CHAPTER XV

Functions on n-Space

XV, §1. PARTIAL DERIVATIVES

Before considering the general case of a differentiable map of a vector space into another, we shall consider the special case of a function, i.e. a real valued map.

We consider functions on $\mathbf{R}^n$. A point of $\mathbf{R}^n$ is denoted by

$$x = (x_1, \ldots, x_n).$$

We use small letters even for points in $\mathbf{R}^n$ to fit the notation of the next chapter. Occasionally we still use a capital letter. In particular, if

$$A = (a_1, \ldots, a_n)$$

is an element of $\mathbf{R}^n$, we write $Ax = A \cdot x = a_1 x_1 + \cdots + a_n x_n$. The reason for using sometimes a capital and sometimes a small letter will appear later, when in fact the roles played by A and by x will be seen to correspond to different kinds of objects. In the special case which interests us in this chapter, we can still take them both in $\mathbf{R}^n$.

Let U be an open set of $\mathbf{R}^n$, and let $f: U \to \mathbf{R}$ be a function. We define its **partial derivative** at a point $x \in U$ by

$$\begin{aligned} D_i f(x) &= \lim_{h \to 0} \frac{f(x + he_i) - f(x)}{h} \\ &= \lim_{h \to 0} \frac{f(x_1, \ldots, x_i + h, \ldots, x_n) - f(x_1, \ldots, x_n)}{h} \end{aligned}$$

if the limit exists. Note here that $e_i = (0, \ldots, 1, \ldots, 0)$ is the unit vector with 1 in the i-th component and 0 at all other components, and $h \in \mathbf{R}$ approaches 0.

We sometimes use the notation

$$D_i f(x) = \frac{\partial f}{\partial x_i}.$$

We see that $D_i f$ is an ordinary derivative which keeps all variables fixed but the i-th variable. In particular, we know that the derivative of a sum, and the derivative of a constant times a function follow the usual rules, that is $D_i(f + g) = D_i f + D_i g$ and $D_i(cf) = cD_i f$ for any constant c.

Example. If $f(x, y) = 3x^3y^2$ then

$$\frac{\partial f}{\partial x} = D_1 f(x, y) = 9x^2y^2$$

and

$$\frac{\partial f}{\partial y} = D_2 f(x, y) = 6x^3y.$$

Of course we may iterate partial derivatives. In this example, we have

$$\frac{\partial^2 f}{\partial x\, \partial y} = D_1 D_2 f(x, y) = 18x^2y,$$

$$\frac{\partial^2 f}{\partial y\, \partial x} = D_2 D_1 f(x, y) = 18x^2y.$$

Observe that the two iterated partials are equal. This is not an accident, and is a special case of the following general theorem.

Theorem 1.1. *Let f be a function on an open set U in $\mathbf{R}^2$. Assume that the partial derivatives $D_1 f$, $D_2 f$, $D_1 D_2 f$ and $D_2 D_1 f$ exist and are continuous. Then*

$$D_1 D_2 f = D_2 D_1 f.$$

Proof. Let (x, y) be a point in U, and let h, k be small non-zero numbers. We consider the expression

$$g(x) = f(x, y + k) - f(x, y).$$

We apply the mean value theorem and conclude that there exists a number s_1 between x and $x + h$ such that

$$g(x + h) - g(x) = g'(s_1)h.$$

This yields

$$\begin{aligned} f(x + h, y + k) - f(x + h, y) - f(x, y + k) + f(x, y) \\ &= g(x + h) - g(x) \\ &= [D_1 f(s_1, y + k) - D_1 f(s_1, y)]h \\ &= D_2 D_1 f(s_1, s_2)kh \end{aligned}$$

with some number s_2 between y and $y + k$.

Applying the same procedure to $g_2(y) = f(x + h, y) - f(x, y)$, we find that there exist numbers t_1, t_2 between x, $x + h$ and y, $y + k$ respectively such that

$$D_2 D_1 f(s_1, s_2)kh = D_1 D_2 f(t_1, t_2)kh.$$

We cancel the kh, and let $(h, k) \to (0, 0)$. Using the continuity of the repeated derivatives yields $D_2 D_1 f(x, y) = D_1 D_2 f(x, y)$, as desired.

Consider a function of three variables $f(x, y, z)$. We can then take three kinds of partial derivatives: D_1, D_2 or D_3; in other notation, $\partial/\partial x$, $\partial/\partial y$, $\partial/\partial z$. Let us assume throughout that all the partial derivatives which we shall consider exist and are continuous, so that we may form as many repeated partial derivatives as we please. Then using Theorem 1.1 we can show that it does not matter in which order we take partials. For example, if we have a function of three variables x_1, x_2, x_3 we find that

$$D_3 D_1 f = D_1 D_3 f.$$

This is simply an application of Theorem 1.1 keeping the second variable fixed. We may then take a further partial derivative, for instance

$$D_1 D_3 D_1 f.$$

Here D_1 occurs twice and D_3 occurs once. Interchanging D_3 and D_1 by Theorem 1.1 we get

$$D_1 D_3 D_1 f = D_1 D_1 D_3 f = D_1^2 D_3 f.$$

In general an iteration of partials can be written

$$D_1^{k_1} D_2^{k_2} D_3^{k_3} f$$

with integers $k_1, k_2, k_3 \geqq 0$. Similar remarks apply to n variables, in which case iterated partials can be written

$$D_1^{k_1} \cdots D_n^{k_n}$$

with integers $k_i \geqq 0$. Such a product expression is called an **elementary partial differential operator**. The sum

$$k_1 + \cdots + k_n$$

is called its **degree** or **order**. For instance $D_1^2 D_3$ has degree $2 + 1 = 3$.

Remark. Some smoothness assumption has to be made in order to have the commutativity of the partial derivatives. See Exercise 12 for a counterexample.

In the exercises, we deal with polar coordinates.
Let $x = r\cos\theta$ and $y = r\sin\theta$. Let

$$f(x, y) = g(r, \theta).$$

We wish to express $\partial g/\partial r$ and $\partial g/\partial\theta$ in terms of $\partial f/\partial x$ and $\partial f/\partial y$. We have

$$g(r, \theta) = f(r\cos\theta, r\sin\theta).$$

Hence

$$\frac{\partial g}{\partial r} = D_1 f(x, y)\frac{\partial x}{\partial r} + D_2 f(x, y)\frac{\partial y}{\partial r}$$

so

$$(*)\qquad \frac{\partial g}{\partial r} = D_1 f(x, y)\cos\theta + D_2 f(x, y)\sin\theta.$$

Similarly,

$$\frac{\partial g}{\partial\theta} = D_1 f(x, y)\frac{\partial x}{\partial\theta} + D_2 f(x, y)\frac{\partial y}{\partial\theta},$$

so

$$(**)\qquad \frac{\partial g}{\partial\theta} = D_1 f(x, y)(-r\sin\theta) + D_2 f(x, y) r\cos\theta.$$

Instead of writing $D_1 f(x, y)$, $D_2 f(x, y)$, we may write $\partial f/\partial x$ and $\partial f/\partial y$ respectively.

In 2-space, the operator

$$\Delta = D_1^2 + D_2^2 \qquad \text{or} \qquad \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

is called the **Laplace operator**. In Exercises 6 and 7 you will be asked to express it in terms of polar coordinates. A function f such that $\Delta f = 0$ is called **harmonic**. Important functions which are harmonic are given in polar coordinates by

$$r^n \cos n\theta \qquad \text{and} \qquad r^n \sin n\theta,$$

where n is a positive integer. You can prove easily that these functions are harmonic by using the formula of Exercise 6. They are fundamental in the theory of harmonic functions because other harmonic functions are expressed in terms of these, as infinite series

$$g(r, \theta) = \sum_{n=0}^{\infty} a_n r^n \cos n\theta + \sum_{n=1}^{\infty} b_n r^n \sin n\theta,$$

with appropriate constant coefficients a_n and b_n. In general, a C^2-function f on an open set of $\mathbf{R}^n$ is called **harmonic** if $(D_1^2 + \cdots + D_n^2)f = 0$.

XV, §1. EXERCISES

In the exercises, assume that all repeated partial derivatives exist and are continuous as needed.

1. Let f, g be two functions of two variables with continuous partial derivatives of order $\leqq 2$ in an open set U. Assume that

$$\frac{\partial f}{\partial x} = -\frac{\partial g}{\partial y} \qquad \text{and} \qquad \frac{\partial f}{\partial y} = \frac{\partial g}{\partial x}.$$

   Show that

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0.$$

2. Let f be a function of three variables, defined for $X \neq O$ by $f(X) = 1/|X|$. Show that

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = 0$$

   if the three variables are (x, y, z). (The norm is the euclidean norm.)

3. Let $f(x, y) = \arctan(y/x)$ for $x > 0$. Show that

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0.$$

4. Let θ be a fixed number, and let

$$x = u \cos \theta - v \sin \theta, \qquad y = u \sin \theta + v \cos \theta.$$

Let f be a function of two variables, and let $f(x, y) = g(u, v)$. Show that

$$\left(\frac{\partial g}{\partial u}\right)^2 + \left(\frac{\partial g}{\partial v}\right)^2 = \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2.$$

5. Assume that f is a function satisfying

$$f(tx, ty) = t^m f(x, y)$$

for all numbers x, y, and t. Show that

$$x^2 \frac{\partial^2 f}{\partial x^2} + 2xy \frac{\partial^2 f}{\partial x\, \partial y} + y^2 \frac{\partial^2 f}{\partial y^2} = m(m-1)f(x, y).$$

[*Hint*: Differentiate twice with respect to t. Then put $t = 1$.]

6. Let $x = r \cos \theta$ and $y = r \sin \theta$. Let $f(x, y) = g(r, \theta)$. Show that

$$\frac{\partial f}{\partial x} = \cos \theta \frac{\partial g}{\partial r} - \frac{\sin \theta}{r} \frac{\partial g}{\partial \theta},$$

$$\frac{\partial f}{\partial y} = \sin \theta \frac{\partial g}{\partial r} + \frac{\cos \theta}{r} \frac{\partial g}{\partial \theta}.$$

[*Hint*: Solve the simultaneous system of linear equations (*) and (**) given in the example of the text.]

7. Let $x = r \cos \theta$ and $y = r \sin \theta$. Let $f(x, y) = g(r, \theta)$. Show that

$$\frac{\partial^2 g}{\partial r^2} + \frac{1}{r}\frac{\partial g}{\partial r} + \frac{1}{r^2}\frac{\partial^2 g}{\partial \theta^2} = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}.$$

This exercise gives the polar coordinate form of the Laplace operator, and we can write symbolically:

$$\boxed{\left(\frac{\partial}{\partial x}\right)^2 + \left(\frac{\partial}{\partial y}\right)^2 = \left(\frac{\partial}{\partial r}\right)^2 + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\left(\frac{\partial}{\partial \theta}\right)^2}$$

[*Hint for the proof*: Start with (*) and (**) and take further derivatives as needed. Then take the sum. Lots of things will cancel out leaving you with $D_1^2 f + D_2^2 f$.]

8. With the same notation as in the preceding exercise, show that

$$\left(\frac{\partial g}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial g}{\partial \theta}\right)^2 = \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2.$$

9. In $\mathbf{R}^2$, suppose that $f(x, y) = g(r)$ where $r = \sqrt{x^2 + y^2}$. Show that

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \frac{d^2 g}{dr^2} + \frac{1}{r}\frac{dg}{dr}.$$

10. (a) In $\mathbf{R}^3$, suppose that $f(x, y, z) = g(r)$ where $r = \sqrt{x^2 + y^2 + z^2}$. Show that

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = \frac{d^2 g}{dr^2} + \frac{2}{r}\frac{dg}{dr}.$$

(b) Assume that f is harmonic except possibly at the origin on $\mathbf{R}^n$, and that there is a C^2 function g such that $f(X) = g(r)$ where $r = \sqrt{X \cdot X}$. Let $n \geqq 3$. Show that there exist constants C, K such that $g(r) = Kr^{2-n} + C$. What if $n = 2$?

