linear transformations corresponds to multiplication of their matrices, we obtain

$$Dh(a) = Df(b)Dg(a), \text{ where } b = g(a) \quad \ldots (6)$$

This is called the matrix form of the chain rule. It can also be written as a set of scalar equations by each matrix in terms of its entries.

Suppose that $a \in \mathbb{R}^p$, $b = g(a) \in \mathbb{R}^n$, and $f(b) \in \mathbb{R}^m$. Then $h(a) \in \mathbb{R}^m$ and we can write

$$g = (g_1, g_2, \ldots, g_n),\ f = (f_1, f_2, \ldots, f_m),\ h = (h_1, h_2, \ldots, h_p)$$

Then $Dh(a)$ is an $m \times p$ matrix, $Df(b)$ is an $m \times n$ matrix and $Dg(a)$ is an $n \times p$ matrix, given by

$$Dh(a) = [D_j h_i(a)],\ i = 1, 2, \ldots m,\ j = 1, 2, \ldots, p$$

$$Df(b) = [D_k f_i(b)],\ i = 1, 2, \ldots, m,\ k = 1, 2, \ldots, n$$

and $$Dg(a) = [D_j g_k(a)],\ k = 1, 2, \ldots, n,\ j = 1, 2, \ldots, p$$

The matrix equation (6) is equivalent to mp scalar equations,

$$D_j h_i(a) = \sum_{k=1}^{n} D_k f_i(b) D_j g_k(a), \text{ for } i = 1, 2, \ldots, m \text{ and } j = 1, 2, \ldots, p.$$

These equations express the partial derivatives of the components of h in terms of the partial derivatives of the components of f and g.

Example 1. Extended chain rule for scalar fields.

Suppose f is a scalar field $(m = 1)$. Then

h is also a scalar field and there are p equations in the chain rule, one for each of the partial derivatives of h:

$$D_j h(a) = \sum_{k=1}^{n} D_k f(b) D_j g_k(a) \quad \text{for } j = 1, 2, \ldots, p$$

The special case $p=1$ was already considered in section 1.11. In This case we get only one equation, $h'(a) = \sum_{k=1}^{n} D_k f(b) g_k'(a)$.

Now take $p=2$ and $n=2$. Write $a = (s,t)$ and $b = (x,y)$. Then the components x and y are related to s and t by the equations

$$x = g_1(s,t), \quad y = g_2(s,t).$$

The chain rule gives a pair of equations for the partial derivatives of h:

$$D_1 h(s,t) = D_1 f(x,y) D_1 g_1(s,t) + D_2 f(x,y) D_1 g_2(s,t),$$

$$D_2 h(s,t) = D_1 f(x,y) D_2 g_1(s,t) + D_2 f(x,y) D_2 g_2(s,t).$$

In the notation, the pair of equations is usually written as

$$\frac{\partial h}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s}, \quad \frac{\partial h}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t}. \quad \cdots (7)$$

as shown in Theorem 1.10.4, page – 44.

Example 2 Polar coordinates. The temperature of a thin plate is described by a scalar field

f, the temperature at (x,y) being $f(x,y)$. Polar coordinates $x = r\cos\theta$, $y = r\sin\theta$ are introduced, and the temperature becomes a function of r and θ determined by the equation $\varphi(r,\theta) = f(r\cos\theta, r\sin\theta)$.
Express the partial derivatives $\frac{\partial\varphi}{\partial r}$ and $\frac{\partial\varphi}{\partial\theta}$ in terms of the partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$.

Solution: We use the chain rule as expressed in Equation (7), writing (r,θ) instead of (s,t) and φ instead of h. The equations

$$x = r\cos\theta, \quad y = r\sin\theta$$

give us $\frac{\partial x}{\partial r} = \cos\theta$, $\frac{\partial y}{\partial r} = \sin\theta$, $\frac{\partial x}{\partial\theta} = -r\sin\theta$, $\frac{\partial y}{\partial\theta} = r\cos\theta$

Substituting these formula in (7) we obtain

$$\frac{\partial\varphi}{\partial r} = \frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\sin\theta, \quad \frac{\partial\varphi}{\partial\theta} = -r\frac{\partial f}{\partial x}\sin\theta + r\frac{\partial f}{\partial y}\cos\theta, \quad \text{---}(8)$$

These are the required formulas for $\frac{\partial\varphi}{\partial r}$ and $\frac{\partial\varphi}{\partial\theta}$.

Example 3 Second order partial derivatives. Refer to Example 2 and express the second order partial derivative $\frac{\partial^2\varphi}{\partial\theta^2}$ in terms of the partial derivatives of f

Solution: We begin with the formula for $\frac{\partial\varphi}{\partial\theta}$ in (8) and differentiate with respect to θ

treating r as a constant. There are two terms on the right, each of which must be differentiated as a product.

Then we have

$$\frac{\partial^2 f}{\partial \theta^2} = -r\frac{\partial f}{\partial x}\frac{\partial(\sin\theta)}{\partial \theta} - r\sin\theta\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial x}\right) + r\frac{\partial f}{\partial y}\frac{\partial(\cos\theta)}{\partial\theta} + r\cos\theta\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial y}\right)$$

$$= -r\cos\theta\frac{\partial f}{\partial x} - r\sin\theta\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial x}\right) - r\sin\theta\frac{\partial f}{\partial y} + r\cos\theta\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial y}\right) \quad \cdots (9)$$

To compute the derivative of $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ with respect to θ we must keep in mind that, as function of r and θ, $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are composite functions given by

$$\frac{\partial f}{\partial x} = D_1 f(r\cos\theta, r\sin\theta) \text{ and } \frac{\partial f}{\partial y} = D_2 f(r\cos\theta, r\sin\theta)$$

Therefore, their derivatives with respect to θ must be determined by use of the chain rule.

We again use (7) with f replaced by $D_1 f$, to obtain

$$\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial(D_1 f)}{\partial x}\frac{\partial x}{\partial\theta} + \frac{\partial(D_1 f)}{\partial y}\frac{\partial y}{\partial\theta} = \frac{\partial^2 f}{\partial x^2}(-r\sin\theta) + \frac{\partial^2 f}{\partial y\partial x}(r\cos\theta).$$

Similarly, using (7) with f replaced by $D_2 f$, we find

$$\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial(D_2 f)}{\partial x}\frac{\partial x}{\partial\theta} + \frac{\partial(D_2 f)}{\partial y}\frac{\partial y}{\partial\theta} = \frac{\partial^2 f}{\partial x\partial y}(-r\sin\theta) + \frac{\partial^2 f}{\partial y^2}(r\cos\theta).$$

When ~~the~~ these formulas are used in (9) we obtain

$$\frac{\partial^2 \phi}{\partial\theta^2} = -r\cos\theta\frac{\partial f}{\partial x} + r^2\sin^2\theta\frac{\partial^2 f}{\partial x^2} - r^2\sin\theta\cos\theta\frac{\partial^2 f}{\partial y\partial x}$$
$$- r\sin\theta\frac{\partial f}{\partial y} - r^2\sin\theta\cos\theta\frac{\partial^2 f}{\partial x\partial y} + r^2\cos^2\theta\frac{\partial^2 f}{\partial y^2}.$$

This is the required formula $\frac{\partial^2\phi}{\partial\theta^2}$. Analogous for the second-order partial derivatives $\frac{\partial^2\phi}{\partial r^2}$, $\frac{\partial^2\phi}{\partial r\partial\theta}$ and $\frac{\partial^2\phi}{\partial\theta\partial r}$