row is $\nabla f_k(a)$, and where y is regarded as an $n \times 1$ column matrix. The matrix $Df(a)$ is called the Jacobian matrix of f at a. Its kjth entry is the partial derivative $D_j f_k(a)$. Then, we have

$$Df(a) = \begin{vmatrix} D_1 f_1(a) & D_2 f_1(a) & \cdots & D_n f_1(a) \\ D_1 f_2(a) & D_2 f_2(a) & \cdots & D_n f_2(a) \\ \vdots & \vdots & & \vdots \\ D_1 f_m(a) & D_2 f_m(a) & \cdots & D_n f_m(a) \end{vmatrix}$$

The Jacobian matrix $Df(a)$ is defined at each point where the mn partial derivatives $D_j f_k(a)$ exist. The total derivative ~~...~~ T_a is also written as $f'(a)$. The derivative $f'(a)$ is a linear transformation; the Jacobian $Df(a)$ is a matrix representation for this transformation. The first-order Taylor formula takes the form $$f(a+v) = f(a) + f'(a)(v) + \|v\| E(a,v), \quad \cdots (4)$$ where $E(a,v) \to 0$ as $v \to 0$. This resembles the one-dimensional Taylor formula. To compute the components of the vector $f'(a)(v)$ we can use the matrix product $Df(a)v$ or formula (3) of Theorem 1.13.1

1.14 Differentiability implies continuity

Theorem 1.14.1 If a vector field f is differentiable at a, then f is continuous at a.

Proof: As in the scalar case, we use the Taylor formula to prove this theorem. If we let $v \to 0$ in (4) the error term $\|v\| E(a, v) \to 0$. The linear part $f'(a)(v)$ also tends to 0 because linear transformations are continuous at 0. This completes the proof.

At this point it is convenient to derive an inequality which will be used in the proof of the chain rule in the next section. The inequality concerns a vector field f differentiable at a; it states that

$$\| f'(a)(v) \| \leq M_f(a) \|v\|, \text{ where } M_f(a) = \sum_{k=1}^{m} \|\nabla f_k(a)\|. \quad (5)$$

To prove this we use Equation (3) along with the triangle inequality and the Cauchy-Schwarz inequality to obtain

$$\| f'(a)(v) \| = \left\| \sum_{k=1}^{m} \nabla f_k(a) \cdot v \, e_k \right\| \leq \sum_{k=1}^{m} |\nabla f_k(a) \cdot v| \leq \sum_{k=1}^{m} \|\nabla f_k(a)\| \|v\|$$

1.15 The chain rule for the derivatives of vector fields.

Theorem 1.15.1 (CHAIN RULE) Let f and g be vector fields such that the composition $h = f \circ g$ is defined in a neighbourhood of a point a. Assume that g is differentiable at a, with total derivative $f'(b)$.

Then h is differentiable at a, and the total derivative

$h'(a)$ is given by

$$h'(a) = f'(b) \circ g'(a),$$

the composition of the linear transformations $f'(b)$ and $g'(a)$.

Proof: We consider the difference $h(a+y) - h(a)$ for small $\|y\|$, and show that we have a first-order Taylor formula. From the definition of h we have

$$h(a+y) - h(a) = f(g(a+y)) - f(g(a)) = f(b+v) - f(b), \quad \cdots (1)$$

where $v = g(a+y) - g(a)$. Taylor formula for $g(a+y)$ gives us

$$v = g'(a)(y) + \|y\| E_g(a,y), \text{ where } E_g(a,y) \to 0 \text{ as } y \to 0 \quad \cdots (2)$$

The ~~Eq~~ Taylor formula for $f(b+v)$ gives us

$$f(b+v) - f(b) = f'(b)(v) + \|v\| E_f(b,v) \quad \cdots (3)$$

where $E_f(b,v) \to 0$ as $v \to 0$. Using (2) in (3) we obtain

$$f(b+v) - f(b) = f'(b)\, g'(a)(y) + f'(b)\left(\|y\| E_g(a,y)\right) + \|v\| E_f(b,v)$$

$$= f'(b)\, g'(a)(y) + \|y\| E(a,y), \quad \cdots (4)$$

where $E(a,0) = 0$ and

$$E(a,y) = f'(b)\left(E_g(a,y)\right) + \frac{\|v\|}{\|y\|} E_f(b,v) \quad \text{if } y \neq 0. \cdots (5)$$

To complete the proof we need to show that

$$E(a,y) \to 0 \quad \text{as } y \to 0$$

The first term on the right of (5) tends to 0 as $y \to 0$ because $E_g(a,y) \to 0$ as $y \to 0$

and linear transformations are continuous at 0.
In the second term on the right of (5) ~~tends to~~
~~as y→0~~ the factor $E_f(b, v) \to 0$ because $v \to 0$
as $y \to 0$. The quotient $\|v\|/\|y\|$ remains bounded
because, by (2) and equation (5) of Theorem 1.14.1
we have

$$\|v\| \le M_g(a)\|y\| + \|y\|\,\|E_g(a,y)\|.$$

Therefore both terms on the right of (5) tends to
0 as $y \to 0$, so $E(a,y) \to 0$.
Thus from (4) and (1) we obtain the Taylor formula

$$h(a+y) - h(a) = f'(b)\,g'(a)(y) + \|y\|\,E(a,y),$$

where $E(a,y) \to 0$ as $y \to 0$. This proves that h
is differentiable at a and that the total derivative
$h'(a)$ is equal to the composition $f'(b) \circ g'(a)$.

<u>1.16</u> Matrix form of the chain rule
Let $h = f \circ g$, where g is differentiable at a
and f is differentiable at $b = g(a)$. The chain
rule states that $h'(a) = f'(b) \circ g'(a)$
We can express the chain rule in terms of the
Jacobian matrices $Dh(a)$, $Df(b)$ and $Dg(a)$ which
represent the linear transformations $h'(a)$, $f'(b)$
and $g'(a)$, respectively. Since composition of