11. Let $r = \sqrt{x^2 + y^2}$ and let r, θ be the polar coordinates in the plane. Using the formula for the Laplace operator in Exercise 7 verify that the following functions are harmonic:
(a) $r^n \cos n\theta = g(r, \theta)$ (b) $r^n \sin n\theta = g(r, \theta)$
As usual, n denotes a positive integer. So you are supposed to prove that the expression

$$\frac{\partial^2 g}{\partial r^2} + \frac{1}{r}\frac{\partial g}{\partial r} + \frac{1}{r^2}\frac{\partial^2 g}{\partial \theta^2}$$

is equal to 0 for the above functions g.

12. For $x \in \mathbf{R}^n$ let $x^2 = x_1^2 + \cdots + x_n^2$. For t real > 0, let

$$f(x, t) = t^{-n/2} e^{-x^2/4t}.$$

If Δ is the Laplace operator, $\Delta = \sum \partial^2/\partial x_i^2$, show that $\Delta f = \partial f/\partial t$. A function satisfying this differential equation is said to be a solution of the **heat equation.**

13. This exercise gives an example of a function whose repeated partials exist but such that $D_1 D_2 f \neq D_2 D_1 f$. Let

$$f(x, y) = \begin{cases} xy\dfrac{x^2 - y^2}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0), \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$$

Prove:
(a) The partial derivatives $\partial^2 f/\partial x\,\partial y$ and $\partial^2 f/\partial y\,\partial x$ exist for all (x, y) and are continuous except at (0, 0).
(b) $D_1 D_2 f(0, 0) \neq D_2 D_1 f(0, 0)$.

Green's function

14. Let (a, b) be an open interval, which may be (a, ∞). Let

$$M_y = -\left(\frac{d}{dy}\right)^2 + p(y),$$

where p is an infinitely differentiable function. We view M_y as a differential operator. If f is a function of the variable y, then we use the notation

$$M_y f(y) = -f''(y) + p(y)f(y).$$

A Green's function for the differential operator M is a suitably smooth function $g(y, y')$ defined for y, y' in (a, b) such that

$$M_y \int_a^b g(y, y')f(y')\,dy' = f(y)$$

for all infinitely differentiable functions f on (a, b) with compact support (meaning f is 0 outside a closed interval contained in (a, b)). Now let $g(y, y')$ be any continuous function satisfying the following additional conditions:

GF 1. *g is infinitely differentiable in each variable except on the diagonal, that is when $y = y'$.*

GF 2. *If $y \neq y'$, then $M_y g(y, y') = 0$.*

Prove:

Let g be a function satisfying **GF 1** *and* **GF 2**. *Then g is a Green's function for the operator M if and only if g also satisfies the jump condition*

GF 3. $D_1 g(y, y+) - D_1 g(y, y-) = 1$.

As usual, one defines

$$D_1 g(y, y+) = \lim_{\substack{y' \to y \\ y' > y}} D_1 g(y, y'),$$

and similarly for $y-$ instead of $y+$, we take the limit with $y' < y$. [*Hint*: Write the integral

$$\int_a^b = \int_a^y + \int_y^b.\Big]$$

15. Assume now that the differential equation $f'' - pf = 0$ has two linearly independent solutions J and K. (You will be able to prove this after reading the chapter on the existence and uniqueness of solutions of differential equations. See Chapter XIX, §3, Exercise 2.) Let $W = JK' - J'K$.
(a) Show that W is constant $\neq 0$.

(b) Show that there exists a unique Green's function of the form

$$g(y, y') = \begin{cases} A(y')J(y) & \text{if } y' < y, \\ B(y')K(y) & \text{if } y' > y, \end{cases}$$

and that the functions A, B necessarily have the values $A = K/W$, $B = J/W$.

16. On the interval $(-\infty, \infty)$ let $M_y = -(d/dy)^2 + c^2$ where c is a positive number, so take $p = c > 0$ constant. Show that e^{cy} and e^{-cy} are two linearly independent solutions and write down explicitly the Green's function for M_y.

17. On the interval $(0, \infty)$ let

$$M_y = -\left(\frac{d}{dy}\right)^2 - \frac{s(1-s)}{y^2}$$

where s is some fixed complex number. For $s \neq \frac{1}{2}$, show that y^{1-s} and y^s are two linearly independent solutions and write down explicitly the Green's function for the operator.

XV, §2. DIFFERENTIABILITY AND THE CHAIN RULE

A function φ defined for all sufficiently small *vectors* $h \in \mathbf{R}^n$, $h \neq 0$, is said to be $o(h)$ for $h \to 0$ if

$$\lim_{h \to 0} \frac{\varphi(h)}{|h|} = 0.$$

Observe that here, $h = (h_1, \ldots, h_n)$ is a *vector* with components h_i which are *numbers*.

We use any norm $|\ |$ on $\mathbf{R}^n$ (usually in practice the euclidean norm or the sup norm). Of course we cannot divide a function by a vector, so we divide by the norm of the vector.

If a function $\varphi(h)$ is $o(h)$, then we can write it in the form

$$\varphi(h) = |h|\psi(h),$$

where

$$\lim_{h \to 0} \psi(h) = 0.$$

All we have to do is to let $\psi(h) = \varphi(h)/|h|$ for $h \neq 0$. Thus at first ψ is defined for sufficiently small $h \neq 0$. However, we may extend the function ψ by continuity so that it is defined at 0 by $\psi(0) = 0$.

We say a function $f: U \to \mathbf{R}$ is **differentiable** at a point x if there exists a vector $A \in \mathbf{R}^n$ such that

$$f(x + h) = f(x) + A \cdot h + o(h).$$

By this we mean that there is a function φ defined for all sufficiently small values of $h \neq 0$ such that $\varphi(h) = o(h)$ for $h \to 0$ and

$$f(x + h) = f(x) + A \cdot h + \varphi(h).$$

In view of our preceding remark, we can express this equality by the condition that there exists a function ψ defined for all sufficiently small h such that

$$\lim_{h \to 0} \psi(h) = 0$$

and

$$f(x + h) = f(x) + A \cdot h + |h|\psi(h).$$

We can include the value of ψ at 0 because when $h = 0$ we have indeed $f(x) = f(x) + A \cdot 0$.

We define the **gradient** of f at any point x at which all partial derivatives exist to be the vector

$$\operatorname{grad} f(x) = (D_1 f(x), \dots, D_n f(x)).$$

One should of course write $(\operatorname{grad} f)(x)$ but we omit one set of parentheses for simplicity.

Sometimes we use the notation $\partial f/\partial x_i$ for the partial derivative, and so

$$\operatorname{grad} f(x) = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}\right).$$

Theorem 2.1. *Let f be differentiable at a point x and let A be a vector such that*

$$f(x + h) = f(x) + A \cdot h + o(h).$$

Then all partial derivatives of f at x exist, and

$$A = \operatorname{grad} f(x).$$

Conversely, assume that all partial derivatives of f exist in some open set containing x and are continuous functions. Then f is differentiable at x.

Proof. Let $A = (a_1, \ldots, a_n)$. The first assertion follows at once by letting $h = te_i$ with real t and letting $t \to 0$. It is then the definition of partial derivatives that $a_i = D_i f(x)$. As to the second, we use the mean value theorem repeatedly as follows. We write

$$\begin{aligned}
&f(x_1 + h_1, \ldots, x_n + h_n) - f(x_1, \ldots, x_n)\\
&\quad = f(x_1 + h_1, \ldots, x_n + h) \qquad\qquad - f(x_1, x_2 + h_2, \ldots, x_n + h_n)\\
&\quad\quad + f(x_1, x_2 + h_2, \ldots, x_n + h_n) - f(x_1, x_2, \ldots, x_n + h_n)\\
&\quad\quad\ \vdots \qquad\qquad\qquad\qquad\qquad\qquad \vdots\\
&\quad\quad + f(x_1, \ldots, x_{n-1}, x_n + h_n) \qquad - f(x_1, \ldots, x_n)\\
&\quad = D_1 f(c_1, x_2 + h_2, \ldots, x_n + h_n)h_1 + \cdots + D_n f(x_1, \ldots, x_{n-1}, c_n)h_n,
\end{aligned}$$

where $c_1, \ldots, c_n$ lie between $x_i + h_i$ and x_i, respectively. By continuity, for each i there exists a function ψ_i such that

$$\lim_{h \to 0} \psi_i(h) = 0$$

and such that

$$D_i f(x_1, \ldots, x_{i-1}, c_i, \ldots, x_n + h_n) = D_i f(x) + \psi_i(h).$$

Hence

$$\begin{aligned}
f(x + h) - f(x) &= \sum_{i=1}^{n} (D_i f(x) + \psi_i(h))h_i\\
&= \sum_{i=1}^{n} D_i f(x)h_i + \sum_{i=1}^{n} \psi_i(h)h_i.
\end{aligned}$$

It is now clear that the first term on the right is nothing but $\operatorname{grad} f(x) \cdot h$ and the second term is $o(h)$, as was to be shown.

Remark. Some sort of condition on the partial derivatives has to be placed so that a function is differentiable. For instance, let

$$f(x, y) = \frac{xy}{x^2 + y^2} \quad \text{if} \quad (x, y) \neq (0, 0),$$

$$f(0, 0) = 0.$$

You can verify that $D_1 f(x, y)$ and $D_2 f(x, y)$ are defined for all (x, y), including at $(0, 0)$, but f is not differentiable at $(0, 0)$. It is not even continuous at the origin.

When the function f is differentiable, we see from Theorem 2.1 that the gradient of f takes the place of a derivative. We say that f is **differentiable on** U if it is differentiable at every point of U. The rules for derivative of a sum hold as usual: If f, g are differentiable, then

$$\operatorname{grad}(f + g) = \operatorname{grad} f + \operatorname{grad} g;$$

and if c is a number,

$$\operatorname{grad}(cf) = c \operatorname{grad} f.$$

One could formulate a rule for the product of two functions as usual, but we leave this to the reader. At the moment, we do not have an interpretation for the gradient. We shall derive one later. We shall use a technique reducing certain questions in several variables to questions in one variable as follows. Suppose f is defined on an open set U, and let

$$\varphi : [a, b] \to U$$

be a differentiable curve. Then we may form the composite function $f \circ \varphi$ given by

$$(f \circ \varphi)(t) = f(\varphi(t)).$$

We may think of φ as parametrizing a curve, or we may think of $\varphi(t)$ as representing the position of a particle at time t. If f represents, say, the temperature function, then $f(\varphi(t))$ is the temperature of the particle at time t. The rate of change of temperature of the particle along the curve is then given by the derivative $df(\varphi(t))/dt$. The chain rule which follows gives an expression for this derivative in terms of the gradient, and generalizes the usual chain rule to n variables.

Theorem 2.2. *Let $\varphi : J \to \mathbf{R}^n$ be a differentiable function defined on some interval, and with values in an open set U of $\mathbf{R}^n$. Let $f : U \to \mathbf{R}$ be a differentiable function. Then $f \circ \varphi : J \to \mathbf{R}$ is differentiable, and*

$$(f \circ \varphi)'(t) = \operatorname{grad} f(\varphi(t)) \cdot \varphi'(t).$$

Proof. By the definition of differentiability, say at a point $t \in J$, there is a function ψ such that

$$\lim_{k \to 0} \psi(k) = 0,$$

and

$$\begin{aligned} f(\varphi(t + h)) - f(\varphi(t)) = \operatorname{grad} f(\varphi(t)) \cdot (\varphi(t + h) - \varphi(t)) \\ + |\varphi(t + h) - \varphi(t)| \psi(k(h)), \end{aligned}$$

where $k(h) = \varphi(t + h) - \varphi(t)$. Divide by the *number* h to get

$$\frac{f(\varphi(t+h)) - f(\varphi(t))}{h} = \operatorname{grad} f(\varphi(t)) \cdot \frac{\varphi(t+h) - \varphi(t)}{h}$$
$$\pm \left| \frac{\varphi(t+h) - \varphi(t)}{h} \right| \psi(k(h)).$$

Take the limit as $h \to 0$ to obtain the statement of the chain rule.

