exists a continuous map $\alpha: [a, b] \longrightarrow U$ such that $\alpha(a) = p$ and $\alpha(b) = q$. This means that two points of U can be joined by a continuous curve in U or that U is made up of one single "piece."

PROPOSITION 9. *Let* $f: U \subset R^n \longrightarrow R$ *be a differentiable function defined on a connected open subset* U *of* R^n. *Assume that* $df_p: R^n \longrightarrow R$ *is zero at every point* $p \in U$. *Then* f *is constant on* U.

Proof. Let $p \in U$ and let $B_\delta(p) \subset U$ be an open ball around p and contained in U. Any point $q \in B_\epsilon(p)$ can be joined to p by the "radial" segment $\beta: [0, 1] \longrightarrow U$, where $\beta(t) = tq + (1 - t)p$, $t \in [0, 1]$ (Fig. A2-6). Since U is open, we can extend β to $(0 - \epsilon, 1 + \epsilon)$. Now, $f \circ \beta: (0 - \epsilon, 1 + \epsilon) \longrightarrow R$ is a function defined in an open interval, and

$$d(f \circ \beta)_t = (df \circ d\beta)_t = 0,$$

since $df \equiv 0$. Thus,

$$\frac{d}{dt}(f \circ \beta) = 0$$

for all $t \in (0 - \epsilon, 1 + \epsilon)$, and hence $(f \circ \beta) = \text{const}$. This means that $f(\beta(0)) = f(p) = f(\beta(1)) = f(q)$; that is, f is constant on $B_\delta(p)$.

Thus, the proposition is proved locally; that is, each point of U has a

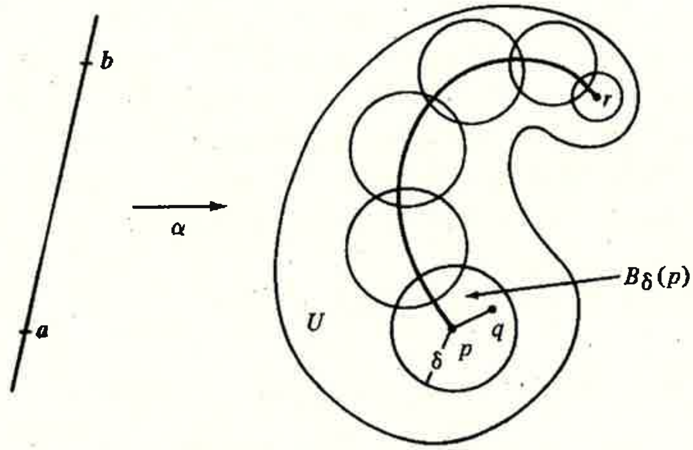

Figure A2-6

neighborhood such that f is constant on that neighborhood. Notice that so far we have not used the connectedness of U. We shall need it now to show that these constants are all the same.

Let r be an arbitrary point of U. Since U is connected, there exists a continuous curve $\alpha: [a, b] \longrightarrow U$, with $\alpha(a) = p, \alpha(b) = r$. The function $f \circ \alpha: [a, b] \longrightarrow R$ is continuous in $[a, b]$. By the first part of the proof, for each $t \in [a, b]$, there exists an interval I_t, open in $[a, b]$, such that $f \circ \alpha$ is constant on I_t. Since $\bigcup_t I_t = [a, b]$, we can apply the Heine-Borel theorem (Prop. 6). Thus, we can choose a finite number $I_1, \ldots, I_k$ of the intervals I_t so that $\bigcup_i I_t = [a, b]$, $i = 1, \ldots, k$. We can assume, by renumbering the intervals, if necessary, that two consecutive intervals overlap. Thus, $f \circ \alpha$ is constant in the union of two consecutive intervals. It follows that f is constant on $[a, b]$; that is,

$$f(\alpha(a)) = f(p) = f(\alpha(b)) = f(r).$$

Since r is arbitrary, f is constant on U. **Q.E.D.**

One of the most important theorems of differential calculus is the so-called inverse function theorem, which, in the present notation, says the following. (Recall that a linear map A is an isomorphism if the matrix of A is invertible.)

INVERSE FUNCTION THEOREM. *Let* $\mathrm{F}: \mathrm{U} \subset \mathrm{R}^n \longrightarrow \mathrm{R}^n$ *be a differentiable mapping and suppose that at* $\mathrm{p} \in \mathrm{U}$ *the differential* $d\mathrm{F}_p: \mathrm{R}^n \longrightarrow \mathrm{R}^n$ *is an isomorphism. Then there exists a neighborhood* V *of* p *in* U *and a neighborhood* W *of* F(p) *in* R^n *such that* $\mathrm{F}: \mathrm{V} \longrightarrow \mathrm{W}$ *has a differentiable inverse* $\mathrm{F}^{-1}: \mathrm{W} \longrightarrow \mathrm{V}$.