Application: interpretation of the gradient

From the chain rule we get a simple example giving a geometric interpretation for the gradient. Let x be a point of U and let v be a fixed vector of norm 1. We define the **directional derivative** of f at x in the direction of v to be

$$\boxed{D_v f(x) = \frac{d}{dt} f(x + tv)\bigg|_{t=0}.}$$

This means that if we let $g(t) = f(x + tv)$, then

$$D_v f(x) = g'(0).$$

By the chain rule, $g'(t) = \operatorname{grad} f(x + tv) \cdot v$, whence

$$\boxed{D_v f(x) = \operatorname{grad} f(x) \cdot v.}$$

From this formula we obtain an interpretation for the gradient. We use the standard expression for the dot product, namely

$$D_v f(x) = |\operatorname{grad} f(x)|\,|v| \cos\theta,$$

where θ is the angle between v and $\operatorname{grad} f(x)$. Depending on the direction of the unit vector v, the number $\cos\theta$ ranges from -1 to $+1$. The maximal value occurs when v has the same direction as $\operatorname{grad} f(x)$, in which case for such **unit vector** v we obtain

$$D_v f(x) = |\operatorname{grad} f(x)|.$$

Therefore we get an interpretation for the direction and norm of the gradient:

> *The direction of* grad $f(x)$ *is the direction of maximal increase of the function* f *at* x.
> *The norm* $|\text{grad}\, f(x)|$ *is equal to the rate of change of* f *in its direction of maximal increase.*

Example. Find the directional derivative of the function $f(x, y) = x^2y^3$ at $(1, -2)$ in the direction of $(3, 1)$.

Let $A = (3, 1)$. Direction is meant from the origin to A. Note that A is not a unit vector, so we have to use a unit vector in the direction of A, namely

$$v = \frac{1}{\sqrt{10}}(3, 1).$$

We have $\text{grad}\, f(x, y) = (2xy^3, 3x^2y^2)$ and $\text{grad}\, f(1, -2) = (-16, 12)$. Hence the desired directional derivative is

$$\begin{aligned} D_v f(1, -2) &= (-16, 12) \cdot \frac{1}{\sqrt{10}}(3, 1) \\ &= \frac{1}{\sqrt{10}}(-36). \end{aligned}$$

Consider the set of all $x \in U$ such that $f(x) = 0$; or given a number c, the set of all $x \in U$ such that $f(x) = c$. This set, which we denote by S_c, is called the **level hypersurface of level** c. Let $x \in S_c$ and assume again that $\text{grad}\, f(x) \neq 0$. It will be shown as a consequence of the implicit function theorem that given any direction perpendicular to the gradient, there exists a differentiable curve

$$\alpha: J \to U$$

defined on some interval J containing 0 such that $\alpha(0) = x$, $\alpha'(0)$ has the given direction, and $f(\alpha(t)) = c$ for all $t \in J$. In other words, the curve is contained in the level hypersurface. Without proving the existence of such a curve, we see from the chain rule that if we have a curve α lying in the hypersurface such that $\alpha(0) = x$, then

$$0 = \frac{d}{dt} f(\alpha(t)) = \text{grad}\, f(\alpha(t)) \cdot \alpha'(t).$$

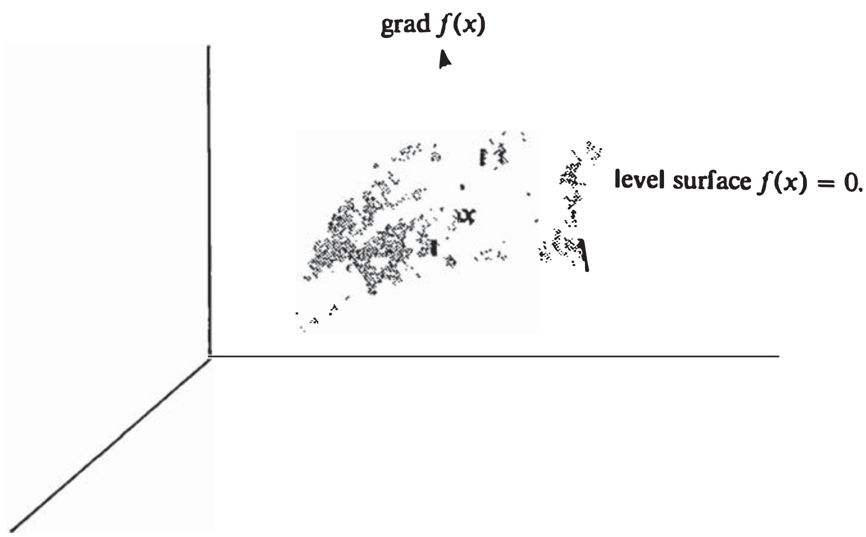
In particular, for $t = 0$,

$$\begin{aligned} 0 &= \operatorname{grad} f(\alpha(0)) \cdot \alpha'(0) \\ &= \operatorname{grad} f(x) \cdot \alpha'(0). \end{aligned}$$

Hence the velocity vector $\alpha'(0)$ of the curve at $t = 0$ is perpendicular to grad $f(x)$. From this we make the geometric conclusion that

grad $f(x)$ *is perpendicular to the level hypersurface at* x.

Thus geometrically the situation looks like this:



Application: the tangent plane

We want to apply the chain rule to motivate a definition of the tangent plane to a surface. For this we need to recall a little more explicitly some properties of linear algebra. We denote n-tuples in $\mathbf{R}^n$ by capital letters. If

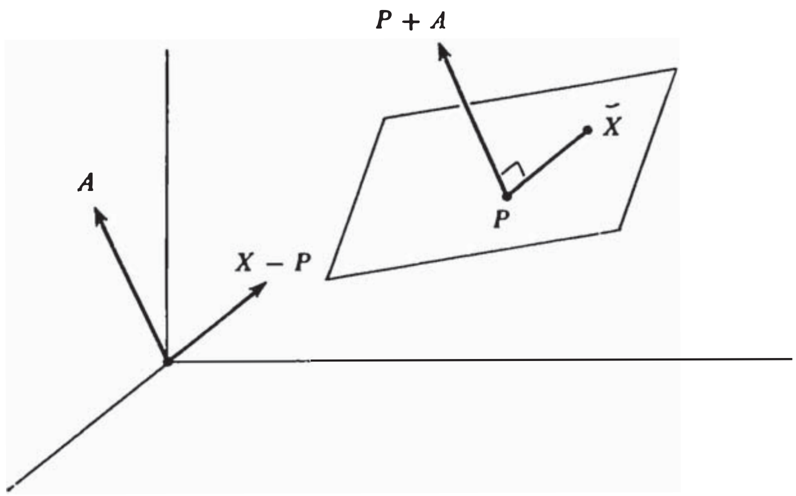
$$A = (a_1, \ldots, a_n) \qquad \text{and} \qquad B = (b_1, \ldots, b_n)$$

are elements of $\mathbf{R}^n$, we have already seen that A is perpendicular to B if and only if $A \cdot B = 0$.

Let $A \in \mathbf{R}^n$, $A \neq O$ and let P be a point in R^n. We define the **hyperplane through P perpendicular to A** to be the set of all points X such that

$$(X - P) \cdot A = 0,$$

or also $X \cdot A = P \cdot A$. This corresponds to the figure as shown. The set of points Y such that $Y \cdot A = 0$ is the hyperplane passing through the origin, perpendicular to A, and the hyperplane through P, perpendicular to A, is a translation by P.



Example. The plane in 3-space passing through $(1, -2, 3)$, perpendicular to $A = (-2, 4, 1)$ has the equation

$$-2x + 4y + z = -2 - 8 + 3 = -7.$$

Let f be a differentiable function on some open set U in $\mathbf{R}^n$. Let c be a number, and let S be the set of points X such that

$$f(X) = c, \qquad \text{but} \qquad \operatorname{grad} f(X) \neq O.$$

The set S is called a **hypersurface** in $\mathbf{R}^n$. Let P be a point of S. We define the **tangent hyperplane** of S at P to be the hyperplane passing through P perpendicular to $\operatorname{grad} f(P)$.

Example. Let $f(x, y, z) = x^2 + y^2 + z^2$. The surface S of points X such that $f(X) = 4$ is the sphere of radius 2 centered at the origin. Let

$$P = (1, 1, \sqrt{2}).$$

We have $\operatorname{grad} f(x, y, z) = (2x, 2y, 2z)$ and so

$$\operatorname{grad} f(P) = (2, 2, 2\sqrt{2}).$$

Hence the tangent plane at P is given by the equation

$$2x + 2y + 2\sqrt{2}z = 8.$$

Functions depending only on the distance. Let f be a differentiable function on $\mathbf{R}^n - \{0\}$, depending only on the distance from the origin, that is, there exists a differentiable function g of one variable $r > 0$ such that

$$f(X) = g(r) \qquad \text{where} \qquad r = \sqrt{X \cdot X} = \sqrt{x_1^2 + \cdots + x_n^2}.$$

Then a routine differentiation using the chain rule shows that

$$\boxed{(\operatorname{grad} f)(X) = \frac{g'(r)}{r} X.}$$

Carry out the differentiation as Exercise 3.

XV, §2. EXERCISES

1. Show that any two points on the sphere of radius 1 (or any radius) in n-space centered at the origin can be joined by a differentiable curve. If the points are not antipodal, divide the straight line between them by its length at each point. Or use another method: taking the plane containing the two points, and using two perpendicular vectors of lengths 1 in this plane, say A, B, consider the unit circle

$$\alpha(t) = (\cos t)A + (\sin t)B.$$

2. Let f be a differentiable function on $\mathbf{R}^n$, and assume that there is a differentiable function h such that

$$(\operatorname{grad} f)(X) = h(X)X.$$

Show that f is constant on the sphere of radius r centered at the origin in $\mathbf{R}^n$. [*Hint*: Use Exercise 1.]

3. Prove the converse of Exercise 2, which is the last statement preceding the exercises, namely if $f(X) = g(r)$, then $\operatorname{grad} f(X) = g'(r)X/r$.

4. Let f be a differentiable function on $\mathbf{R}^n$ and assume that there is a positive integer m such that $f(tX) = t^m f(X)$ for all numbers $t \neq 0$ and all points X in $\mathbf{R}^n$. Prove **Euler's relation**:

$$x_1 \frac{\partial f}{\partial x_1} + \cdots + x_n \frac{\partial f}{\partial x_n} = mf(X).$$

5. Let f be a differentiable function defined on all of space. Assume that

$$f(tP) = tf(P)$$

for all numbers t and all points P. Show that

$$f(P) = \operatorname{grad} f(O) \cdot P.$$

6. Find the equation of the tangent plane to each of the following surfaces at the specified point.

(a) $x^2 + y^2 + z^2 = 49$ at $(6, 2, 3)$
(b) $x^2 + xy^2 + y^3 + z + 1 = 0$ at $(2, -3, 4)$
(c) $x^2y^2 + xz - 2y^3 = 10$ at $(2, 1, 4)$
(d) $\sin xy + \sin yz + \sin xz = 1$ at $(1, \pi/2, 0)$

7. Find the directional derivative of the following functions at the specified points in the specified directions.
(a) $\log(x^2 + y^2)^{1/2}$ at $(1, 1)$, direction $(2, 1)$
(b) $xy + yz + xz$ at $(-1, 1, 7)$, direction $(3, 4, -12)$

8. Let $f(x, y, z) = (x + y)^2 + (y + z)^2 + (z + x)^2$. What is the direction of greatest increase of the function at the point $(2, -1, 2)$? What is the directional derivative of f in this direction at that point?

9. Let f be a differentiable function defined on an open set U. Suppose that P is a point of U such that $f(P)$ is a maximum, that is suppose we have

$$f(P) \geqq f(X) \qquad \text{for all } X \text{ in } U.$$

Show that $\operatorname{grad} f(P) = O$.

10. Let f be a function on an open set U in 3-space. Let g be another function, and let S be the surface consisting of all points X such that

$$g(X) = 0 \qquad \text{but} \qquad \operatorname{grad} g(X) \neq O.$$

Suppose that P is a point of the surface S such that $f(P)$ is a maximum for f on S, that is

$$f(P) \geqq f(X) \qquad \text{for all } X \text{ on } S.$$

Prove that there is a number λ such that

$$\operatorname{grad} f(P) = \lambda \operatorname{grad} g(P).$$

11. Let $f : \mathbf{R}^2 \to \mathbf{R}$ be the function such that $f(0, 0) = 0$ and

$$f(x, y) = \frac{x^3}{x^2 + y^2} \quad \text{if} \quad (x, y) \neq (0, 0).$$

Show that f is not differentiable at $(0, 0)$. However, show that for any differentiable curve $\varphi: J \to \mathbf{R}^2$ passing through the origin, $f \circ \varphi$ is differentiable.

XV, §3. POTENTIAL FUNCTIONS

Let U be an open set in $\mathbf{R}^n$. By a **continuous path** in U we shall mean a continuous map $\alpha: J \to U$ from some closed interval $J = [a, b]$ into U. By a **piecewise continuous path** in U we shall mean a finite sequence

$$\{\alpha_1, \ldots, \alpha_r\}$$

of continuous paths, defined on closed intervals $J_1, \ldots, J_r$ such that if $J_i = [a_i, b_i]$ then

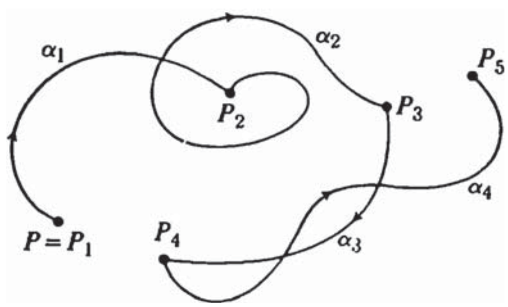
$$\alpha_{i+1}(a_{i+1}) = \alpha_i(b_i).$$

We call $\alpha_i(a_i)$ the beginning point of α_i and $\alpha_i(b_i)$ the end point of α_i. We call $\alpha_1(a_1)$ the **beginning point** of the path, and $\alpha_r(b_r)$ its **end point**. We often use a short symbol like γ to denote a path. We say that the path

$$\gamma = \{\alpha_1, \ldots, \alpha_r\}$$

is piecewise C^1 if each α_i has a continuous derivative. *For the rest of this chapter, by a path we shall mean a piecewise C^1 path.*

A path looks like this:



We say that an open set U is **connected** if given two points P, Q in the set, there exists a path in U whose beginning point is P and whose end point is Q.

Theorem 3.1. *Let U be an open set in $\mathbf{R}^n$ and assume that U is connected. Let f, g be two differentiable functions on U. If* $\operatorname{grad} f = \operatorname{grad} g$ *on U, then there exists a constant C such that*

$$f = g + C.$$

Proof. We note that $\operatorname{grad}(f - g) = \operatorname{grad} f - \operatorname{grad} g = 0$, so it will suffice to prove that if ψ is a differentiable function on U with $\operatorname{grad} \psi = 0$ then ψ is constant.

Let P, Q be any two points of U, and let $\{\alpha_1, \ldots, \alpha_r\}$ be a path between P and Q, that is P is its beginning point and Q is its end point. Then for each i,

$$(\psi \circ \alpha_i)'(t) = \operatorname{grad} \psi(\alpha_i(t)) \cdot \alpha_i'(t) = 0.$$

Hence $\psi \circ \alpha_i$ is constant on its interval of definition. In particular, let P_i be the beginning point of α_i. If α_1 is defined on $[a_1, b_1]$ then

$$\psi(P_1) = \psi(\alpha_1(a_1)) = \psi(\alpha_1(b_1)) = \psi(P_2),$$

By induction, we obtain

$$\psi(P_1) = \psi(P_2) = \cdots = \psi(P_{r+1}),$$

thereby proving the theorem.

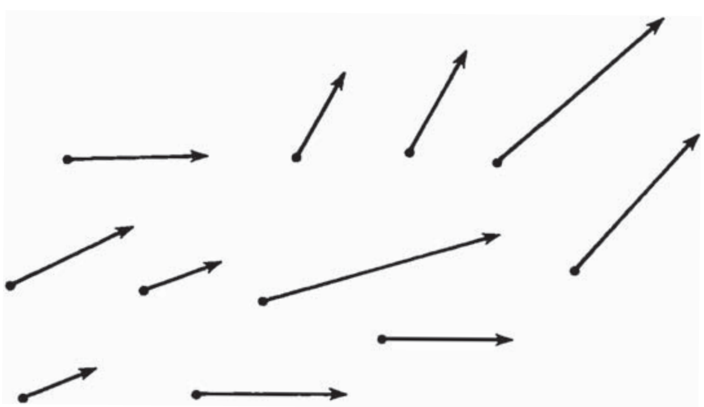
Again let U be an open set in $\mathbf{R}^n$. A **vector field** on U is a map

$$F: U \to \mathbf{R}^n$$

(which therefore associates with each point of U an element of $\mathbf{R}^n$). The map F is represented by coordinate functions, $F = (f_1, \dots, f_n)$. We say that F is continuous (resp. differentiable) if each f_i is continuous (resp. differentiable).

Example. Let $F(x, y) = (x^2y, \sin xy)$. Then F is a vector field which to the point (x, y) associates $(x^2y, \sin xy)$, having the same number of coordinates, namely two of them in this case.

A vector field in physics is often interpreted as a field of forces. A vector field may be visualized as a field of arrows, which to each point associates the arrow as shown on the figure. Each arrow points in the direction of



the force, and the length of the arrow represents the magnitude of the force.

If f is a differentiable function on U, then we observe that $\operatorname{grad} f$ is a vector field, which associates the vector $\operatorname{grad} f(P)$ to the point in U.

If F is a vector field and if there exists a differentiable function φ such that $F = -\operatorname{grad} \varphi$, then φ is called the **potential energy** of the vector field, and F is called **conservative**, for the following reason. Suppose that a particle of mass m moves along a differentiable curve $\alpha(t)$ in U, and let us assume that this particle obeys **Newton's law**:

$$F(\alpha(t)) = m\alpha''(t)$$

for all t where $\alpha(t)$ is defined. In other words, force equals mass times acceleration. Physicists define the **kinetic energy** to be

$$\tfrac{1}{2}m\alpha'(t)^2 = \tfrac{1}{2}mv(t)^2$$

where $v(t)$ is the speed (norm of the velocity).

Conservation law. *If $F = -\operatorname{grad}\varphi$, then the sum of the potential energy φ and the kinetic energy is constant.*

Proof. We have to prove that

$$\varphi(\alpha(t)) + \tfrac{1}{2}m\alpha'(t)^2$$

is constant. To see this, we differentiate this sum. By the chain rule, we see that its derivative is equal to

$$\operatorname{grad}\varphi(\alpha(t))\cdot\alpha'(t) + m\alpha'(t)\cdot\alpha''(t).$$

By Newton's law, $m\alpha''(t) = F(\alpha(t)) = -\operatorname{grad}\varphi(\alpha(t))$. Hence this derivative is equal to 0. This proves what we wanted.

It is not true that all vector fields are conservative. We shall discuss below the problem of determining which ones are conservative. The fields of classical physics are conservative.

Example. Consider a force $F(X)$ which is inversely proportional to the square of the distance from the point X to the origin, and in the direction of X. Then there is a constant k such that for $X \neq 0$ we have

$$F(X) = k\frac{1}{|X|^2}\frac{X}{|X|}$$

because $X/|X|$ is the unit vector in the direction of X. Thus

$$F(X) = k\frac{1}{r^3}X,$$

where $r = |X|$. A potential energy for F is given by

$$\varphi(X) = \frac{k}{r}.$$

This is immediately verified by taking the partial derivatives of this function. Cf. Exercise 3 of the preceding section, and Exercises 1 and 2 below.

By a **potential function** for a vector field F we shall mean a differentiable function φ such that $F = \text{grad}\, \varphi$. Thus a potential function is equal to *minus* the potential energy (if it exists). Theorem 3.1 shows that a potential function is uniquely determined up to a constant if U is connected.

Example. When a vector field comes from a single source of whatever (heat, electricity, etc.) located at a point, then the potential function $\varphi(X)$ depends only on the distance from this point. Suppose this point is the origin. Then there exists a function g of one variable such that

$$\varphi(X) = g(r) \qquad \text{where} \qquad r = |X| = \sqrt{X \cdot X}.$$

Conversely, suppose given a function g of one variable r, defined for $r > 0$, and of class C^1. Define $\varphi(X) = g(|X|) = g(r)$. Then in the preceding section you saw that

$$(\text{grad}\, \varphi)(X) = \frac{g'(r)}{r} X.$$

Now do Exercise 2.

From Theorem 1.1, we are able to deduce a criterion for the existence of a potential function.

Theorem 3.2. *Let $F = (f_1, \ldots, f_n)$ be a C^1 vector field on an open set U of $\mathbf{R}^n$. (That is, each f_i has continuous partial derivatives.) If F has a potential function, then*

$$D_i f_j = D_j f_i$$

for every $i, j = 1, \ldots, n$.

Proof. This is an immediate corollary of Theorem 1.1. Indeed, if φ is a potential function for F, then $f_i = D_i \varphi$. Hence

$$D_j f_i = D_j D_i \varphi = D_i D_j \varphi = D_i f_j,$$

as was to be shown.

Example. We conclude that if, say in two variables, we have a vector field F with $F(x, y) = (f(x, y), g(x, y))$ such that f, g have continuous partials, and $\partial f/\partial y \neq \partial g/\partial x$, then the vector field does *not* have a potential function. For instance, the vector field

$$F(x, y) = (x^2 y, x + y^3)$$

does not have a potential function. In this case, $f(x, y) = x^2y$ and

$$g(x, y) = x + y^3,$$

and $\partial f/\partial y = x^2$ while $\partial g/\partial x = 1$.

For the converse of Theorem 3.2, in general, we need some condition on the open set U. However, in many special cases, we can find a potential function by ordinary integration. The most important case is the following.

Theorem 3.3. *Let $a < b$ and $c < d$ be numbers. Let F be a C^1 vector field on the rectangle of all points (x, y) with $a < x < b$ and $c < y < d$. Assume that $F = (f, g)$ with coordinate functions f, g such that*

$$D_2 f = D_1 g.$$

Then F has a potential function on the rectangle.

Proof. Let (x_0, y_0) be a point of the rectangle. Define

$$\varphi(x, y) = \int_{x_0}^{x} f(t, y)\, dt + \int_{y_0}^{y} g(x_0, u)\, du.$$

Then the second integral on the right does not depend on the variable x. Consequently we have

$$D_1\varphi(x, y) = f(x, y)$$

by the fundamental theorem of calculus. On the other hand, by Theorem 7.1 of Chapter X, we can differentiate under the first integral sign, and obtain

$$\begin{aligned} D_2\varphi(x, y) &= \int_{x_0}^{x} D_2 f(t, y)\, dt + g(x_0, y) \\ &= \int_{x_0}^{x} D_1 g(t, y)\, dt + g(x_0, y) \\ &= g(x, y) - g(x_0, y) + g(x_0, y) \\ &= g(x, y), \end{aligned}$$

as was to be shown.

The theorem generalizes to n variables as follows.

Theorem 3.4. *Let F be a C^1 vector field defined on a rectangular box $a_i < x_i < b_i$ for $i = 1, \ldots, n$. Let*

$$F = (f_1, \ldots, f_n)$$

be its coordinates, and assume that $D_i f_j = D_j f_i$ for all pairs of indices i, j. Then F has a potential function.