A differentiable mapping $F: V \subset R^n \longrightarrow W \subset R^n$, where V and W are open sets, is called a *diffeomorphism* of V with W if F has a differentiable inverse.

The inverse function theorem asserts that if at a point $p \in U$ the differential dF_p is an isomorphism, then F is a diffeomorphism in a neighborhood of p. In other words, an assertion about the differential of F at a point implies a similar assertion about the behavior of F in a neighborhood of the point.

This theorem will be used repeatedly in this book. A proof can be found, for instance, in Buck, *Advanced Calculus*, p. 285.

Example 11. Let $F: R^2 \to R^2$ be given by

$$F(x, y) = (e^x \cos y, e^x \sin y), \qquad (x, y) \in R^2.$$

The component functions of F, namely, $u(x, y) = e^x \cos y$, $v(x, y) = e^x \sin y$, have continuous partial derivatives of all orders. Thus, F is differentiable.

It is instructive to see, geometrically, how F transforms curves of the xy plane. For instance, the vertical line $x = x_0$ is mapped into the circle $u = e^{x_0} \cos y$, $v = e^{x_0} \sin y$ of radius e^{x_0}, and the horizontal line $y = y_0$ is mapped into the half-line $u = e^x \cos y_0$, $v = e^x \sin y_0$ with slope $\tan y_0$. It follows that (Fig. A2-7)

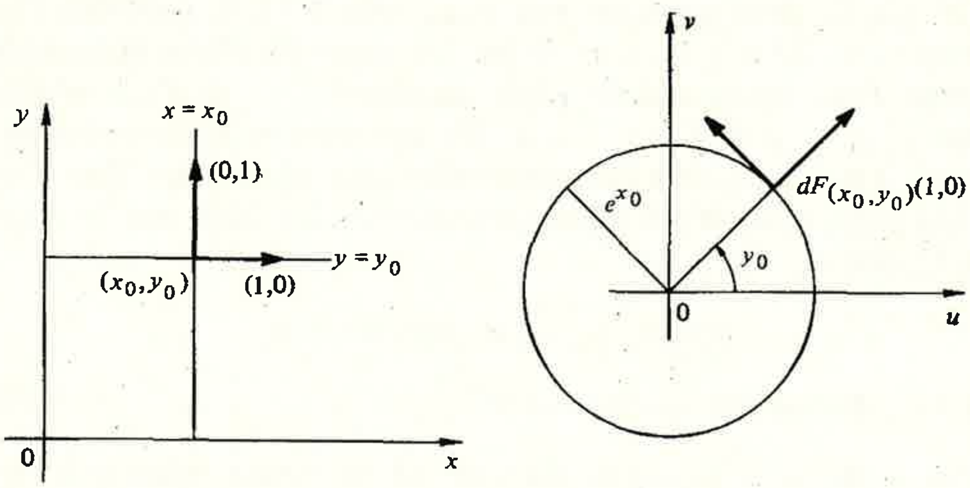

Figure A2-7

$$dF_{(x_0,y_0)}(1, 0) = \frac{d}{dx}(e^x \cos y_0, e^x \sin y_0)\big|_{x=x_0}$$
$$= (e^{x_0} \cos y_0, e^{x_0} \sin y_0),$$
$$dF_{(x_0,y_0)}(1, 0) = \frac{d}{dy}(e^{x_0} \cos y, e^{x_0} \sin y)\big|_{y=y_0}$$
$$= (-e^{x_0} \sin y_0, e^{x_0} \cos y_0).$$

This can be most easily checked by computing the Jacobian matrix of F,

$$dF_{(x,y)} = \begin{pmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{pmatrix} = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix},$$

and applying it to the vectors (1, 0) and (0, 1) at (x_0, y_0).

We notice that the Jacobian determinant $\det(dF_{(x,y)}) = e^x \neq 0$, and thus dF_p is nonsingular for all $p = (x, y) \in R^2$ (this is also clear from the previous geometric considerations). Therefore, we can apply the inverse function theorem to conclude that F is locally a diffeomorphism.

Observe that $F(x, y) = F(x, y + 2\pi)$. Thus, F is not one-to-one and has no global inverse. For each $p \in R^2$, the inverse function theorem gives neighborhoods V of p and W of $F(p)$ so that the restriction $F\colon V \longrightarrow W$ is a diffeomorphism. In our case, V may be taken as the strip $\{-\infty < x < \infty, 0 < y < 2\pi\}$ and W as $R^2 - \{(0, 0)\}$. However, as the example shows, even if the conditions of the theorem are satisfied everywhere and the domain of definition of F is very simple, a global inverse of F may fail to exist.