Proof. Exercise, following the same pattern as in Theorem 3.3. For example, if $n = 3$, one defines

$$\varphi(x, y, z) = \int_{x_0}^{x} f_1(t, y, z)\, dt + \int_{y_0}^{y} f_2(x_0, t, z)\, dt + \int_{z_0}^{z} f_3(x_0, y_0, t)\, dt,$$

where (x_0, y_0, z_0) is a fixed point in the rectangular box. The same technique of differentiating under the integral sign shows that φ is a potential function for F.

XV, §3. EXERCISES

1. Let $X = (x_1, \ldots, x_n)$ denote a vector in $\mathbf{R}^n$. Let $|X|$ denote the euclidean norm. Find a potential function for the vector field F defined for all $X \neq O$ by the formula

$$F(X) = r^k X$$

where $r = |X|$. (Treat separately the cases $k = -2$, and $k \neq -2$.)

2. Again let $r = |X|$. Let g be a differentiable function of one variable. Show that the vector field defined by

$$F(X) = \frac{g'(r)}{r} X$$

on the open set of all $X \neq O$ has a potential function, and determine this potential function.

3. Let

$$G(x, y) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2}\right).$$

This vector field is defined on the plane $\mathbf{R}^2$ from which the origin has been deleted.
(a) For this vector field $G = (f, g)$ show that $D_2 f = D_1 g$.
(b) Why does this vector field have a potential function on every rectangle not containing the origin?
(c) Verify that the function $\psi(x, y) = -\arctan x/y$ is a potential function for G on any rectangle not intersecting the line $y = 0$.

(d) Verify that the function $\psi(x, y) = \arccos x/r$ is a potential function for this vector field in the upper half plane.

In the next section you will see that this vector field does not admit a potential function on the whole plane from which the origin has been deleted.

XV, §4. CURVE INTEGRALS

Let $F = (f, g)$ be a vector field such that $D_2 f = D_1 g$.

On more general domains than rectangles there does not always exist a potential function because the domain does not allow for the simple type of integration which we performed. In Theorem 3.3 we could integrate the function repeatedly without difficulty, with an ordinary integral. We shall now see how to extend this integration, and formulate whatever is true in general.

Let U be an open set in $\mathbf{R}^n$ and let $\alpha: J \to \mathbf{R}^n$ be a C^1 curve (so with continuous derivative) defined on a closed interval J, with say $J = [a, b]$. Assume that α takes its values in U. Let F be a continuous vector field on U. We wish to define the integral of F along α. We **define**

$$\boxed{\int_\alpha F = \int_a^b F(\alpha(t)) \cdot \alpha'(t)\, dt.}$$

Note. $\alpha(t)$ is a point of U, so we can take $F(\alpha(t))$ which is a vector. Dotting with the vector $\alpha'(t)$ yields a number for each t. Thus the expression inside the integral is a function of t, and is continuous, so we can integrate it. If P, Q are the beginning and end points of α respectively, that is

$$P = \alpha(a) \qquad \text{and} \qquad Q = \alpha(b),$$

then we shall also write the integral in the form

$$\int_{P,\alpha}^{Q} F \qquad \text{or} \qquad \int_P^Q F \cdot d\alpha.$$

Example. Let $F(x, y) = (x^2 y, y^3)$. Let α parametrize the straight line between $(0, 0)$ and $(1, 1)$, so $\alpha(t) = (t, t)$ for $0 \leqq t \leqq 1$. To find the integral of F along α from the origin to $(1, 1)$, we have $F(\alpha(t)) = (t^3, t^3)$ and $\alpha'(t) = (1, 1)$. Hence

$$F(\alpha(t)) \cdot \alpha'(t) = 2t^3.$$

Hence

$$\int_\alpha F = \int_0^1 2t^3\,dt = \tfrac{1}{2}.$$

Remark. Occasionally one commits an abuse of language in speaking of the integral of a vector field along a path. For instance, let

$$F(x, y) = (y^2, -x)$$

be a vector field in the plane $\mathbf{R}^2$. We wish to find the integral of F along the parabola $x = y^2$, from (0, 0) to (1, 2). Strictly speaking, this is a meaningless statement since the parabola is not given in parametric form by a map from an interval into the plane. However, in such cases, we usually mean to take the integral along some naturally selected path whose set of points is the given portion of the curve between (0, 0) and (1, 2). In this case, we would take the path defined by

$$\alpha(t) = (t^2, t),$$

which parametrizes the parabola, between $t = 0$ and $t = 1$. Thus the desired integral is equal to

$$\int_0^1 (t^2, -t^2)\cdot(2t, 1)\,dt = \int_0^1 (2t^3 - t^2)\,dt = \tfrac{1}{6}.$$

The **straight line segment** between two points P and Q is usually parametrized by

$$\alpha(t) = P + t(Q - P) \qquad \text{with} \quad 0 \leqq t \leqq 1.$$

The **circle** of radius $a > 0$ around the origin is parametrized by

$$\beta(t) = (a\cos t, a\sin t).$$

The integral along a curve is independent of the parametrization. This is essentially proved in the next theorem.

Theorem 4.1. *Let $J_1 = [a_1, b_1]$ and $J_2 = [a_2, b_2]$ be two intervals, and let $g\colon J_1 \to J_2$ be a C^1 map such that $g(a_1) = a_2$ and $g(b_1) = b_2$. Let $\alpha\colon J_2 \to U$ be a C^1 path into an open set U of $\mathbf{R}^n$. Let F be a continuous vector field on U. Then*

$$\int_\alpha F = \int_{\alpha\circ g} F.$$

Proof. This is nothing more than the chain rule. By definition,

$$\begin{aligned}\int_{\alpha\circ g} F &= \int_{a_1}^{b_1} F(\alpha(g(t)))\cdot\frac{d\alpha(g(t))}{dt}\,dt\\ &= \int_{g(a_1)}^{g(b_1)} F(\alpha(u))\cdot\frac{d\alpha(u)}{du}\,du\\ &= \int_{a_2}^{b_2} F(\alpha(u))\cdot\alpha'(u)\,du = \int_\alpha F.\end{aligned}$$

This proves our theorem.

Suppose the vector field is on $\mathbf{R}^2$, say

$$F(x, y) = (f(x, y), g(x, y)).$$

Then one denotes the integral of F along a curve α formally by the expression

$$\boxed{\int_\alpha F = \int_\alpha f\,dx + g\,dy.}$$

The curve α can be represented by coordinates,

$$\alpha(t) = (x(t), y(t)) \qquad \text{with } a \leqq t \leqq b,$$

and therefore in terms of the parameter t the integral is given by

$$\int_\alpha F = \int_a^b \left[f\frac{dx}{dt} + g\frac{dy}{dt}\right] dt.$$

Note that the expression inside the integral sign is precisely the dot product:

$$f\frac{dx}{dt} + g\frac{dy}{dt} = F(\alpha(t))\cdot\frac{d\alpha}{dt}.$$

Example. Let us go back to the vector field $F(x, y) = (x^2y, y^3)$ to be integrated along the line segment between $(0, 0)$ and $(1, 1)$. Then we can write the integral in the form

$$\begin{aligned}\int_\alpha F &= \int_\alpha x^2y\,dx + y^3\,dy \qquad [\text{with } x = t, y = t]\\ &= \int_0^1 t^3\,dt + t^3\,dt\\ &= \tfrac{1}{2}.\end{aligned}$$

Example. We want to find the integral of the vector field

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

around the circle of radius 2, counterclockwise. We parametrize the circle by $x = 2 \cos \theta$, $y = 2 \sin \theta$, so the integral is equal to

$$\begin{aligned}\int_C F &= \int_C \frac{x}{x^2 + y^2}\, dx + \frac{y}{x^2 + y^2}\, dy \\ &= \int_0^{2\pi} \cos \theta(-\sin \theta)\, d\theta + \sin \theta(\cos \theta)\, d\theta \\ &= 0.\end{aligned}$$

If $\alpha = \{\alpha_1, \ldots, \alpha_r\}$ is a path such that each α_i is C^1, we define

$$\int_\alpha F = \int_{\alpha_1} F + \cdots + \int_{\alpha_r} F$$

to be the sum of the integrals of F taken over each α_i, $i = 1, \ldots, r$.

We shall say that the path α is **closed** if its beginning point is equal to its end point. The next theorem is concerned with closed paths, and with the dependence of an integral on the path between two points. For this we make a remark.

Let $\alpha: J \to U$ be a C^1 path between two points of U. Say α is defined on $J = [a, b]$ and $P = \alpha(a)$, $Q = \alpha(b)$. We can define a path going in reverse direction by letting

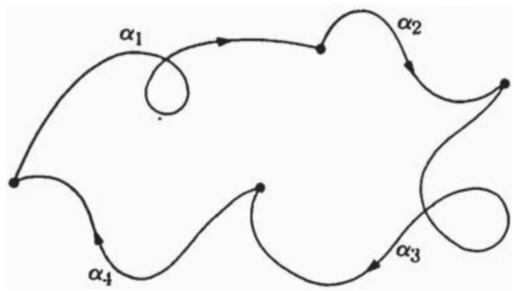
$$\alpha^-(t) = \alpha(a + b - t).$$

When $t = a$ we have $\alpha^-(a) = \alpha(b)$, and when $t = b$ we have $\alpha^-(b) = \alpha(a)$. Also, α^- is defined on the interval $[a, b]$. A simple change of variables in the integral shows that

$$\int_{\alpha^-} F = -\int_\alpha F.$$

We leave this to the reader. We call α^- the **opposite path** of α, or **inverse path**.

The piecewise C^1 path consisting of the pair $\{\alpha, \alpha^-\}$ is a closed path, which comes back to the beginning point of α. More general closed paths look like this:



Theorem 4.2. *Let U be a connected open set in $\mathbf{R}^n$. Let F be a continuous vector field on U. Then the following conditions are equivalent:*

(1) *F has a potential function on U.*
(2) *The integral of F between any two points of U is independent of the path.*
(3) *The integral of F along any closed path in U is equal to* 0.

Proof. Assume condition (1), and let φ be a potential function for F on U. Let α first be a C^1 path in U defined on an interval $[a, b]$. Then using the chain rule, we find:

$$\begin{aligned}\int_\alpha F = \int_a^b F(\alpha(t)) \cdot \alpha'(t)\, dt &= \int_a^b (\operatorname{grad} \varphi(\alpha(t))) \cdot \alpha'(t)\, dt \\ &= \int_a^b \frac{d}{dt} \varphi(\alpha(t))\, dt \\ &= \varphi(\alpha(b)) - \varphi(\alpha(a)).\end{aligned}$$

Thus if $P = \alpha(a)$ and $Q = \alpha(b)$ are the beginning and end points of α respectively, we find that

$$\boxed{\int_{P,\alpha}^{Q} F = \int_\alpha F = \varphi(Q) - \varphi(P).}$$

From this we conclude first that the integral is independent of the path, and depends only on the values of φ at Q and P.

Now suppose that $\alpha = \{\alpha_1, \alpha_2, \ldots, \alpha_r\}$ is a piecewise C^1 path between points $P = P_1$ and $Q = P_{r+1}$, where P_i is the beginning point of α_i (or the end point of α_{i-1}). By definition and by what we have just seen, we find that

$$\begin{aligned}\int_{P,\alpha}^{Q} F &= \int_{\alpha_1} F + \cdots + \int_{\alpha_r} F \\ &= \varphi(P_2) - \varphi(P_1) + \varphi(P_3) - \varphi(P_2) + \cdots + \varphi(P_{r+1}) - \varphi(P_r) \\ &= \varphi(P_{r+1}) - \varphi(P_1) = \varphi(Q) - \varphi(P).\end{aligned}$$

Hence the same result holds in the general case.

In particular, if α is a closed path, then $P = Q$ and we find

$$\int_{P,\alpha}^{Q} F = \varphi(P) - \varphi(P) = 0.$$

Thus we have shown that condition (1) implies both (2) and (3).

It is obvious that (2) implies (3). Conversely, assume that the integral of F along any closed path is equal to 0. We shall prove (2). Intuitively, given two points P, Q and two paths α, β from P to Q, we go from P to Q along α, and back along the inverse of β. The integral must be equal to 0. To see this formally, let β^- be the path opposite to β. Then Q is the beginning point of β^- and P is its end point. Hence the path $\{\alpha, \beta^-\}$ is a closed path, and by hypothesis,

$$\int_\alpha F + \int_{\beta^-} F = 0.$$

However,

$$\int_\alpha F + \int_{\beta^-} F = \int_\alpha F - \int_\beta F = 0.$$

Hence

$$\int_\alpha F = \int_\beta F$$

thus proving that (3) implies (2).

There remains to prove that if we assume (2), that is if the integral is independent of the path, then F admits a potential function.

Let P be a fixed point of U. It is natural to define for any point Q of U the value

$$\varphi(Q) = \int_P^Q F$$

taken along *any* path α, since this value is independent of the path. We now contend that φ is a potential function for F. To verify this, we must compute the partial derivatives of φ. If

$$F = (f_1, \ldots, f_n)$$

is expressed in terms of its coordinate functions f_i, we must show that

$$D_i\varphi = f_i \qquad \text{for } i = 1, \ldots, n.$$

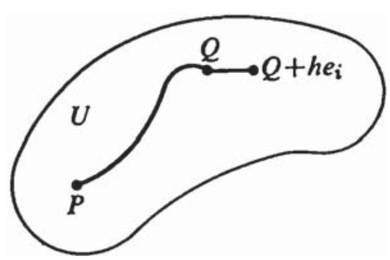
Let $Q = (x_1, \ldots, x_n)$, and let e_i be the i-th unit vector. We must show that

$$\lim_{h\to 0} \frac{\varphi(Q + he_i) - \varphi(Q)}{h} = f_i(Q).$$

We have

$$\begin{aligned}\varphi(Q + he_i) - \varphi(Q) &= \int_P^{Q+he_i} F - \int_P^Q F \\ &= \int_Q^{Q+he_i} F\end{aligned}$$

where the integrals are taken along any path. Since they are independent of the path, we do not specify a path in the notation. Now the integral between Q and $Q + he_i$ will be taken along the most natural path, namely the straight line segment between Q and $Q + he_i$.



Since U is open, taking h sufficiently small, we know that this line segment lies in U. Thus we select the path α such that $\alpha(t) = Q + the_i$ with

$$0 \leqq t \leqq 1.$$

Then $\alpha(0) = Q$ and $\alpha(1) = Q + he_i$. Furthermore, $\alpha'(t) = he_i$. We find:

$$\frac{\varphi(Q + he_i) - \varphi(Q)}{h} = \frac{1}{h}\int_0^1 F(Q + the_i) \cdot he_i \, dt.$$

But for any vector $v \in \mathbf{R}^n$ we have $F(v) \cdot e_i = f_i(v)$. Consequently our expression is equal to

$$\frac{1}{h}\int_0^1 f_i(Q + the_i)h\,dt.$$

We change variables, letting $u = ht$ and $du = h\,dt$. We find:

$$\frac{1}{h}\int_0^h f_i(Q + ue_i)\,du.$$

Let $g(u) = f_i(Q + ue_i)$ and let G be an indefinite integral for g, so that $G' = g$. Then

$$\frac{1}{h}\int_0^h f_i(Q + ue_i)\,du = \frac{G(h) - G(0)}{h}.$$

Taking the limit as $h \to 0$, we obtain $G'(0) = g(0) = f_i(Q)$, thus showing that the i-th partial derivative of φ exists and is equal to f_i. This concludes the proof of our theorem.

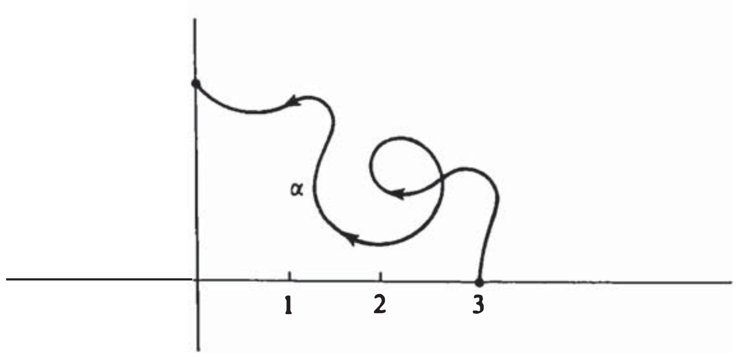
Example. The theorem allows us to show that the vector field

$$G(x, y) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2}\right)$$

defined on the plane $\mathbf{R}^2$ from which the origin is deleted does not have a potential function. Indeed, if you integrate this vector field around a circle centered at the origin, you will find 2π (see Exercise 4). This circle is a closed path, so there cannot be a potential function.

On the other hand, this vector field has a potential function on the upper half plane as given in Exercise 3 of the preceding section. Therefore the integral of G along the path shown on the figure is easy to determine:

$$\int_\alpha G = \arccos 0 - \arccos 1 = \frac{\pi}{2}.$$



Observe, however, that some vector fields are defined on the same domain, but do admit potential functions, for instance any vector field of the form

$$F(x, y) = \frac{g'(r)}{r}(x, y),$$

where $r = \sqrt{x^2 + y^2}$ and g is a differentiable function of one variable. The potential function is $g(r) = f(x, y)$, which you can check by direct differentiation in computing its gradient.

XV, §4. EXERCISES

Compute the curve integrals of the vector field over the indicated curves.

1. $F(x, y) = (x^2 - 2xy, y^2 - 2xy)$ along the parabola $y = x^2$ from $(-2, 4)$ to $(1, 1)$.

2. $(x, y, xz - y)$ over the line segment from $(0, 0, 0)$ to $(1, 2, 4)$.

3. (x^2y^2, xy^2) along the closed path formed by parts of the line $x = 1$ and the parabola $y^2 = x$, counterclockwise.

4. Let

$$G(x, y) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2}\right).$$

(a) Find the integral of this vector field counterclockwise along the circle $x^2 + y^2 = 2$ from $(1, 1)$ to $(-\sqrt{2}, 0)$.
(b) Counterclockwise around the whole circle.
(c) Counterclockwise around the circle $x^2 + y^2 = a^2$ for $a > 0$.

5. Let $r = (x^2 + y^2)^{1/2}$ and $F(X) = r^{-1}X$ for $X = (x, y)$. Find the integral of F over the circle of radius 2, centered at the origin, taken in the counterclockwise direction.

6. Let C be a circle of radius 20 with center at the origin. Let $F(X)$ be a vector field on $\mathbf{R}^2$ such that $F(X)$ has the same direction as X (that is there exists a differentiable function $g(X)$ such that $F(X) = g(X)X$, and $g(X) > 0$ for all X). What is the integral of F around C, taken counterclockwise?

7. Let P, Q be points in 3-spaces. Show that the integral of the vector field given by

$$F(x, y, z) = (z^2, 2y, 2xz)$$

from P to Q is independent of the curve selected between P and Q.

8. Let $F(x, y) = (x/r^3, y/r^3)$ where $r = (x^2 + y^2)^{1/2}$. Find the integral of F along the curve

$$\alpha(t) = (e^t \cos t, e^t \sin t)$$

from the point $(1, 0)$ to the point $(e^{2\pi}, 0)$.

9. Let $F(x, y) = (x^2y, xy^2)$.
   (a) Does this vector field admit a potential function?
   (b) Compute the integral of this vector field from (0, 0) to the point

$$P = (1/\sqrt{2}, 1/\sqrt{2})$$

   along the line segment from (0, 0) to P.
   (c) Compute the integral of this vector field from (0, 0) to P along the path which consists of the segment from (0, 0) to (1, 0), and the arc of circle from (1, 0) to P. Compare with the value found in (b).

10. Let

$$F(x, y) = \left(\frac{x \cos r}{r}, \frac{y \cos r}{r}\right),$$

   where $r = \sqrt{x^2 + y^2}$. Find the value of the integral of this vector field:
   (a) Counterclockwise along the circle of radius 1, from (1, 0) to (0, 1).
   (b) Counterclockwise around the entire circle.
   (c) Does this vector field admit a potential function? Why?

11. Let

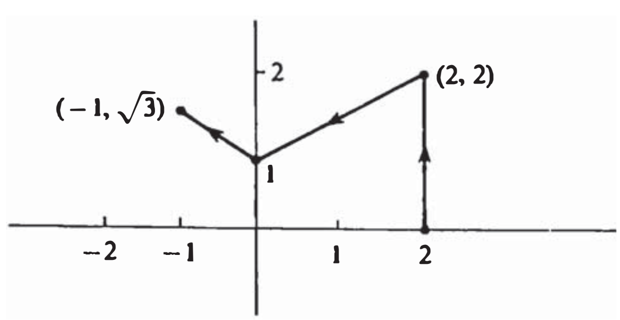
$$F(x, y) = \left(\frac{xe^r}{r}, \frac{ye^r}{r}\right).$$

   Find the value of the integral of this vector field:
   (a) Counterclockwise along the circle of radius 1 centered at the origin.
   (b) Counterclockwise along the circle of radius 5 centered at the point $(14, -17)$.
   (c) Does this vector field admit a potential function? Why?

12. Let

$$G(x, y) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2}\right).$$

   (a) Find the integral of G along the line $x + y = 1$ from (0, 1) to (1, 0).
   (b) From the point (2, 0) to the point $(-1, \sqrt{3})$ along the path shown on the figure.



13. Let F be a smooth vector field on $\mathbf{R}^2$ from which the origin has been deleted, so F is not defined at the origin. Let $F = (f, g)$. Assume that $D_2 f = D_1 g$ and let

$$k = \frac{1}{2\pi} \int_C F.$$

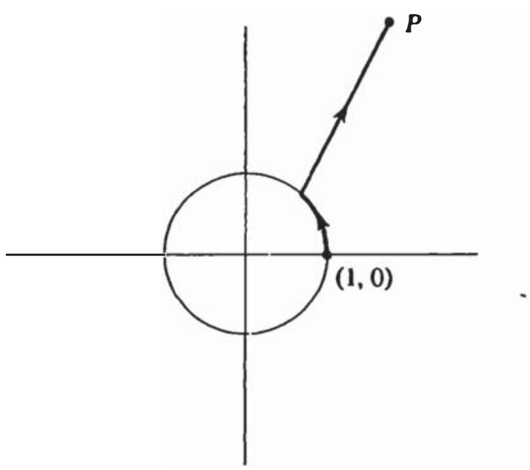
where C is the circle of radius 1 centered at the origin. Let G be the vector field

$$G(x, y) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right).$$

Show that there exists a function φ defined on $\mathbf{R}^2$ from which the origin has been deleted such that

$$F = \operatorname{grad} \varphi + kG$$

[*Hint*: Follow the same method as in the proof of Theorem 4.2 in the text, but define $\varphi(P)$ by integrating $F - kG$ from (1,0) to P as shown on the figure.]



XV, §5. TAYLOR'S FORMULA

Let f be a function on an open set U of $\mathbf{R}^n$. We may take iterated partial derivatives (if they exist) of the form

$$D_1^{i_1} \cdots D_n^{i_n} f$$

where $i_1, \ldots, i_n$ are integers $\geqq 0$. It does not matter in which order we take the partials (provided they exist and are continuous) according to Theorem 1.1.

If $c_{i_0 \cdots i_n}$ are numbers, we may form finite sums

$$\sum c_{i_1 \cdots i_n} D_1^{i_1} \cdots D_n^{i_n}$$

which we view as applicable to functions which have enough partial derivatives. More precisely, we say that a function f on U is of **class C^p** (for some integer $p \geqq 0$) if all partial derivatives

$$D_1^{i_1} \cdots D_n^{i_n} f$$

exist for $i_1 + \cdots + i_n \leqq p$ and are continuous. It is clear that the functions of class C^p form a vector space. Let $i_1, \ldots, i_n$ be integers $\geqq 0$ such that $i_1 + \cdots + i_n = r \leqq p$. Let F_p be the vector space of functions of class C^p. (For $p = 0$, this is the vector space of continuous functions on U.) Then any monomial $D_1^{i_1} \cdots D_n^{i_n}$ may be viewed as a linear map $F_p \to F_{p-r}$ given by

$$f \mapsto D_1^{i_1} \cdots D_n^{i_n} f.$$

We say that f is of **class C^∞** if it is of class C^p for every positive integer p. If f is of class C^∞, then $D_1^{i_1} \cdots D_n^{i_n} f$ is also of class C^∞. We can take the sum of linear maps in the usual way, and thus

$$(\sum c_{i_1 \cdots i_n} D_1^{i_1} \cdots D_n^{i_n}) f = \sum c_{i_1 \cdots i_n} D_1^{i_1} \cdots D_n^{i_n} f,$$

if the sum is taken over all n-tuples of integers $(i_1, \ldots, i_n)$ such that

$$i_1 + \cdots + i_n \leqq r.$$

A linear map such as the above, expressed as a sum of monomials of partial derivatives with constant coefficients, will be called a **partial differential operator with constant coefficients**.

We multiply such operators in the obvious way using distributivity. For example,

$$(D_1 + D_2)^2 = D_1^2 + 2D_1D_2 + D_2^2.$$

In terms of two variables (x, y), say, we write this also in the form

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

Similarly, we write in terms of n variables $x_1, \ldots, x_n$:

$$D_1^{i_1} \cdots D_n^{i_n} = \left(\frac{\partial}{\partial x_1}\right)^{i_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{i_n} = \frac{\partial^{i_1 + \cdots + i_n}}{\partial x_1^{i_1} \cdots \partial x_n^{i_n}}.$$

In Taylor's formula, we shall use especially the expansion

$$(h_1D_1 + \cdots + h_nD_n)^r = \sum c_{i_1 \cdots i_n} h_1^{i_1} \cdots h_n^{i_n} D_1^{i_1} \cdots D_n^{i_n}$$

if $h_1, \ldots, h_n$ are numbers. In the special case where $n = 2$ we have

$$(hD_1 + kD_2)^r = \sum_{i=0}^{r} \binom{r}{i} h^i k^{r-i} D_1^i D_2^{r-i}.$$

In the general case, the coefficients are generalizations of the binomial coefficients, which we don't need to write down explicitly.

It will be convenient to use a vector symbol

$$\nabla = (D_1, \ldots, D_n) = \left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\right),$$

its form in terms of $x_1, \ldots, x_n$ being used when the variables are called $x_1, \ldots, x_n$. If $H = (h_1, \ldots, h_n)$ is an n-tuple of numbers, then we agree to let

$$H \cdot \nabla = h_1 D_1 + \cdots + h_n D_n.$$

We view $H \cdot \nabla$ as a linear map, applicable to functions. Observe that

$$(H \cdot \nabla) f = h_1 D_1 f + \cdots + h_n D_n f = H \cdot \operatorname{grad} f,$$

or in terms of a vector $X = (x_1, \ldots, x_n)$,

$$\begin{aligned}(H \cdot \nabla) f(X) &= h_1 D_1 f(X) + \cdots + h_n D_n f(X) \\ &= H \cdot (\operatorname{grad} f)(X).\end{aligned}$$

This last dot product is the old dot product between the vectors H and $\operatorname{grad} f(X)$. Of course one should write $((H \cdot \nabla) f)(X)$, but as usual we omit the extra parentheses.

This notation will be useful in the following application. Let f be a C^1 function on an open set U in $\mathbf{R}^n$. Let $P \in U$, and let H be a vector. For some open interval of values of t, the vectors $P + tH$ lie in U. Consider the function g of t defined by

$$g(t) = f(P + tH).$$

By a trivial application of the chain rule, we find that

$$\begin{aligned}\frac{dg(t)}{dt} = g'(t) &= \operatorname{grad} f(P + tH) \cdot H \\ &= h_1 D_1 f(P + tH) + \cdots + h_n D_n f(P + tH) \\ &= (H \cdot \nabla) f(P + tH).\end{aligned}$$

We can generalize this to higher derivatives:

Theorem 5.1. *Let r be a positive integer. Let f be a function of class C^r on an open set U in n-space. Let $P \in U$. Let H be a vector. Then*

$$\left(\frac{d}{dt}\right)^r (f(P + tH)) = (H \cdot \nabla)^r f(P + tH).$$

Proof. For $r = 1$ we have just proved our formula. By induction, assume it proved for $1 \leqq k < r$. Let $\varphi = (H \cdot \nabla)^k f$ and apply the derivative d/dt to the function $t \mapsto \varphi(P + tH)$. By the case $k = 1$, we find

$$\frac{d}{dt}(\varphi(P + tH)) = (H \cdot \nabla)\varphi(P + tH).$$

Substituting $\varphi = (H \cdot \nabla)^k f$, we find that this expression is equal to

$$(H \cdot \nabla)^{k+1} f(P + tH),$$

as was to be proved.

Taylor's formula. *Let f be a C^r function on an open set U of $\mathbf{R}^n$. Let $P \in U$ and let H be a vector. Assume that the line segment*

$$P + tH, \qquad 0 \leqq t \leqq 1,$$

is contained in U. Then there exists a number τ between 0 and 1 such that

$$f(P + H) = f(P) + \frac{(H \cdot \nabla)f(P)}{1!} + \cdots + \frac{(H \cdot \nabla)^{r-1} f(P)}{(r-1)!} + \frac{(H \cdot \nabla)^r f(P + \tau H)}{r!}.$$

Proof. Let $g(t) = f(P + tH)$. Then g is differentiable as a function of t in the sense of functions of one variable, and we can apply the ordinary Taylor formula to g and its derivatives between $t = 0$ and $t = 1$. In that case, all powers of $(1 - 0)$ are equal to 1. Hence Taylor's formula in one variable applied to g yields

$$g(1) = g(0) + \frac{g'(0)}{1!} + \cdots + \frac{g^{(r-1)}(0)}{(r-1)!} + \frac{g^{(r)}(\tau)}{r!}$$

for some number τ between 0 and 1. The successive derivatives of g are given by Theorem 5.1. If we evaluate them for $t = 0$ in the terms up to order $r - 1$ and for $t = \tau$ in the r-th term, then we see that the Taylor formula for f simply drops out.

Estimate for Taylor's formula. *Let the remainder term be*

$$R(H) = \frac{(H \cdot \nabla)^r f(P + \tau H)}{r!}$$

for $0 \leqq \tau \leqq 1$. *Let* C *be a bound for all partial derivatives of* f *on* U *of order* $\leqq r$. *Then there exists a number* K *depending only on* r *and* n *such that*

$$|R(H)| \leqq \frac{CK}{r!} |H|^r.$$

Proof. If we expand out $(H \cdot \nabla)^r$, we obtain a sum

$$\sum c_{i_1 \cdots i_n} h_1^{i_1} \cdots h_n^{i_n} D_1^{i_1} \cdots D_n^{i_n}$$

where the $c_{(i)}$ are fixed numbers coming from generalized multinomial coefficients depending only on r and n, and the exponents satisfy

$$i_1 + \cdots + i_n = r.$$

The estimate is then obvious, since each term can be estimated as indicated, and the number of terms in the sum depends only on r and n.

Using another notation, we obtain

$$\begin{aligned} f(X) &= f(O) + D_1 f(O)x_1 + \cdots + D_n f(O)x_n + \cdots + f_{r-1}(X) + R_r(X) \\ &= f(O) + f_1(X) + \cdots + f_{r-1}(X) + R_r(X) \end{aligned}$$

where $f_1, \ldots, f_{r-1}$ are homogeneous polynomials of degrees $1, \ldots, r-1$, respectively, and R_r is a remainder term which we can write as

$$|R_r(X)| = O(|X|^r) \qquad \text{for} \quad |X| \to 0.$$

The sum

$$f_0(X) + \cdots + f_{r-1}(X)$$

is the **Taylor polynomial** in several variables of total degree $\leqq r - 1$.

XV, §5. EXERCISES

1. Let f be a differentiable function defined for all of $\mathbf{R}^n$. Assume that $f(O) = 0$ and that $f(tX) = tf(X)$ for all numbers t and vectors $X = (x_1, \ldots, x_n)$. Show that for all $X \in \mathbf{R}^n$ we have $f(X) = \operatorname{grad} f(O) \cdot X$.

2. Let f be a function with continuous partial derivatives of order $\leqq 2$, that is of class C^2 on $\mathbf{R}^n$. Assume that $f(O) = 0$ and $f(tX) = t^2 f(X)$ for all numbers t and all vectors X. Show that for all X we have

$$f(X) = \frac{(X \cdot \nabla)^2 f(O)}{2}.$$

3. Let f be a function defined on an open ball centered at the origin in $\mathbf{R}^n$ and assume that f is of class C^∞. Show that one can write

$$f(X) = f(O) + g_1(X)x_1 + \cdots + g_n(X)x_n$$

where $g_1, \ldots, g_n$ are functions of class C^∞. [*Hint*: Use the fact that

$$f(X) - f(O) = \int_0^1 \frac{d}{dt} f(tX)\, dt.]$$

4. Let f be a C^∞ function defined on an open ball centered at the origin in $\mathbf{R}^n$. Show that one can write

$$f(X) = f(O) + \operatorname{grad} f(O) \cdot X + \sum_{i,j} g_{ij}(X) x_i x_j$$

where g_{ij} are C^∞ functions. [*Hint*: Assume first that $f(O) = 0$ and $\operatorname{grad} f(O) = O$. In Exercises 3 and 4, use an integral form for the remainder.]

5. Generalize Exercise 4 near an arbitrary point $A = (a_1, \ldots, a_n)$, expressing

$$f(X) = f(A) + \sum_{i=1}^n D_i f(A)(x_i - a_i) + \sum_{i,j} h_{ij}(X)(x_i - a_i)(x_j - a_j).$$

This expression or that of Exercise 4 is often more useful than the expression of Taylor's formula.

6. Let F_∞ be the set of all C^∞ functions defined on an open ball centered at the origin in $\mathbf{R}^n$. By a **derivation** D of F_∞ into itself, one means a map $D\colon F_\infty \to F_\infty$ satisfying the rules

$$D(f + g) = Df + Dg, \qquad D(cf) = cDf,$$

$$D(fg) = fD(g) + D(f)g$$

for C^∞ functions f, g and constant c. Let $\lambda_1, \ldots, \lambda_n$ be the coordinate functions, that is $\lambda_i(X) = x_i$ for $i = 1, \ldots, n$. Let D be a derivation as above, and let $\psi_i = D(\lambda_i)$. Show that for any C^∞ function f on the ball, we have

$$D(f) = \sum_{i=1}^n \psi_i D_i f$$

where $D_i f$ is the i-th partial derivative of f. [*Hint*: Show first that $D(1) = 0$ and $D(c) = 0$ for every constant c. Then use the representation of Exercise 5.]

7. Let $f(X)$ and $g(X)$ be polynomials in n variables $(x_1, \dots, x_n)$ of degrees $\leqq s - 1$. Assume that there is a number $a > 0$ and a constant C such that

$$|f(X) - g(X)| \leqq C|X|^s$$

for all X such that $|X| \leqq a$. Show that $f = g$. In particular, the polynomial of Taylor's formula is uniquely determined.

8. Let U be open in $\mathbf{R}^n$ and let $f: U \to \mathbf{R}$ be a function of class C^p. Let $g: \mathbf{R} \to \mathbf{R}$ be a function of class C^p. Prove by induction that $g \circ f$ is of class C^p. Furthermore, assume that at a certain point $P \in U$ all partial derivatives

$$D_{i_1} \cdots D_{i_r} f(P) = 0$$

for all choices of $i_1, \dots, i_r$ and $r \leqq k$. In other words, assume that all partials of f up to order k vanish at P. Prove that the same thing is true for $g \circ f$. [*Hint*: Induction.]

XV, §6. MAXIMA AND THE DERIVATIVE

In this section, we assume that the reader knows something about the dimension of vector spaces. Furthermore, if we have a subspace F of $\mathbf{R}^n$ and if we denote by $F^\perp$ the set of all vectors $w \in \mathbf{R}^n$ which are perpendicular to all elements of F, then $F^\perp$ is a subspace, and

$$\dim F + \dim F^\perp = n.$$

In particular, suppose that $\dim F = n - 1$. Then $\dim F^\perp = 1$, and hence $F^\perp$ consists of all scalar multiples of a single vector w, which forms a basis for $F^\perp$.

Let U be an open set of $\mathbf{R}^n$ and let $f: U \to \mathbf{R}$ be a function of class C^1 on U. Let S be the subset of U consisting of all $x \in U$ such that $f(x) = 0$ and $\operatorname{grad} f(x) \neq O$. We call S the **hypersurface** determined by f. The next lemma will follow from the inverse function theorem, proved later.

Lemma 6.1. *Given $x \in S$ and given a vector $w \in \mathbf{R}^n$ perpendicular to $\operatorname{grad} f(x)$, there exists a curve $\alpha: J \to U$ defined on an open interval J containing 0 such that $\alpha(0) = x$, $\alpha(t) \in S$ for all $t \in J$ (so the curve is contained in the hypersurface), and $\alpha'(t) = w$.*

Theorem 6.2. *Let $f: U \to \mathbf{R}$ be a function of class C^1, and let S be the subset of U consisting of all $x \in U$ such that $f(x) = 0$ and $\operatorname{grad} f(x) \neq O$. Let $P \in S$. Let g be a differentiable function on U and assume that P is a maximum for g on S, that is $g(P) \geqq g(x)$ for all $x \in S$. Then there exists* *a number μ such that*

$$\operatorname{grad} g(P) = \mu \operatorname{grad} f(P).$$

Proof. Let $\alpha: J \to S$ be a differentiable curve defined on an open interval J containing 0 such that $\alpha(0) = P$, and such that the curve is contained in S. We have a maximum at $t = 0$, namely

$$g(\alpha(0)) = g(P) \geqq g(\alpha(t))$$

for all $t \in J$. By an old theorem concerning functions of one variable, we have

$$\begin{aligned} 0 = (g \circ \alpha)'(0) &= \operatorname{grad} g(\alpha(0)) \cdot \alpha'(0) \\ &= \operatorname{grad} g(P) \cdot \alpha'(0). \end{aligned}$$

By the lemma, we conclude that grad $g(P)$ is perpendicular to every vector w which is perpendicular to grad $f(P)$, and hence that there exists a number μ such that

$$\operatorname{grad} g(P) = \mu \operatorname{grad} f(P)$$

since the dimension of the orthogonal space to grad $f(P)$ is equal to $n - 1$. This concludes the proof.

The number μ in Theorem 6.2 is called a **Lagrange multiplier**. We shall give an example how Lagrange multipliers can be used to solve effectively a maximum problem.

Example. Find the minimum of the function

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

subject to the constraint $x^2 + 2y^2 - z^2 - 1 = 0$.

The function is the square of the distance from the origin, and the constraint defines a surface, so at a minimum for f, we are finding the point on the surface which is at minimum distance from the origin. Computing the partial derivatives of the functions f and

$$g(X) = x^2 + 2y^2 - z^2 - 1$$

we find that we must solve the system of equations $g(X) = 0$, and:

(a) $2x = \mu 2x$.
(b) $2y = \mu 4y$.
(c) $2z = \mu(-2z)$.

Let (x_0, y_0, z_0) be a solution. If $z_0 \neq 0$ then from (c) we conclude that $\mu = -1$. The only way to solve (a) and (b) with $\mu = -1$ is that $x = y = 0$. In that case, from the equation $g(X) = 0$ we must have

$$z_0^2 = -1,$$

which is impossible. Hence any solution must have $z_0 = 0$.

If $x_0 \neq 0$ then from (a) we conclude that $\mu = 1$. From (b) and (c) we then conclude that $y_0 = z_0 = 0$. From the equation $g(X) = 0$ we must have $x_0 = \pm 1$. In this manner we have obtained two solutions satisfying our conditions, namely

$$(1, 0, 0) \qquad \text{and} \qquad (-1, 0, 0).$$

Similarly if $y_0 \neq 0$, we find two more solutions, namely

$$(0, \sqrt{\tfrac{1}{2}}, 0) \qquad \text{and} \qquad (0, -\sqrt{\tfrac{1}{2}}, 0).$$

These four solutions are therefore the extrema of the function f subject to the constraint g (or on the surface $g = 0$).

If we ask for the minimum of f, then a direct computation of $f(P)$ for P any one of the above four points shows that the two points

$$P = (0, \pm\sqrt{\tfrac{1}{2}}, 0)$$

are the only possible solutions because $1 > \frac{1}{2}$.

Next we give a more theoretical application of the Lagrange multipliers to the minimum of a quadratic form. Let $A = (a_{ij})$ be a symmetric $n \times n$ matrix of real numbers. "Symmetric" means that $a_{ij} = a_{ji}$. If $\langle x, y\rangle$ denotes the ordinary dot product between elements x, y of $\mathbf{R}^n$, then we have $\langle Ax, y\rangle = \langle Ay, x\rangle$. The function

$$f(x) = \langle Ax, x\rangle$$

is called a **quadratic form**. If one expresses x in terms of coordinates $x_1, \ldots, x_n$ then $f(x)$ has the usual shape

$$f(x) = \sum_{i,j=1}^{n} a_{ij}x_i x_j.$$

But this expression in terms of coordinates is not needed for the statement and proof of the next theorem.

A vector $v \in \mathbf{R}^n$, $v \neq 0$ is called an **eigenvector** of A if there exists a number c such that $Av = cv$.

Theorem 6.3. *Let A be a symmetric matrix and let $f(x) = \langle Ax, x\rangle$. Let v be a point of the sphere of radius 1 centered at the origin such that v is a maximum for f, that is*

$$f(v) \geqq f(x) \qquad \textit{for all } x \textit{ on the sphere.}$$

Then v is an eigenvector for A.

Proof. Let α be a differentiable curve passing through v (that is $\alpha(0) = v$) and contained in the sphere S. Using the rules for the derivative of a product, and composition with a linear map, we know that

$$\begin{aligned}\frac{d}{dt} f(\alpha(t)) &= \frac{d}{dt}\langle A\alpha(t), \alpha(t)\rangle \\ &= \langle A\alpha'(t), \alpha(t)\rangle + \langle A\alpha(t), \alpha'(t)\rangle \\ &= 2\langle A\alpha(t), \alpha'(t)\rangle\end{aligned}$$

using the fact that A is symmetric. Since $\alpha(0) = v$ is a maximum for f, we conclude that

$$0 = (f \circ \alpha)'(0) = 2\langle A\alpha(0), \alpha'(0)\rangle = 2\langle Av, \alpha'(0)\rangle.$$

Now by the lemma, we see that Av is perpendicular to $\alpha'(0)$ for every differentiable curve α as above, and hence that $Av = cv$ for some number c. The theorem is proved.

XV, §6. EXERCISES

1. Find the maximum of $6x^2 + 17y^4$ on the subset of $\mathbf{R}^2$ consisting of those points (x, y) such that

$$(x - 1)^3 - y^2 = 0.$$

2. Find the maximum value of $x^2 + xy + y^2 + yz + z^2$ on the sphere of radius 1 centered at the origin.

3. Let f be a differentiable function on an open set U in $\mathbf{R}^n$, and suppose that P is a minimum for f on U, that is $f(P) \leqq f(X)$ for all X in U. Show that all partial derivatives $D_i f(P) = 0$.

4. Let A, B, C be three distinct points in $\mathbf{R}^n$. Let

$$f(X) = (X - A)^2 + (X - B)^2 + (X - C)^2.$$

Find the point where f reaches its minimum and find the minimum value.

5. Find the maximum of the function $f(x, y, z) = xyz$ subject to the constraints $x \geqq 0, y \geqq 0, z \geqq 0$, and $xy + yz + xz = 2$.

6. Find the shortest distance from a point on the ellipse $x^2 + 4y^2 = 4$ to the line $x + y = 4$.

7. Let S be the set of points $(x_1, \ldots, x_n)$ in $\mathbf{R}^n$ such that

$$\sum x_i = 1 \quad \text{and} \quad x_i > 0 \quad \text{for all } i.$$

Show that the maximum of $g(x) = x_1 \cdots x_n$ occurs at $(1/n, \ldots, 1/n)$ and that

$$g(x) \leqq n^{-n} \quad \text{for all } x \in S.$$

[*Hint*: Consider $\log g$.] Use the result to prove that the geometric mean of n positive numbers is less than or equal to the arithmetic mean.

8. Find the point nearest the origin on the intersection of the two surfaces

$$x^2 - xy + y^2 - z^2 = 1 \quad \text{and} \quad x^2 + y^2 = 1.$$

9. Find the maximum and minimum of the function $f(x, y, z) = xyz$:
(a) on the ball $x^2 + y^2 + z^2 \leqq 1$;
(b) on the plane triangle $x + y + z = 4$, $x \geqq 1$, $y \geqq 1$, $z \geqq 1$.

10. Find the maxima and minima of the function

$$(ax^2 + by^2)e^{-x^2-y^2}$$

if a, b are numbers with $0 < a < b$.

11. Let A, B, C denote the intercepts which the tangent plane at (x, y, z)

$$(x > 0, y > 0, z > 0)$$

on the ellipsoid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

makes on the coordinate axes. Find the point on the ellipsoid such that the following functions are a minimum:
(a) $A + B + C$.
(b) $\sqrt{A^2 + B^2 + C^2}$.

12. Find the maximum of the expression

$$\frac{x^2 + 6xy + 3y^2}{x^2 - xy + y^2}.$$

Because there are only two variables, the following method will work: let $y = tx$, and reduce the question to the single variable t.

Exercise 12 can be generalized to more variables, in which case the above method has to be replaced by a different conceptual approach, as follows.

13. Let A be a symmetric $n \times n$ matrix. Denote column vectors in $\mathbf{R}^n$ by X, Y, etc. For $X \in \mathbf{R}^n$ let $f(X) = \langle AX, X \rangle$, so f is a quadratic form. Prove that the maximum of f on the sphere of radius 1 is the largest eigenvalue of A.

Remark. If you know some linear algebra, you should know that the roots of the characteristic polynomial of A are precisely the eigenvalues of A.

14. Let C be a symmetric $n \times n$ matrix, and assume that $X \mapsto \langle CX, X \rangle$ defines a symmetric positive definite scalar product on $\mathbf{R}^n$. Such a matrix is called **positive definite**. From linear algebra, prove that there exists a symmetric positive definite matrix B such that for all $X \in \mathbf{R}^n$ we have

$$\langle CX, X \rangle = \langle BX, BX \rangle = \|BX\|^2.$$

Thus B is a square root of C, denoted by $C^{1/2}$. [*Hint*: The vector space $V = \mathbf{R}^n$ has a basis consisting of eigenvectors of C, so one can define the square root of C by the linear map operating diagonally by the square roots of the eigenvalues of C.]

15. Let A, C be symmetric $n \times n$ matrices, and assume that C is positive definite. Let $Q_A(X) = \langle AX, X \rangle$ and $Q_C(X) = \langle CX, X \rangle = \langle BX, BX \rangle$ with $B = C^{1/2}$. Let

$$f(X) = Q_A(X)/Q_C(X) \qquad \text{for} \quad X \neq 0.$$

Show that the maximum of f (for $X \neq 0$) is the maximal eigenvalue of $B^{-1}AB^{-1}$. [*Hint*: Change variables, write $X = BY$.]

16. Let a, b, c, e, f, g be real numbers. Show that the maximum value of the expression

$$\frac{ax^2 + 2bxy + cy^2}{ex^2 + 2fxy + gy^2} \qquad (eg - f^2 > 0)$$

is equal to the greater of the roots of the equation

$$(ac - b^2) - T(ag - 2bf + ec) + T^2(eg - f^2) = 0.$$

For Exercise 16, both methods, that of Exercise 12 and the one coming from quadratic forms in several variables, work. If you don't mind computations, check that they give the same answer.